

Yingmin Jia
Weicun Zhang
Yongling Fu *Editors*

Proceedings of 2020 Chinese Intelligent Systems Conference

Volume II

Lecture Notes in Electrical Engineering

Volume 706

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India
Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany
Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany

Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Università di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt
Torsten Kroeger, Stanford University, Stanford, CA, USA

Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA

Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore

Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany

Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA

Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany

Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand

Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan

Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy

Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institute of Telecommunications, Universität Stuttgart, Stuttgart, Germany

Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China

Junjie James Zhang, Charlotte, NC, USA

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Executive Editor (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, SCOPUS, MetaPress, Web of Science and Springerlink ****

More information about this series at <http://www.springer.com/series/7818>

Yingmin Jia · Weicun Zhang ·
Yongling Fu
Editors

Proceedings of 2020 Chinese Intelligent Systems Conference

Volume II



Springer

Editors

Yingmin Jia
School of Automation Science
and Electrical Engineering
Beihang University
Beijing, Beijing, China

Weicun Zhang
School of Automation
and Electrical Engineering
University of Science
and Technology Beijing
Beijing, Beijing, China

Yongling Fu
School of Mechanical Engineering
and Automation
Beihang University
Beijing, Beijing, China

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-15-8457-2

ISBN 978-981-15-8458-9 (eBook)

<https://doi.org/10.1007/978-981-15-8458-9>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Contents

Application of Real-Time Wavelet De-noising Based on Sliding Window in LADRC	1
Huayang Qin, Zengqiang Chen, Mingwei Sun, and Qinglin Sun	
Financial Topic Detection Algorithm Based on Multi-feature Fusion	11
Qiang Zhang, Junping Du, and Feifei Kou	
Research on Sentiment Analysis of Financial Text Based on Semantic Matching	20
Xunpu Yuan, Junping Du, and Feifei Kou	
A User Portrait Construction Method for Financial Credit	29
Chong Liu and Junping Du	
E-Product Recommendation Algorithm Based on Knowledge Graph and Collaborative Filtering	38
Jianhai Lv, Yawen Li, Junping Du, and Lei Shi	
Financial Sentiment Analysis Based on Pre-training and TextCNN	48
Xunpu Yuan, Yawen Li, Zhe Xue, and Feifei Kou	
Formation Terminal Sliding Mode Control for Multiple Mobile Robots	57
Lisen Gong and Wuxi Shi	
Mining Correlation Features of User Financial Behavior Based on Attention Mechanism and Dual Channel	66
Pengchao Cheng, Yawen Li, and Zhe Xue	
Current Control of Magneto-Rheological Clutch for Robotics Application	75
Jiangtao Ran, Guangzeng Chen, Chenguang Bai, and Yunjiang Lou	

Research on Open-pit Mine Virtual Environment Construction and Working Vehicle Modeling Simulation Based on PreScan	85
Yixin Guo, Guizhen Yu, Bin Zhou, Yucan Wang, and Guoqiang Liu	
Research on Hardware-in-the-Loop Test System for the Main Control Unit of Unmanned Vehicles	97
Haoyuan Ni, Guizhen Yu, Bin Zhou, and Guoqiang Liu	
Keywords Extraction Algorithm of Financial Review Based on Dirichlet Multinomial Model	107
Jinxuan Li, Yawen Li, and Zhe Xue	
Modelling of Wind Turbine with Battery Energy Storage System Based on Two-Level Switched System	117
Cheng Peng and Yang Song	
Integrated Marine Environment Observation System Based on Iridium Satellite Communication	126
Fei Yu and Jia Sun	
Intelligent Recognition of Bladder Cancer Based on Convolutional Neural Network	135
Xinguang Jin and Qi Zhang	
Distributed Consensus Backstepping Control in Networked Flexible Joint Manipulator Systems	143
Chengye Ying, Lixia Liu, Zhonghua Miao, and Jin Zhou	
Event-Triggered Stochastic Consensus for Networked Lagrangian Systems	152
Suying Pan, Zhiyong Ye, Lan Xiang, and Jin Zhou	
Autonomous Operation of Elevator Buttons for Multi-floor Navigation	161
Zhen Zhao, Junxin Zhao, and Yunjiang Lou	
Optimal Design of a Five-Bar Leg Mechanism for a Quadruped Robot	171
Chenguang Bai, Guangzeng Chen, Jiangtao Ran, and Yunjiang Lou	
Research on Fault Diagnosis Method of Train Wheelset Based on Deep Learning and Big Data Analysis	180
WanQing OuYang, Jian Zhang, Rongchao Fang, HengBo Jiang, Yujie Luo, and Xiaoyu Lu	
MEMS Gyroscope Noise Analysis and Calibration Using Allan Variance and Improved Hann Filter	193
Zhanyi Yan, Senchun Chai, Liu Yang, Baihai Zhang, and Lingguo Cui	

Monocular Dense 3D Reconstruction Algorithm Based on Inverse Depth Filter	203
Jingyun Duo and Long Zhao	
Multi-scale Superpixel Based Homomorphic Filtering for the Component Image Enhancement	212
Pingting Song, Jianlin Wang, Yongqi Guo, Liqiang Zhao, and Kepeng Qiu	
Multi-scale Visual Saliency Fusing with Spatiotemporal Features for Fire Detection	222
Huanyu Guo, Jianlin Wang, Yongqi Guo, Xinjie Zhou, and Kepeng Qiu	
Summary and Prospect of Predictive Control for a Class of Economic System	232
Yiwen Yang and Haotian Peng	
The Conjugate Gradient Algorithm for Control Systems with a Sine Excitation	240
Ling Xu	
Privacy-Preserving Consensus of Continuous Multi-agent Systems with Communication Delay	249
Yize Yang, Hongyong Yang, Meiyang Yu, and Yujiao Sun	
Output Feedback Stabilization for a Class of Nonlinear Systems with Unknown Output Function	257
Jiao Dai and Junyong Zhai	
Adaptive Control for Planar Nonlinear Systems with Input Quantization and Unknown Control Directions	266
Ce Liu and Junyong Zhai	
Numerical Modelling and Velocity Tracking Control for Autonomous Heavy-Duty Trucks	275
Tongtong Lei and Mingxing Li	
A Scheduling Problem of Joint Mobile Charging Sequence and Charging Start Time Control in Wireless Rechargeable Sensor Networks	284
Tao Cao, Chengpeng Jiang, and Wendong Xiao	
Model-Based Battery SOC Estimation Based on GA-UKF Algorithm	298
Zhijie Zhong and Junyong Zhai	
Study on the Vulnerability of Power Grid Cascade Failures Based on Complex Network Theory	307
Yirong Long and Chaoyang Chen	

Multiple Unmanned Armored Vehicle Formation Transform	316
Miqi Huang, Dixin Tian, and Xuting Duan	
Event-Triggered Control for Linear Multi-agent Systems Under DoS Attacks	326
Lianghui Sun, Jianting Lyu, Dai Gao, and Xin Wang	
A Celestial Navigation Approach Based on Star Pattern Recognition	336
Boning Wang, Wei Shao, Hanxue Zhao, Lingfei Dou, and Wenlong Yao	
Consensus Control of High-Dimensional Multi-agent Systems: A Large-Scale System Approach	346
Bin Zhang	
A New Method of Power Quality Cloud Evaluation Considering Dynamic Characteristics	353
Cheng Guo and Yan-chao Yin	
A Control Method of Fixed-Wing UAVs Under Close Formation	363
Yafei Zhao and Yanjie Zhao	
3D Path Planning for UAV with Improved Double Deep Q-Network	374
Liping Zhao, Yaofei Ma, and Jie Zou	
Fuzzy Adaptive Triangular Formation Control for a Class of Nonlinear System	384
Hao Wang, Qin Wang, Hong Shen, and Tianping Zhang	
Deep Learning Based Pathologic Images Recognition Upon Invasive Bladder Cancer	395
Tong Li and Qi Zhang	
An Autonomous Visual-Inertial-Based Navigation System for Quadrotor	404
Sansan Zhao, Rui Li, Yingjing Shi, and Haoran Li	
Unstructured Road Segmentation Method Based on Super Pixel and Region Growing Algorithm	415
Tao Liu, Xi Zhong, and Lixin Zhang	
Acoustical Field Modeling for Communication Through Steel Based on FDTD	424
Tao Liu, Jiajia Liu, Zongmei Bai, and Ouming Liu	
Defect Sample Generation System Based on DCGAN for Glass Package Electrical Connectors	434
Qunpo Liu, Mengke Wang, and Naohiko Hanajima	

A Review of Fault Features for Inter-turn Short Circuit Fault of PM Motor	442
Liuyang Shen, Ziyi Fu, and Sumin Han	
Development of On-Line Monitoring System for Cremation Machine Based on PXI Bus	452
Lin Tian, Fengguang Huang, Feng Shi, and Bingjie Li	
Generating Synthesized CT from Cone-Beam Computed Tomography (CBCT) Using Artifact Disentanglement Network for Image-Guided Radiotherapy (IGRT)	460
Hanlin Cheng, Jiwei Liu, Jianfei Liu, Ronghu Mao, and Pengjian Sun	
Design and Application of Wireless Temperature Monitoring System for Diesel Locomotive in Reconditioning Field Based on ZigBee Network	470
Wenlong Yao, Zhen Pang, Keyu Zhuang, and Wei Shao	
IMC-PID Controller Based on Extended State Observer for Second-Order Delayed Unstable Processes with Two Unstable Poles	479
Yachao Liu, Jian Gao, and Lanyu Zhang	
Fast Load Control of Supercritical Unit Through ADRC Optimized by Cross Quantum Pigeon	487
Guolian Hou, Ting Huang, Huilin Su, Linjuan Gong, Congzhi Huang, Bo Hu, and Bing Wang	
Finite-Time Containment Consensus Control of Second-Order Multi-agent Systems	496
Yan Cui and Qi Xue	
Track Tracking Control of Unmanned Intelligent Sweeping Vehicles Based on Improved MFAC	506
Wenlong Yao, Zhen Pang, Ronghu Chi, and Wei Shao	
Rapid Trajectory Design for Air-Launched High Speed Near Space Vehicle	516
Jianhui Liu, Mingang Zhang, Dongfei Hu, and Ningning Zhang	
Advanced Simulation of Electric Trimmable Horizontal Stabilizer Actuation System Based on AMESim	524
Wei Li, Xudong Han, Wensen Zhang, Jian Fu, and Liming Yu	
Characteristic Analysis of Mechatronic Equipment of Flight Control System for More Electric Aircraft	533
Tuanhui Guo, Kunxu Cui, Jian Fu, and Yongling Fu	
Modeling and Simulation of Redundancy Management of Electro-Hydrostatic Actuator	542
Jixin Song, Liming Yu, Xudong Han, and Jian Fu	

Global Exponential Rendezvous Control of Nonholonomic Unicycle Vehicles with Directed Communication Topology	552
Yixin Zhao, Bin Xing, Wei Huang, Shiyu Luo, Peng Li, and Wenjing Xie	
Spectral Criterion for Stability of Mean-Field Stochastic Periodic Systems.....	562
Yuechen Cui and Hongji Ma	
Research on Sliding Mode Control of Multi-differential Omnidirectional Mobile Platform.....	571
Huimin Zhang, Yue Ma, and Koanhee Cho	
Object Tracking Algorithm of UAV Based on Fast Kernel Correlation Filter	584
Xukuai Liu, Mingjian Sun, and Lingbo Meng	
Develop Trajectory Tracking Controller for 4WS Vehicle	594
Koanhee Cho, Yue Ma, and Huimin Zhang	
Design of Fatigue Driving Detection Algorithm Based on Image Processing.....	602
Shangzheng Liu, Yalei Wu, Qunpo Liu, and Qinghui Zhu	
Optimal Design for Active Vibration Control of Aero-engines	611
Ting Fang and Jiqiang Wang	
Stabilization of Linear Systems by a Novel Event-Triggered Control with Time-Varying Threshold	621
Xinyu Jiang and Fei Hao	
Modelling of SFR for Wind-Thermal Power Systems via Improved RBF Neural Networks	630
Jianhua Zhang, Hongrui Li, Bo Hu, Yong Min, Qun Chen, Guolian Hou, and Congzhi Huang	
An Efficient Action Recognition Framework Based on ELM and 3D CNN	641
Yiping Zou and Xuemei Ren	
Dynamic Modelling and Co-simulation for Active Suspension Systems	649
Huidong Hou, Shichang Han, Yingbo Huang, and Jing Na	
Design of Terminal Sliding Mode Adaptive Controller for Multiple Mobile Manipulators with Flexible Joints	657
Baigeng Wang and Shurong Li	
CT Tracing Algorithm Based on Optical Flow Field	667
Lei Xiao, Minghai Xu, and Zhongyi Hu	

Photovoltaic Power Forecasting Based on Randomized Multi-scale Kernels	674
Yang Deng, Yeqiong Liu, and Xuemei Dong	
Environmental Sound Recognition Based on Residual Network and Stacking Algorithm	682
Haoyuan Wang, Xuemei Ren, and Zhen Zhao	
Robust Online Filter Based on a Second-Order Adaptive Model	691
Shenglun Yi and Xuemei Ren	
Extended-State-Observer-Based Funnel Control for Motor Servo System with Quantized Input	699
Yun Cheng and Xuemei Ren	
Control and Application of Tail-Like Mechanism in Self-balance Robot	709
Lufeng Zhang, Xuemei Ren, and Yun Cheng	
Intermittent Fault Detection and Isolation for Discrete System with Unknown Disturbance Based on Interval Observer	717
Hao Li, Jianfeng Qu, Xiaoyu Fang, Jinzhuo Liu, and Hongpeng Yin	
Research on Simulation Technology of Cooperative Guidance for Aircraft Formation	725
Jiahui Tong, Jin Liu, Jing Li, and Tianming Zhan	
On Fatigue Driving Detection System Based on Deep Learning	734
Yasheng Yuan, Fengzhi Dai, Yunzhong Song, and Jichao Zhao	
On Improvement of Target Recognition eCCA Method Based on SSVEP	742
Di Yin, Fengzhi Dai, Yunzhong Song, Jichao Zhao, and Dejin Wang	
Review of Relative Navigation for Noncooperative Spacecraft in Close Range	750
Dayi Wang, Qiyang Hu, Wenbo Li, Haidong Hu, and Kebei Zhang	
Anomaly Detection of Satellite Momentum Wheel Based on Fast Dynamic Time Warping	773
Pengjie Cheng, Jianshen Li, Yan Hu, Na Li, Zifeng Qiu, and Shuangxin Wang	
Reliability Analysis of Flexible Test System for the Products of Aerospace Pyromechanical Device Based on Dynamic Bayesian Network	784
Jianxiong Zhou, Shanbi Wei, and Yi Chai	
Research and Development of Vision System of SMT Machine	793
Xujie Hou, Yonghui Xie, and Lei Pan	

Air Pollutants NO₂ Concentration Prediction Based on LSTM Neural Network method	801
Jihan Li, Xiaoli Li, Jian Liu, and Kang Wang	
Fractional Integral Sliding Mode Control for Trajectory Tracking of Baxter Robot Manipulators	809
Xiru Wu and Jingtong Jia	
Fast Detection of Passion Fruit with Multi-class Based on YOLOv3	818
Xiru Wu and Rongchai Tang	
Author Index	827



Application of Real-Time Wavelet De-noising Based on Sliding Window in LADRC

Huayang Qin^{1,2}, Zengqiang Chen^{1,2(✉)}, Mingwei Sun^{1,2}, and Qinglin Sun^{1,2}

¹ College of Artificial Intelligence, Nankai University, Tianjin 300350, China
chenzq@nankai.edu.cn

² Tianjin Key Laboratory of Intelligent Robotics, Nankai University,
Tianjin 300350, China

Abstract. The paper applies wavelet de-noising in linear active disturbance rejection control (LADRC). To decrease the noise in real time, the method of sliding window is designed. When the frequency and intensity of disturbance noise in the system are large, the control effect of LADRC will be reduced greatly. To solve that, the real-time de-noising algorithm is concatenated between the noisy output and the linear extended state observer (LESO). The results of simulations show that, in LADRC, the applied real-time de-noising algorithm is effective.

Keywords: Real-time wavelet de-noising · Linear active disturbance rejection control (LADRC) · Linear extended state observer (LESO)

1 Introduction

As the rapid development of various advanced control technologies, a system requires more and more control accuracy. However, any system is subject to both disturbance and noise. For example, in almost all applications, both desired and noisy signals are present in the observed data. There are many reasons for the noise: it can be generated from adjacent sources, nature, human beings, or the process of observation and so on [1]. There are many kinds of noise, such as gaussian white noise and harmonic noise. The existence of noise will reduce the control accuracy of the system and affect the control effect of the system.

In practical application, the core problem of control should be to suppress the interference and uncertainty in the system, that is, how to suppress the disturbance [2]. In order to solve such problems, ADRC came into being [3], which is to eliminate the disturbance information from the signal of the controlled object with the control signal as soon as possible, so as to greatly reduce disturbance's influence on the system. To make parameters adjustment easily, ADRC is simplified to LADRC [4], which promotes the practical application of ADRC [5–8].

ADRC is to suppress disturbance and noise from the angle of control algorithm. However, when the frequency and intensity of disturbance noise in the

system are large, the control effect of LADRC will be greatly reduced. To solve this problem, we can choose an appropriate mathematical model for the noise from the point of view of analog filtering [9] and adopt an effective method to minimize the negative impact of noise on the system, so as to greatly improve the performance of the control system [10]. In recent years, with the development of wavelet theory, signal de-noising by wavelet transform has been widely used in image processing, fault diagnosis, satellite navigation and stock market analysis. Nonetheless, because of its non-causality and non-translational invariance, it is difficult for wavelet transform to do recursive calculation, which influences the real-time application of wavelet de-noising. The most current wavelet de-noising method is based on the off-line batch processing of measurement data, to effectively reduce noise. It is little to see that wavelet de-noising is applied to control system's real-time noise reduction [11]. In this paper, inspired by the research of Sun Pengyue in [12], the real-time wavelet de-noising based on sliding window is applied in LADRC, to improve the control effect of the systems contain high frequency noise.

2 Algorithm Description

2.1 Wavelet De-noising

The simplified model of noise signal is as follows:

$$f(k) = s(k) + n(k) \quad (1)$$

where time k is equally spaced, $s(k)$ is the original signal, and $n(k)$ is a noise.

The goal of noise reduction is to reduce the degree of contamination to $s(k)$ by the noise $n(k)$.

After wavelet transform or decomposition, discrete signals show different characteristics at different resolutions, so different thresholds and wavelet coefficients can be used to reduce noise [13]. Therefore, $f \in L^2(R)$ (square-integral real number vector space) is a finite power function. The wavelets multi-resolution analysis tells, we can express f as a series expansion of scale function and wavelet in the given vector space V_J :

$$f = \sum_k a_{J,k} \phi_{J,k} = \sum_k a_{J-1,k} \phi_{J-1,k} + \sum_k b_{J-1,k} \psi_{J-1,k} \quad (2)$$

where a is the scale coefficients, b is the wavelet coefficient, ϕ represents the scaling function, ψ represents the wavelet function, j, k, l belong to the set of integers. Consider the independence of k , we can assume $f = \sum_l a_{J,l} \phi_{J,l}$, then

$$a_{J-1,k} = \langle \sum_l a_{J,l} \phi_{J,l}, \phi_{J-1,k} \rangle = \sum_l \langle \phi_{J,l}, \phi_{J-1,k} \rangle a_{J,l} \quad (3)$$

where

$$\langle \phi_{J,l}, \phi_{J-1,k} \rangle = \int_R 2^{\frac{J}{2}} \phi(2^J x - l) \cdot 2^{\frac{J-1}{2}} \phi(2^{J-1} x - k) dx = h_{l-2k} \quad (4)$$

and we can get

$$a_{J-1,k} = \sum_l h_{l-2k} a_{J,l}$$

similarly, we get

$$b_{J-1,k} = \langle \sum_l a_{J,l} \phi_{J,l}, \psi_{J-1,k} \rangle = \sum_l \langle \phi_{J,l}, \psi_{J-1,k} \rangle a_{J,l} = \sum_l g_{l-2k} a_{J,l} \quad (5)$$

There are three main steps for wavelet denoising:

1. Determine the wavelet basis type and decomposition level, then perform wavelet transform and obtain a set of wavelet coefficients $b_{j,k}$ of the noisy signal $f(k)$, j is the level of the wavelet decomposition.
2. Apply the threshold disposing to $b_{j,k}$, and get the estimated wavelet coefficients $\hat{b}_{j,k}$, so as to keep $\|\hat{b}_{j,k} - u_{j,k}\|$ as small as possible, $u_{j,k}$ is the wavelet coefficients of $s(k)$.
3. Use $\hat{b}_{j,k}$ to do wavelet reconstruction, then get the estimated signal $\hat{f}(k)$, which is the de-noised signal.

2.2 Real-Time De-noising Based on Sliding Window

Because of the non-causality and non-translational invariance of wavelet transform, wavelet de-noising cannot achieve the recursive de-noising of time series, which limits the application of wavelet de-noising in real-time systems. For the noise reduction in LADRC, the noisy output need to add sliding window to realize the recursive calculation.

Assuming that the data length required for de-noising is L_x , the sliding window design is as

$$W_i = \begin{cases} \emptyset & i < L_x \\ \{x(i-L_x+1), \dots, x(i)\} & i \geq L_x \end{cases} \quad (6)$$

where L_x is the width of window, $x(i)$ is the noisy output. Equation (6) indicates that when the length of output is less than the width of the window, no de-noising operation is performed. When the length is greater than L_x , always using the latest L_x output for de-noising, and outputting the wavelet de-noising result $\hat{x}(L_x)$.

However, the above real-time noise reduction algorithm always takes the data at the right end of the window as the latest denoised output, it will appear boundary effect and false Gibbs phenomenon [14], so the filtering effect after adding a window is not ideal. In order to solve this problem, window extension can be used to reduce the signal distortion caused by boundary effect. The most commonly used method is symmetric window extension, as follows

$$W_i = \begin{cases} \emptyset & i < L_x \\ \{x(i-L_x+1), \dots, x(i), x(i), \dots, x(i-L_x+1)\} & i \geq L_x \end{cases} \quad (7)$$

It can be seen that the length of the window extension is the same as the length of the window, while the extended data is still the latest L_x output, but in reverse order. This design can ensure the continuity of the original data before de-noising. Because the de-noised value is located in the middle of the window, it can effectively reduce the influence of the boundary effect. However, the data length of the above window design has been doubled, and the computing amount will increase, so the real-time performance cannot be satisfied. In view of this, we adopt the following window extension design, which can further reduce the computation on the basis of refraining from the boundary effect. The window design is as follows

$$W_i = \begin{cases} \emptyset & i < L_x \\ \{x(i - L_x + 1), \dots, x(i), x(i), \dots, x(i - L_e + 1)\} & i \geq L_x \end{cases} \quad (8)$$

where L_e is the length of window extension. Compared with the $2L_x$ symmetric window design shown in Eq. (7), the total width of the above window is $L_w = L_x + L_e$, as long as L_e is selected properly, it can effectively reduce the amount of computation.

The wavelet de-noising based on sliding window mentioned above can be used to complete the de-noising of the noisy output in LADRC. The algorithm process is shown in Fig. 1.

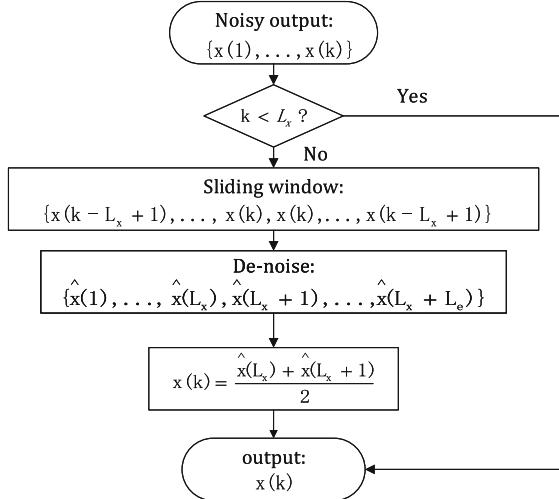


Fig. 1. Real-time wavelet de-noising based on extended sliding window

Firstly, when the length of the noisy output is less than L_x , the algorithm is in initialization, and no de-noising operation is performed. Secondly, when the length reaches L_x , the output sequence is extended through the sliding window

design shown in Eq. (8). Then, according to the description in Sect. 2.1, the real-time de-noising of the noisy output is completed through the three steps: Wavelet decomposition, threshold noise reduction and wavelet reconstruction. Thirdly, outputting the wavelet de-noising result $[\hat{x}(L_x) + \hat{x}(L_x + 1)]/2$ at the current moment, and the noisy output at the next moment is processed as above. Finally, the de-noising of the noisy output in LADRC can be completed.

3 Control Scheme

Consider a second order plant [15]

$$\ddot{y} = -1.41\dot{y} + 23.2T_d + (23.2 - 40)u + 40u = w + 40u \quad (9)$$

$$w = -1.41\dot{y} + 23.2T_d + (23.2 - 40)u$$

where w is referred to as the generalized disturbance, y is the output, u is the control signal, and T_d is a torque disturbance. The LESO is

$$\dot{z} = \begin{bmatrix} -3\omega_o & 1 & 0 \\ -3\omega_o^2 & 0 & 1 \\ -\omega_o^3 & 0 & 0 \end{bmatrix} z + \begin{bmatrix} 0 & 3\omega_o \\ 40 & 3\omega_o^2 \\ 0 & \omega_o^3 \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} \quad (10)$$

where ω_o is the bandwidth of the LESO, and

$$\begin{aligned} z_1 &\rightarrow y \\ z_2 &\rightarrow \dot{y} \\ z_3 &\rightarrow w, \text{ as } t \rightarrow \infty \end{aligned} \quad (11)$$

define the controller as

$$\begin{aligned} u &= \frac{u_0 - z_3}{40} \\ u_0 &= k_p(r - z_1) - k_d z_2 \end{aligned} \quad (12)$$

$$\begin{aligned} k_d &= 2\xi\omega_c, \quad \xi = 1 \\ k_p &= \omega_c^2 \end{aligned} \quad (13)$$

where ξ is the damping ratio and ω_c is the bandwidth of the state feedback system.

The above is LADRC. And LADRC with de-noising is designed by concatenating the real-time wavelet de-noising based on extended sliding window, which is described in Sect. 2.2, between the noisy output and LESO, as Fig. 2, where b_0 is the gain of plant. The stability analysis of LADRC is given in [15].

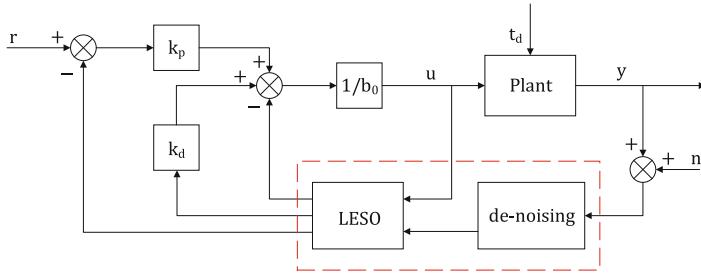


Fig. 2. Diagram of LADRC with de-noising

4 Simulation Results

In simulations, using the solver ode1, the fundamental sample time is 1 ms, add a noise $n(t)$ of 1 ms sampling period. Add a step torque disturbance of -1.16 at $t = 3$ s. Select $\omega_o = 40\text{rad/s}$, $\omega_c = 20\text{rad/s}$ and $b_0 = 40$.

To test the performance of the LADRC(with de-noising), we use a trapezoidal transient profile with a settling time of 1 s. Set the noise to the sinusoidal noise $n(t) = 0.1\sin(500\pi t)$ and the white noise whose power(power spectral density, PSD) is 10^{-5} , respectively. The simulation results are displayed from Figs. 3, 4, 5, 6, 7 and 8.

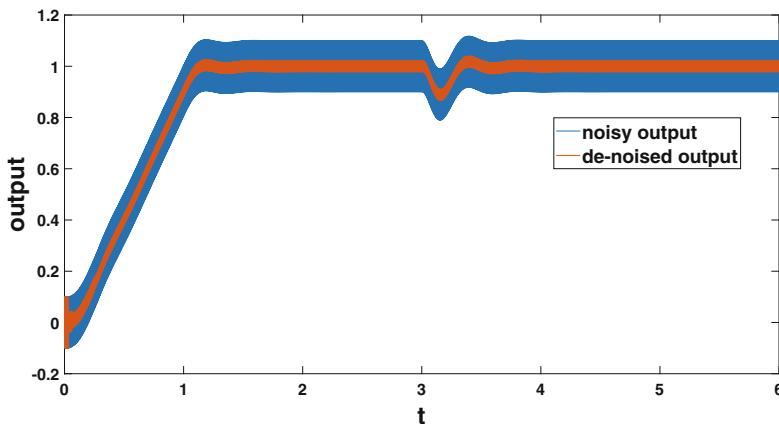


Fig. 3. Noisy output and de-noised output($n(t) = 0.1\sin(500\pi t)$)

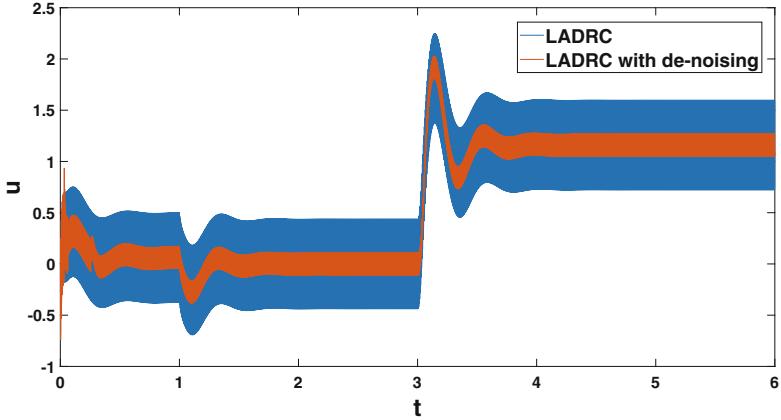


Fig. 4. Control signals of LADRC(with de-noising, $n(t) = 0.1\sin(500\pi t)$)

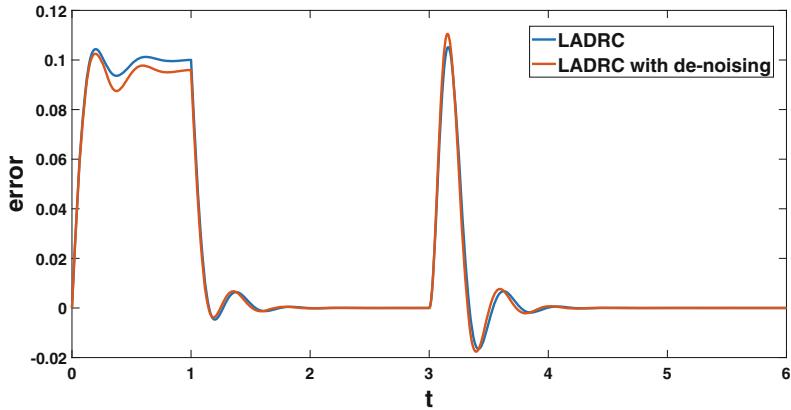


Fig. 5. Error of LADRC(with de-noising, $n(t) = 0.1\sin(500\pi t)$)

From Fig. 3 and Fig. 6, we can see the real-time wavelet de-noising algorithm have great performance, which can filter out about 70% noise. Figure 4 and Fig. 7 indicate LADRC with de-noising is of smaller control signals, which decreases controlled quantify by half, on average. Figure 5 and Fig. 8 show that the error of the two methods are similar.

Taken together, the results indicate the effectiveness and accuracy of the applied real-time wavelet de-noising algorithm in LADRC.

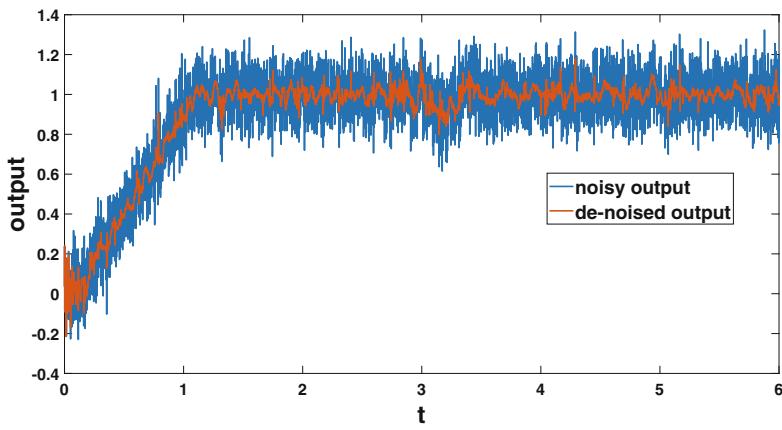


Fig. 6. Noisy output and de-noised output (white noise, $PSD = 10^{-5}$)

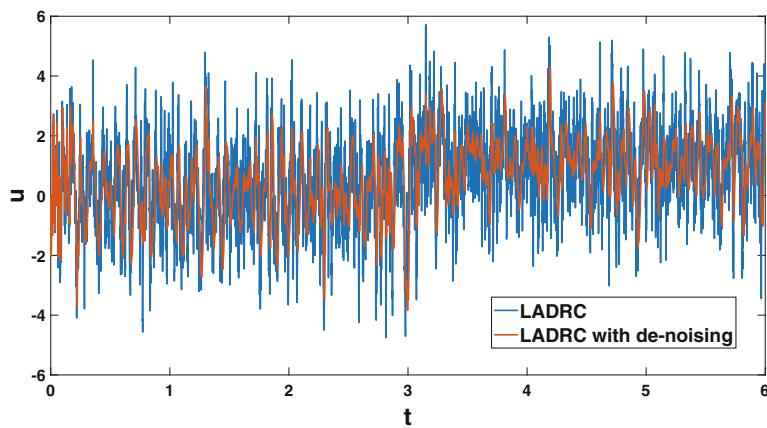


Fig. 7. Control signals of LADRC (with de-noising, white noise, $PSD = 10^{-5}$)

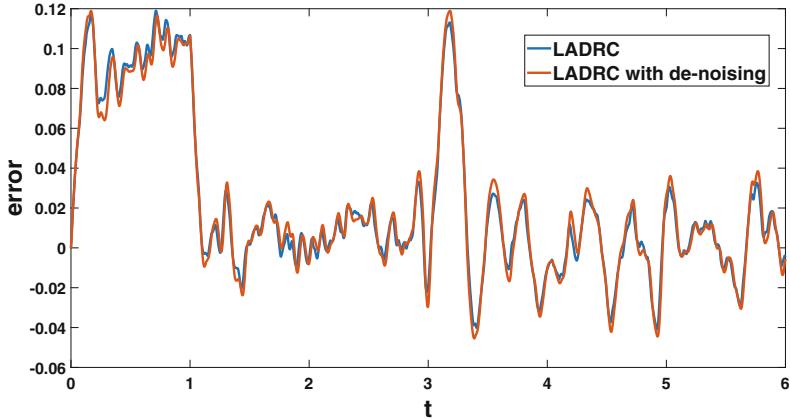


Fig. 8. Error of LADRC(with de-noising, white noise, $PSD = 10^{-5}$)

5 Conclusions

This paper successfully applies the real-time wavelet de-noising to LADRC. Taking into account that the control effect of LADRC will be greatly reduced when the frequency and intensity of the noise in the system are large, LADRC with denoising is utilized instead. Numerical simulation results show the proposed method is effective, especially in filtering out noise and decreasing the controlled quantify. Although the simulation is about a second order system, it also works for higher order systems: concatenate the real-time de-noising algorithm between the noisy output and LESO.

Acknowledgments. This work was supported partly by the Natural Science Foundation of China under grants 61973175, 61973172, and the Key Technologies R&D Program of Tianjin Grant No. 19JCZDJC32800.

References

1. Li, S.: New algorithm and performance analysis of communication signal processing under stable distributed noise. Dalian University of Technology (2011)
2. Gao, Z.: Research on thought of ADRC. Control Theory Appl. **30**(12), 1498–1510 (2013)
3. Han, J.Q.: From PID to active disturbance rejection control. IEEE Trans. Ind. Electron. **56**(3), 900–906 (2009). <https://doi.org/10.1109/TIE.2008.2011621>
4. Gao, Z.Q.: On the centrality of disturbance rejection in automatic control. ISA Trans. **53**(4), 850–857 (2014). <https://doi.org/10.1016/j.isatra.2013.09.012>
5. Sun, M.W., Wang, Z.H., Wang, Y.K., et al.: On low-velocity compensation of brushless DC servo in the absence of friction model. IEEE Trans. Ind. Electron. **60**(9), 3897 (2013)

6. Qiu, D.M., Sun, M.W., Wang, Z.H., et al.: Practical wind-disturbance rejection for large deep space observatory antenna. *IEEE Trans. Control Syst. Technol.* **22**(5), 1983 (2014)
7. Dong, J.Y., Sun, L., Li, D.H.: Linear active disturbance rejection control for ball mill coal-pulverizing systems. *Chin. J. Eng.* **37**(4), 509 (2015)
8. Liu, J.J., Chen, Z.Q., Sun, M.W., Sun, Q.L.: Application of active disturbance rejection control in high-angle-of-attack maneuver for aircraft with thrust vector. *Chinese J. Eng.* **41**(9), 1187–1193 (2019). <https://doi.org/10.13374/j.issn2095-9389.2019.09.010>
9. Wang, R., Li, Y., Sun, H., Chen, Z.: Integrated cabin contaminant monitoring network based on Kalman Consensus Filter. *ISA Trans.* **71**, 112–120 (2017)
10. Wang, R., Sun, H., Ma, Z.: Stability and performance analysis of a jump linear control system subject to digital upsets. *Chin. Phys. B* **24**(4), 040201 (2015)
11. Jiang, D., Chen, M.: A real-time wavelet de-noising algorithm. *J. Instrum.* (06), 781–783 + 795 (2004)
12. Sun, P.: Research on the key technology of anti-ionospheric scintillation carrier tracking of satellite navigation signal. National University of Defense Science and Technology (2017)
13. Zhou, Z.R., et al.: Improvement of the signal to noise ratio of Lidar echo signal based on wavelet de-noising technique. *Opt. Lasers Eng.* **51**(8) (2013)
14. Li, Q., Wang, X., Song, Y., Wei, F.: Image translation invariant wavelet denoising method and its application. *J. Hohai Univ. Changzhou* (04), 17–20 (2006)
15. Gao, Z.: Scaling and bandwidth-parameterization based controller tuning. In: Proceedings of the 2003 American Control Conference, Denver, CO, USA, pp. 4989–4996 (2003). <https://doi.org/10.1109/ACC.2003.1242516>



Financial Topic Detection Algorithm Based on Multi-feature Fusion

Qiang Zhang, Junping Du^(✉), and Feifei Kou

Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia,
School of Computer Science, Beijing University of Posts and Telecommunications,
Beijing 100876, China
junpingdu@126.com

Abstract. In order to help investors get effective information from a large amount of financial news data as soon as possible, a topic detection model based on multi-view text semantics and clustering topic detection algorithm is proposed. In the financial news data set, different models are used to extract the characteristics of the news, and the characteristics of various models are merged. The clustering algorithm is improved by introducing JS divergence and time decay factors. The experimental results show that compared with the traditional topic detection model, the proposed method has higher accuracy of topic detection and shorter runtime of clustering algorithm.

Keywords: Financial news · Topic detection · Text representation · Text clustering

1 Introduction

The rapid development of Internet media has expanded the audience of financial news, making the relationship between the securities market and financial news more and more closely. Studies have found that stock returns and transactions in the stock market are directly related to the number of Dow Jones announcements per day. Highly reported stocks have lower returns than those that have received little media attention. From this we can see that stock trends are more susceptible to financial news. How to extract effective knowledge from a large number of financial news has become a problem. For this reason, some scholars have proposed topic detection and tracking (TDT).

Topic detection is one of the most important tasks of TDT. Its main goal is to detect topics described in the data stream and classify the text describing the same topic into one category. In the field of topic detection, the number of topics in the data to be detected is often unknown. K-means clustering and other methods require K values to be determined in advance, and the clustering results are often sensitive to K values. Therefore, traditional methods often cluster by hierarchical clustering and set a threshold to end the clustering. By setting the

threshold method, although a high-precision result can be obtained, there may still be some topics underfitting because the internal similarities between different topics are not the same; if the threshold is lowered, it may cause other topics to overfit. Aiming at the problems in the above research, this paper proposes a topic detection algorithm based on text semantics and clustering, which will detect the same topic with a large time interval, and improve the traditional single-pass clustering method to make the cluster. The class process is more balanced, which improves the accuracy of topic detection.

2 Related Work

Topic detection and tracking (topic detection and tracking, TDT) is mainly to detect unknown topics in the system. Since the introduction of TDT, researchers have proposed various methods. On the whole, these methods are mainly aimed at the topic model and topic detection methods.

Chen et al. [1] Proposed a knowledge-based semi-supervised hierarchical online topic detection framework. This framework can detect topics in a hierarchical manner online. Wang et al. [2] Proposed a topic detection model based on multi-dimensional public opinion network. Based on the actual process of generating and disseminating multiple topics related to the same topic, a multi-dimensional network model oriented to the topology of self-media network public opinion was constructed. Zhu et al. [3] Proposed an TF-IDF algorithm based on time distribution and user attention, using a new method of segmentation and combination of words to extract hot topics based on hot items. Li [4] and others proposed a distributed, incremental Weibo topic model BBE +, which uses the distributed framework Spark to process data streams in real time, processes the streams in an incremental manner, and tracks topic changes. Zhang et al. [5] Proposed a four-stage framework to improve the performance of topic detection, using deep learning to enrich short text information through image understanding, and then, using improved potential Dirichlet assignment to optimize effective word pairs to improve extraction The accuracy of the keywords.

Zhao [6] and others proposed a background-based topic detection model to improve the performance of single-domain corpus topic detection. First, in order to improve accuracy, the distribution of words and the relationship between words in the entire corpus are used as background knowledge. Secondly, a dynamic probability model based on probability gain is introduced to detect potential topics by learning a model based on dynamic human memory cognitive processes. Zhang [7] and others proposed an improved TopicSketch model, which uses a new topic quality measurement method to estimate the word intrusion and topic consistency based on two or two points of mutual information between topic words. Zhong [8] and others proposed a Top-k burst topic detection method combined with regional characteristics, using four types of indicators: word appearance frequency, word associated users, word distribution area, and word social behavior, and proposed a novel Weibo network. Word burst calculation model. Zhang [9] and others proposed a similarity-based topic detection and analysis system (PTDAS), which uses the FP growth algorithm to mine the

microblog with the smallest cosine similarity from a microblog data set, and then summarizes it as the topic. At the same time, in order to discover topics in real time, parallelize algorithms on Spark for efficient mining.

3 Topic Detection Algorithm Based on Text Semantics and Clustering

This paper proposes a topic detection algorithm based on multi-view feature fusion and improved single-pass clustering. The framework of the proposed algorithm is shown in Fig. 1.

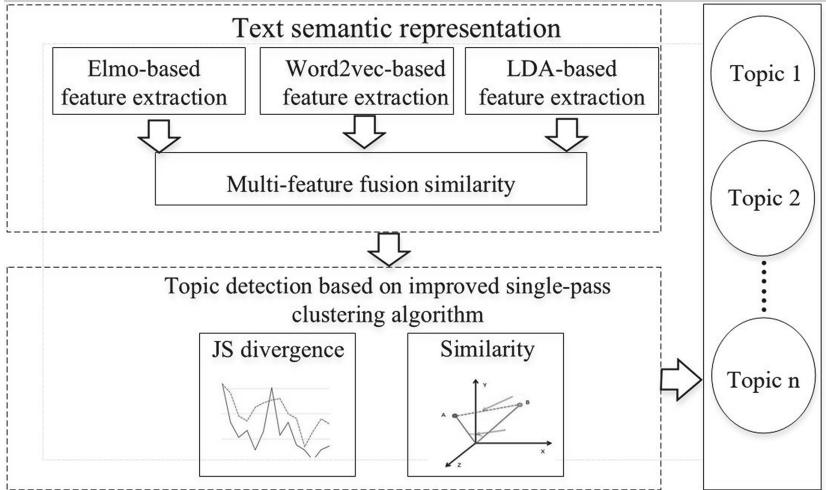


Fig. 1. The framework of the proposed topic detection method.

The algorithm consists of two parts: the acquisition of multi-perspective feature fusion semantic representation and the clustering process based on time decay. The acquisition of semantic representations is constructed by the fusion of multiple semantic representations of text. Improved single-pass clustering can obtain more accurate clustering results by adding factors that cause topics to decay over time in the clustering process.

3.1 Text Semantic Representation Based on Multi-view Feature Fusion

If text data cannot be directly input to a computer for processing, structured processing is required and the text is represented using vectors. In order to better represent text, we extract text features from multiple perspectives and build a vector model.

For each piece of financial news, three models are established to calculate the similarity between news text and topic. According to the three models established, the similarity between text and topic is calculated respectively, and the three similarities are combined to obtain the optimal similarity. The multi-view feature fusion formula is shown in Eq. (1).

$$\begin{aligned} \text{similarity}(v, t) = & \lambda_1 \text{similarity}_{\text{Elmo}}(v, t) \\ & + \lambda_2 \text{similarity}_{\text{word2vec}}(v, t) + \lambda_3 \text{similarity}_{\text{lda}}(v, t) \end{aligned} \quad (1)$$

Where $\text{similarity}_{\text{Elmo}}(v, t)$ is the similarity of news text v and topic t based on the Elmo model, $\text{similarity}_{\text{word2vec}}(v, t)$ is the similarity of news text v and topic t based on the word2vec model, $\text{similarity}_{\text{lda}}(v, t)$ is the similarity of news text v and topic t based on the LDA model, and $\lambda_1, \lambda_2, \lambda_3$ is the multi-view feature fusion coefficient.

The Elmo model is a new type of deep contextualized word representation, which can model the complex features of words and the changes of words in the language context. Our word vector is a function of the internal state of the deep bidirectional language model (biLM), pre-trained on a large text corpus.

Elmo is a bidirectional LSTM model, and a forward and backward language model. The objective function is to take the maximum likelihood of the two-direction language model. Among them, the forward model is shown in Eq. (2):

$$p(t_1, t_2 \dots t_N) = \prod_{k=1}^N p(t_k | t_1, t_2 \dots t_{k-1}) \quad (2)$$

The backward model is shown in Eq. (3):

$$p(t_1, t_2 \dots t_N) = \prod_{k=1}^N p(t_k | t_{k+1}, t_{k+2} \dots t_N) \quad (3)$$

The maximum likelihood function is shown in Eq. (4):

$$\sum_{k=1}^N (\log(t_k | t_1, t_2 \dots t_{k-1}) + \log(t_k | t_{k+1}, t_{k+2} \dots t_N)) \quad (4)$$

3.2 Clustering Algorithm Based on Time Decay

Single-pass is an incremental clustering algorithm for processing stream data, which is often used for topic detection and tracking. For arriving text streams, each text will be vectorized in the input order. Text is divided based on the similarity between the current text and the existing topic categories. Merge existing text into a high similarity topic class, or create a new topic class to achieve incremental and dynamic clustering of data. This article uses JS divergence to measure the similarity between topics. Use a time decay function to improve the similarity between topics.

When we calculate the similarity of topic tracking news topics, we think that the closer the two topics are, the more similar they should be. We introduce time as a parameter in the calculation of topic similarity. The larger the time span, the less likely the similarity is. We introduce the following time decay function to calculate the effect of time span on topic similarity. As shown in Eq. (5):

$$H(\Delta t) = e^{(-\Delta t/L)} \quad (5)$$

Where Δt is the time interval between the text to be detected and the topic that has been detected. Parameter L is used to control the speed of time decay.

For the calculation of similarity between text and body, we introduce JS divergence. The value of JS divergence is between 0 and 1. The smaller the JS divergence between two topics, the higher the similarity between the two topics.

The improved algorithm in this paper is divided into two steps. The first step is to calculate the time decay value of the text to be detected, and calculate the similarity between the text and the detected topic. Compared with the set threshold, if the value is greater than the threshold, the text is a new topic; otherwise, the modified text belongs to the topic with the highest similarity. The algorithm flow is as follows:

Algorithm 1. Topic detection algorithm based on multi-view feature fusion

Input: Dataset of news $D = d_1, d_2, d_3, \dots, d_n$, threshold U Time decay threshold V .

Output: A list of topic

- 1: Determine if d is the first news, if yes, go to step 2, otherwise go to step 3
 - 2: Create a new topic and add d to the new topic
 - Step 3. Calculate bursty score of term, $W(i, j)$
 - 3: UPreprocess text d , vectorize text
 - 4: Calculate the time decay value and JS divergence between d and the detected topics $t_1, t_2, t_3, \dots, t_n$. If the time decay value of d and t is greater than V , go to step 5.
 - 5: Record the minimum value of $JS(d||t_i)$. If $\min(JS(d||t_i))$ is less than the threshold u , add d to topic t_i , otherwise go to step 2.
 - 6: Return topic list.
-

4 Analysis of Experimental Results

The data used in the experiment comes from news data under the financial topic of Sina, People's Network and other websites, including 8036 pieces of news data. Each news includes its news title, content, time of publication, and number of comments. Different from the traditional method, considering the intuitiveness and brevity of financial news, the headline and the first paragraph of the body usually contain the main information of the news. Therefore, the headline and the first five sentences in the body of each news article are selected for experiments. The jieba tokenizer is used to tokenize and part-of-speech tags, and then remove

Table 1. Some representative topics in the dataset.

Topic ID	Topic name	Numbers of news
1	China-U.S. Trade war	841
2	Virtual currency trading	738
3	Online loan	947
4	Blockchain	868

stop words and retain verbs, nouns, adjectives, and idioms. Several main topics are shown in Table 1.

In this paper, we compared BHDC algorithm with 4 topic detection algorithms on our dataset.

Table 2. Precision, Recall and F1-Measure of the topic detection results.

Method	Precision	Recall	F1-measure
LDA	0.72	0.756	0.738
CTM	0.752	0.716	0.734
TUS-LDA	0.801	0.775	0.788
CGTM	0.783	0.832	0.807
BHDC	0.847	0.863	0.855

Table 2 show the precision, recall and F1-measure values of different methods. Clearly, As can be seen from Table 2, the precision of the BTDF algorithm is higher than that of the other algorithms, which is 0.127 better than LDA, 0.095 better than CTM, 0.046 better than TUS-LDA and 0.064 better than CGTM. The reason of BHDC further improves the performance may be that our method fuses multiple features for better semantic expression.

We use the Pointwise Mutual Information (PMI) index to evaluate topic quality and compare it with algorithms such as CTM, TUS-LDA, and CTM. The comparison results are shown in Figs. 2, 3, and 4.

Experimental results show that the quality of topics detected by the proposed BHDC algorithm is better than other comparison algorithms. Among them, the topic quality of CGTM and BHDC increases with the number of topics, and CGTM is more sensitive to the number of topic clusters. BHDC is superior to other methods. This is because BHDC integrates multiple text features, obtains more semantic representations, adds the factor of topic time decay, and introduces the concept of JS divergence in the calculation of similarity, which can better distinguish similar new topics from old topics.

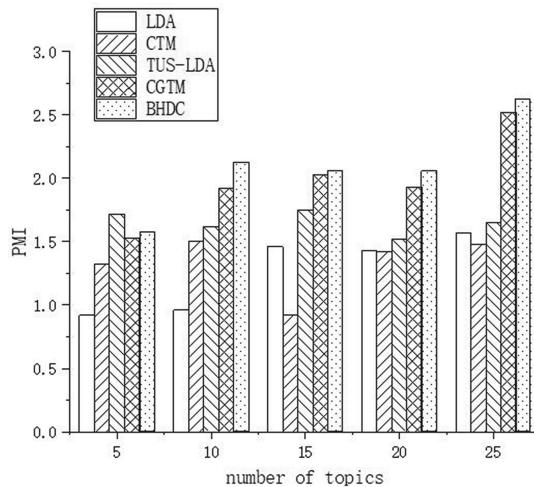


Fig. 2. The PMI values of topic virtual currency trading.

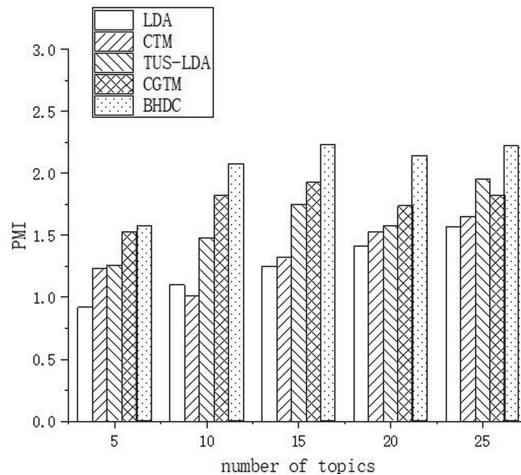


Fig. 3. The PMI values of topic online loan.

In our experiments, the parameters that affect performance include the similarity threshold u . The choice of this parameter also directly affects the performance of the algorithm, and the impact on the algorithm is verified through controlled experiments.

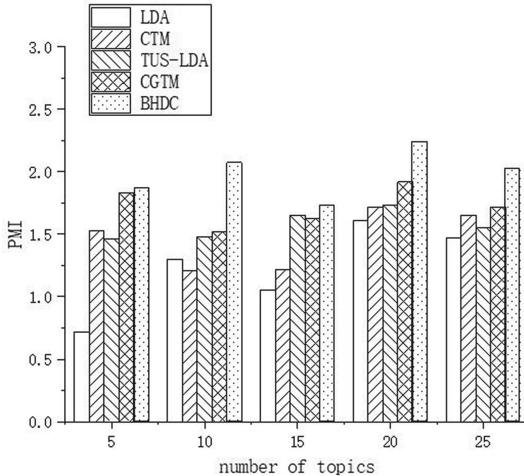


Fig. 4. The PMI values of topic blockchain.

5 Conclusion and Future Work

This paper proposes a topic detection algorithm based on multi-view feature fusion and clustering. It uses different models to extract your semantic features of the text, and calculates the similarity through weighted fusion to obtain a better text semantic representation. At the same time, a time decay factor and JS divergence are introduced in the clustering process. This paper uses multiple sets of experiments on the financial news dataset to verify that the proposed method has higher accuracy, recall and F1-Score than other comparisons.

However, this paper still has some shortcomings in financial topic detection. In summary, it includes two points: the first point is that the topic detection method proposed in this article does not analyze the evolution of the topic, and the second is that it has a good effect on text clustering. However, there is no analysis of the emotional tendency of the topic.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant (No. 61532006, No. 61772083, No. 61802028), and Science and Technology Major Project of Guangxi (GuikeAA18118054).

References

1. Ling, C., Ding, T., Mingqi, L., et al.: A knowledge-based semisupervised hierarchical online topic detection framework. *IEEE Trans. Cybern.* 1–15 (2018)
2. Wang, G., Chi, Y., Liu, Y., et al.: Studies on a multidimensional public opinion network model and its topic detection algorithm. *Inf. Process. Manag.* **56**(3), 584–608 (2019)

3. Zhu, Z., Liang, J., Li, D., Hai, Y., Liu, G.: Hot topic detection based on a refined TF-IDF algorithm. *IEEE Access* **7**, 26996–27007 (2019)
4. Li, J., Wen, J., Tai, Z., Zhang, R., Weiren, Y.: Bursty event detection from microblog: a distributed and incremental approach. *Concurr. Comput.: Practice Exp.* **28**(11), 3115–3130 (2016)
5. Zhang, C., Lu, S., Zhang, C., et al.: A novel hot topic detection framework with integration of image and short text information from Twitter. *IEEE Access* **1** (2018)
6. Taotao, Z., Xiangfeng, L., Wei, Q., et al.: Topic detection model in a single-domain corpus inspired by the human memory cognitive process. *Concurr. Comput.: Practice Exp.* e4642 (2018)
7. Zhang, T., Zhou, B., Huang, J., Jia, Y., Zhang, B., Li, Z.: A refined method for detecting interpretable and real-time bursty topic in microblog stream. *WISE* (1), pp. 3–17 (2017)
8. Zhong, Z., Guan, Y., Li, C., et al.: Localized top-k bursty event detection in microblog. *Chin. J. Comput.* **42**7(07), 76–88 (2018)
9. Zhang, L., Wu, Z., Bu, Z., et al.: A pattern-based topic detection and analysis system on Chinese tweets. *J. Comput. Sci.* (2017). S1877750317309511



Research on Sentiment Analysis of Financial Text Based on Semantic Matching

Xunpu Yuan, Junping Du^(✉), and Feifei Kou

Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia,
School of Computer Science, Beijing University of Posts and Telecommunications,
Beijing 100876, China
junpingdu@126.com

Abstract. With the development of the economy, more and more people on social platforms share financial-related information and hope to understand the development of financial markets through relevant financial data. Aiming at this problem, this thesis proposes a sentiment analysis algorithm of financial texts based on semantic matching. First, preliminary text vectorization is performed on financial text data. Then further training and fine-tuning are performed through pre-trained models in order to fully tap the associations between text contexts and better grasp the semantic focus, so as to better model the text content and obtain higher-level financial text semantic representation. Next, through the improved Siamese network semantic matching model, financial text vectors are trained for semantic matching, so that the distance between financial text vectors with the same emotional category is closer and the distance between different classes is farther, which further optimizes the semantic representation of financial text. Finally, support vector machines are used as classifier to perform sentiment classification of financial texts. Experimental results show that compared with other classic sentiment analysis algorithms, the proposed algorithm has the best sentiment analysis effect on financial text.

Keywords: Sentiment analysis · Semantic matching · Financial text · Vector representation

1 Introduction

With the popularity of the Internet and social development, more and more people are paying attention to the financial field. On the one hand, people will share some of their views on the financial field on social networking platforms, including investment concepts, some personal experience of investment, and promising financial industry information. By sharing, financial information data on social platforms will increase. On the other hand, people also want to get the information they want from these financial data, but how to get the relevant information

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 20–28, 2021.

https://doi.org/10.1007/978-981-15-8458-9_3

they want from the massive financial text data is a difficult problem, which requires the financial text data analysis. However, it is time-consuming and labor-intensive to deal with human resources. Therefore, training a model that can automatically process financial text data has important research value and application value.

Aiming at one aspect of financial text data processing, namely, sentiment analysis, this thesis proposes a sentiment analysis algorithm of financial text based on semantic matching. The sentiment classification of financial texts can provide financial investment institutions, related departments, and investment individuals with relevant information about the direction of the financial market, investment directions, and financial company operating conditions, so as to provide certain assistance for their investment. The contribution of this thesis is to fuse the semantic matching model and the improved Siamese network model to perform vectorization and sentiment analysis on financial text. This improves the effect of extracting the emotional features of financial texts, and also makes the emotional tendency of financial texts more obvious, so that the emotional classification of financial texts can be better achieved. The experimental results show that the algorithm proposed in this thesis has the best sentiment analysis effect on financial text data.

2 Related Work

In the study of semantic representation and matching of texts, literature [1] fuses word embeddings and weight information to obtain a textual representation. Literature [2] proposed a feature word vector representation method based on deep learning Word2vec, then calculated the weight of the feature words through an improved TF-IDF algorithm, and finally multiplied the word weight with the word vector to get the vector representation of the word. Literature [3] proposed a method to learn sentence representation vectors from unlabeled text. Literature [4] proposed a short text modeling method to enhance the semantic learning effect by generating the co-occurrence pattern of words in the corpus. Literature [5] proposed a semantic matching similarity algorithm based on dependency tree through sentence analysis, and generated a new dependency tree model.

In terms of sentiment analysis, the literature [6] uses SVM to perform sentiment analysis of text. Specifically, it finds the support vector in the text vector to realize the sentiment classification of text. Literature [7] introduced weighting schemes into the TextCNN network to achieve text classification and improve the classification effect. Literature [8] used the CRNN-based algorithm to perform sentiment analysis on the text, specifically replacing the last two layers of convolution in the network with a bidirectional LSTM, and then adding a classification layer after the fully connected layer for sentiment classification, improving the classification effect. Literature [9] uses Bi-LSTM for text sentiment classification. Specifically, it uses forward LSTM and backward LSTM to obtain context features in two directions, and then stitches the acquired features to achieve sentiment classification. But these methods can still be improved to further improve the effect.

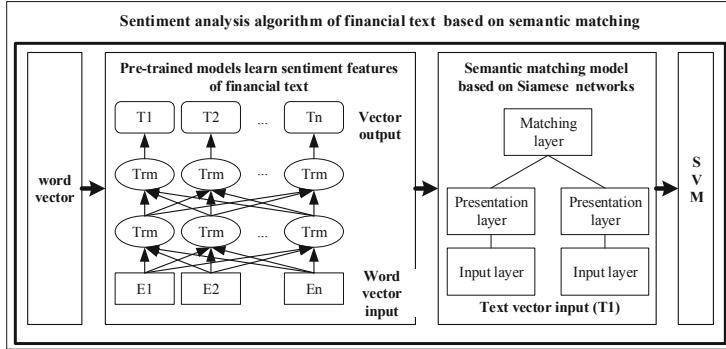


Fig. 1. The architecture of the sentiment analysis of financial text based on semantic matching.

3 Implementation of Sentiment Analysis Algorithm of Financial Text Based on Semantic Matching (BSMSA)

The architecture of the sentiment analysis of financial text based on semantic matching is shown in Fig. 1. For financial text data, the word vector model is first obtained using the word vector model. Then the pre-trained model is used to further mine the association between text contexts and emotional characteristics, so as to obtain a higher-level sentence-level semantic representation. Then construct similar pairs of financial texts, and use the improved Siamese network for semantic matching training, so that the vector distance between financial texts with the same emotional type is closer. Finally, a classifier SVM is added to realize the sentiment classification of financial text.

3.1 Introduction to Pre-trained Model BERT

The financial text semantic representation is obtained through the pre-trained model BERT. BERT combines the two-way Transformer and inputs the entire text sequence into the model at one time in the input part, instead of the previous model, which only inputs the text in a single direction. This input mode allows the BERT model to be trained in both directions when training, and can capture the feature relationships between all text vectors, so that it can better extract the semantic features in the text.

The pre-trained model BERT is shown in Fig. 1. Its input is a sequence, which refers to E1, E2, ..., En in the figure. Therefore, we first need to construct a word vector input sequence. The input sequence consists of three parts, including the corresponding word vector itself, segment representation and position representation. The word vector itself is a preliminary financial text vector obtained from the word vector model. A CLS flag is added at the beginning to indicate the beginning of the sentence, that is, E1. The corresponding output vector can be used as a sentence vector for subsequent emotion classification.

tasks. Add a SEP flag at the end to indicate the end of the sentence, which is En. Segment representation is used to distinguish two sentences, and position representation is used to show that the vectors at different positions are different, which are obtained through model training. Therefore, by inputting the word vector sequence of a sentence in the BERT model, a sentence vector can be obtained, that is, the output vector T1 corresponding to the CLS at the beginning of the sentence.

3.2 Introduction to Improved Semantic Matching Model of Siamese Network

A financial corpus is regarded as a whole sentence for word segmentation to obtain a sequence of word vectors, and then a token sequence is constructed and input into the BERT network to obtain a preliminary financial text sentence vector. Then, the financial text similarity pairs are constructed according to sentiment classification, and the semantic matching training of financial text vectors is performed through the Siamese network semantic matching model to make the distance between financial texts of the same category smaller, thereby completing the semantic matching of financial texts. The implementation framework of the Siamese network semantic matching model is shown in Fig. 2.

The semantic matching model based on Siamese network includes an input layer, a presentation layer, and a matching layer. The input is a 768-dimensional financial text vector obtained using a BERT pre-trained model. The presentation

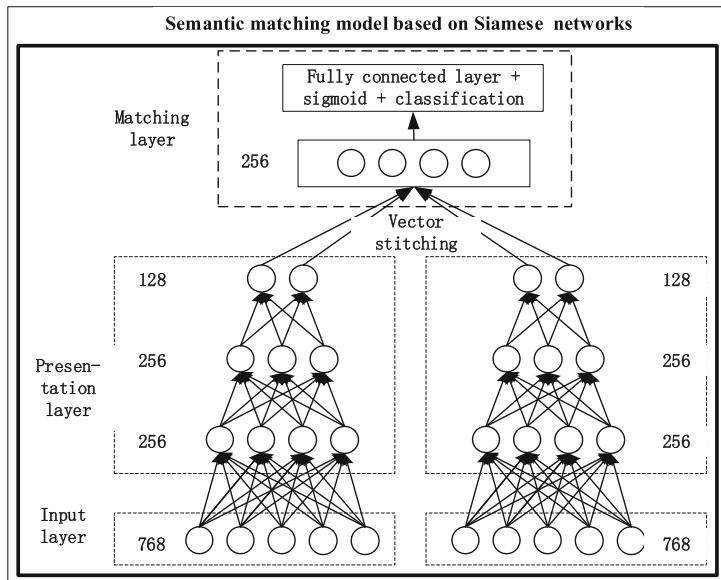


Fig. 2. The implementation framework of the semantic matching model of the Siamese network.

layer uses a DNN neural network with multiple hidden layers. The first hidden layer vector is 256 dimensions, the second hidden layer vector is 256 dimensions, and the output vector is 128 dimensions. Therefore, after inputting a pair of financial text vectors into the network, a pair of 128-dimensional semantic vectors will be obtained.

In the matching layer, the two 128-dimensional vectors are first stitched to obtain a 256-dimensional vector, and then the sigmoid activation function is used to convert the semantic similarity between the two into a similar probability for classification prediction. The specific sigmoid activation function is shown in formula (1) and the loss function is shown in formula (2). Based on this loss function, the model is trained to complete the semantic matching of financial text. Finally, all financial text data is input into the model, and then the 128-dimensional output vector of the presentation layer is taken as the final financial text semantic representation, and all financial text semantic representations can be obtained.

$$\alpha = \sigma(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

where x is the output of the fully connected layer of the matching layer, a value obtained using the sigmoid function.

$$J = -\frac{1}{N} \sum_{i=1}^N [y_i * \ln \alpha_i + (1 - y_i) * \ln(1 - \alpha_i)] \quad (2)$$

where α_i is the predicted value, y_i is the true value, and N is the number of samples.

After the financial text semantic representation is obtained, the text sentiment classification is realized by SVM.

4 Experiments

4.1 Dataset and Evaluation Indicators

First, a corpus of financial texts is constructed. This article collected more than 10,000 financial text corpora, and selected 5,000 positive corpora and 5,000 negative corpora from them. Because two financial texts can form a pair of similar or dissimilar text pairs, the number of text pairs exceeding the number of financial texts can be constructed, so 5000 pairs of similar texts and 5000 pairs of dissimilar texts are constructed. The specific experimental data is shown in Table 1.

In this thesis, precision, recall and F1 are used as evaluation indicators. When the three indicators are relatively high, it shows that the algorithm is effective, and their calculation formulas are as follows.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

Table 1. Experimental data of sentiment analysis

Dataset	Total	Bullish	Bearish	Similar pairs	Dissimilar pairs
Training set	8000	4000	4000	4000	4000
Test set	2000	1000	1000	1000	1000
Total	10000	5000	5000	5000	5000

$$\text{Pr } \text{ecision} = \frac{TP}{TP + FP} \quad (4)$$

$$F1 = \frac{2 \times \text{Pr } \text{ecision} \times \text{Recall}}{\text{Pr } \text{ecision} + \text{Recall}} \quad (5)$$

4.2 Results and Analysis of Semantic Matching Experiments

In the experiment, the parameter adjustment of the output vector dimension of the presentation layer of the Siamese network is first performed, and the vector dimensions are respectively 32, 64, 128, 256, and 512. Then the obtained vectors are stitched, and then the full connected layer and the sigmoid function are used to perform binary classification, that is, to judge whether the financial text pairs are similar or not. Finally, three different evaluation indicators are used for comparison, and the experimental results obtained are shown in Fig. 3. The classification works best when the vector dimension is 128 dimensions. Therefore, the output vector dimension of the representation layer of the Siamese network is selected as 128 dimensions, and the financial text also takes this output vector as the final semantic representation vector.

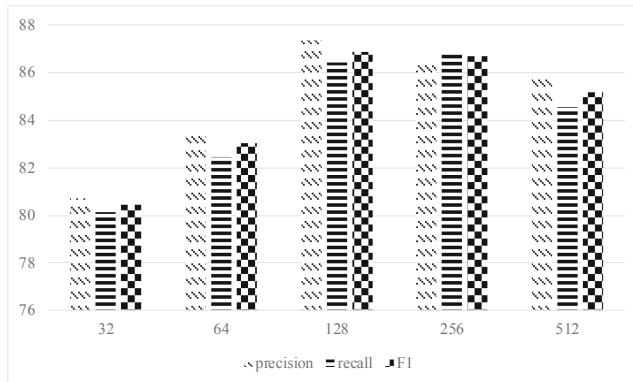


Fig. 3. Comparison of precision, recall, and F1 values for the output vector dimensions of different presentation layers.

4.3 Results and Analysis of Sentiment Analysis Experiments

After obtaining the semantic representation, the same financial text data set was used, and several classic text sentiment analysis methods were selected for comparison experiments. The experimental results are shown in Table 2, the BSMSA algorithm proposed in this thesis has the highest precision, recall and F1 value, and it is about 6% points higher than the worst SVM algorithm evaluation.

Table 2. Comparison of five algorithms in precision, recall, and F1

Algorithm	Precision	Recall	F1
SVM	83.37	82.63	83
TextCNN	85.08	84.17	84.62
CRNN	86.21	85.59	85.9
Bi-LSTM	88.19	87.39	87.79
BSMSA	89.07	88.75	88.91

Then consider the comparison of the two sentiment classifications on the three evaluation indicators, as shown in Fig. 4, the BSMSA algorithm proposed in this thesis performs best. Because the SVM model is a traditional machine learning algorithm, only the initial financial text vector is used for emotion classification, and the overall effect is poor. TextCNN uses convolutional neural networks to extract emotional features in financial texts. Through a number of different convolution kernels, deeper emotional features can be obtained, and the effect is further improved, but it cannot obtain features beyond the convolution kernel length. The CRNN model is a fusion of a CNN model and a bidirectional LSTM, which improves the model, but is not comprehensive enough. The Bi-LSTM model can obtain the features in both directions of the text sequence through the bidirectional LSTM structure, and can also capture the features between the words with a longer distance, so the effect is somewhat improved. But when the distance between words is getting farther and farther, it is difficult to obtain this feature. The algorithm model proposed in this thesis uses the pre-trained model BERT to obtain financial text vectors. When processing financial text sequences, it inputs all text word vectors into the network at one time, and can capture the semantic features between any words. It is not limited by distance, so it can obtain semantic feature vectors that can better express financial text. Then use the semantic matching model to perform semantic matching of financial texts, so that the distance between financial texts of the same emotion is closer, and the distance between different types of financial texts is longer. In this way, when the classifier is used for emotion classification, the experimental effect can be further improved.

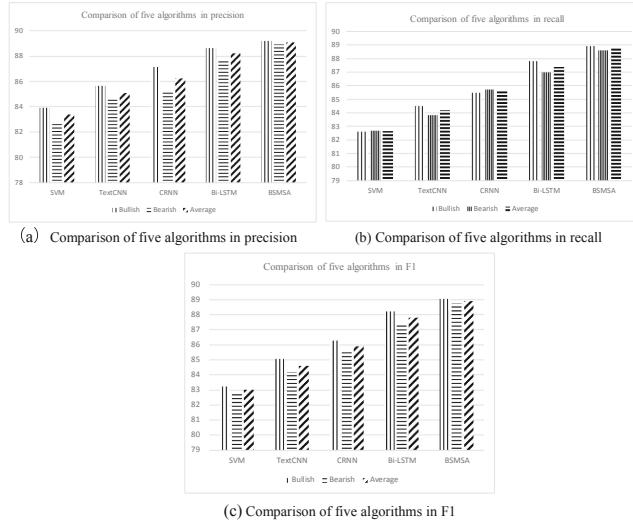


Fig. 4. Comparison of five algorithms in precision, recall, and F1

5 Conclusions

This thesis takes financial texts as research objects, and proposes a sentiment analysis algorithm of financial text based on semantic matching. The pre-trained model and semantic matching model are introduced into the sentiment analysis model, which effectively improves the accuracy of financial text sentiment classification. The main work of this article is as follows:

- (1) Construct financial text vector representation. First, a preliminary financial text word vector is obtained through the word embedding model, and then the financial text word vector is further optimized and trained through a pre-trained model to capture the emotional information in the financial text, thereby obtaining the overall feature vector that can better express the financial text.
- (2) The semantic matching model of the Siamese network is integrated into the financial text sentiment analysis model. First, the obtained financial text vector is input into the Siamese network model. Financial text pairs of the same sentiment category are used as similar text pairs, and different categories are used as similar text pairs. Then the semantic matching training can make the distance between texts with the same emotion classification closer, and the distance between different text pairs farther, so as to achieve emotional focus and make the emotional tendency more obvious. Finally, the financial text vector after semantic matching is input into the classifier to perform sentiment classification. Experimental results show that the algorithm in this thesis performs better.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant (No. 61532006, No. 61772083, No. 61802028), and Science and Technology Major Project of Guangxi (GuikeAA18118054).

References

1. De Boom, C., Van Canneyt, S., Demeester, T., et al.: Representation learning for very short texts using weighted word embedding aggregation. *Pattern Recogn. Lett.* **80**, 150–156 (2016)
2. Liu, C., Sheng, Y., Wei, Z., et al.: Research of text classification based on improved TF-IDF algorithm. In: 2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE), pp. 218–222. IEEE (2018)
3. Arroyo-Fernndez, I., Mndez-Cruz, C.F., Sierra, G., et al.: Unsupervised sentence representations as word information series: revisiting TF-IDF. *Comput. Speech Lang.* **56**, 107–129 (2019)
4. Cheng, X., Yan, X., Lan, Y., et al.: BTM: topic modeling over short texts. *IEEE Trans. Knowl. Data Eng.* **26**(12), 2928–2941 (2014)
5. Kun, H.E., Wei, L.I., Huang, B.: Semantic matching similarity algorithm based on dependency trees. *DEStech Trans. Eng. Technol. Res.* (2018). (icmeit)
6. Li, X., Tuergen, Y., Kahaerjiang, A., et al.: Emotion analysis of active learning based on SVM in Uyghur language. *J. Xinjiang Univ.* (2015)
7. Guo, B., Zhang, C., Liu, J., et al.: Improving text classification with weighted word embeddings via a multi-channel TextCNN model. *Neurocomputing* 363 (2019)
8. Fu, X., Ch'Ng, E., Aickelin, U., et al.: CRNN: a joint neural network for redundancy detection. *Social Science Electronic Publishing* (2017)
9. Xu, G., et al.: Sentiment analysis of comment texts based on BiLSTM. *IEEE Access* **7**, 51522–51532 (2019)



A User Portrait Construction Method for Financial Credit

Chong Liu and Junping Du^(✉)

Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia,
School of Computer Science, Beijing University of Posts and Telecommunications,
Beijing 100876, China
jupingdu@126.com

Abstract. Credit reporting has always been an important issue in the financial field. It is widely used to collect credits from users, portray user portraits, and determine whether to lend to users. This paper proposes a user portrait construction algorithm based on multi-model fusion. This algorithm uses stacking to fuse Word2Vec and text-CNN networks, as well as LDA topic models and BP neural networks to model the user's social media information, and then obtains the user's basic attributes. At the same time, the algorithm uses the stacking method to fuse SVM and KNN to predict the user's credit value from the user's past financial data records. Finally, in order to obtain the final user portrait, the user's social relationship is used to modify the user attributes by applying the PageRank algorithm. The experimental results show that the multi-model fusion method proposed in this paper has the best effect on user portrait construction.

Keywords: User portrait · Financial credit · Social network · Multi model

1 Introduction

The main job of user portraits is to tag users reasonably. Its main purpose is to abstract user information and express user information in the form of tags so that computers can programmatically process user-related information in batches. This requires us to label users objectively and truthfully, then the premise is that we must establish a complete label system for users. Each label in the label system is a representation of some type of user characteristics. The purpose of user portraits is to label users more accurately, facilitate computer processing, and facilitate people to understand such user characteristics.

Lin and Xie [1] construct user portraits of various groups on Weibo, and propose corresponding suggestions for the management of Internet public opinion, personalized services, and marketing strategies. Liu et al. [2] proposed a recommendation method based on dynamic user portraits. Li et al. [3] proposed a secondary fusion algorithm framework for constructing user portraits, which

further improved the average prediction accuracy by about 2%. Zhu et al. [4] collected user behavior data to form user portraits, used machine learning to learn the behavior of normal users. He et al. [5] proposed a design scheme for user behavior portrait of big data network security. Qiu et al. [6] adopted a convolutional neural network method to fuse network structure and text content to represent network users as space vectors, combined with the modularity calculation method based on the Kmeans algorithm. Chen et al. [7] proposed a user portrait method based on an integrated learning framework. This method divided the overall architecture into integrated learning modules and semantic coding modules, and added voting mechanism when making decisions. Zhang et al. [8] used the stacking integration method to fuse multiple cross-modal learning joint representation networks, so that the model can learn the difference of the contribution of different modal representations to the prediction results.

But the current methods mainly use single model to construct user portraits, which is insufficient, and the data source is not adequate, so the prediction can have much bias. So based on the two problems, we proposed a method based on multi model fusion which adopts advantages of different kinds of machine learning algorithms. We also utilize the social network information of user and the social relationship to estimate users credit rating, which makes the result more accurate. Using the embedding and stacking mechanism, we reduce the prediction bias to a certain degree.

2 Algorithm Implementation

2.1 User Financial Portrait Generation Based on Social Networks

The first part predicts the basic attributes of the user based on the microblog sent by the user. Use Text-CNN to extract features from Word2Vec-trained word vectors and predict basic user information. The main idea is to use convolutional neural networks for feature extraction. The core idea of convolutional neural networks is to capture local features. For text, in other words, a local feature is a sliding window composed of several words, similar to N-gram. The advantage of a convolutional neural network is that it can automatically combine and filter N-gram features to obtain semantic information at different levels of abstraction.

Each word vector is obtained as an unknown parameter by training the neural network with the Word2Vec method. The word vector trained by the current network can better capture the features associated with the current task. It is first randomly initialized and is used as network parameters changing with the training process of the network.

The convolutional neural network processes a word matrix composed of word vectors, and the width of the convolution kernel is the same as the width of the word matrix. The width is the size of the word vector, and the convolution kernel only moves in the vertical direction. Therefore, each time the position where the convolution kernel slides is a complete word, a part of the vector of

several words will not be convolved, and the rows of the matrix represent the words in the text, which guarantees the words as the smallest granularity in the language rationality.

Then complete the convolution and pooling process of the user's Weibo. Because the width of the convolution kernel and the word embedding are the same, for a sentence, the result of a convolution kernel is a vector, and its size depends on the height of the convolution kernel filter. It is obtained after max-pooling, and it is a constant. The use of multiple filter sizes enables the kernel to obtain the relationship of words in different ranges, and obtains vertical difference information, which is similar to n-gram, that is, different information will be obtained when words appear in different ranges of a sentence. According to empirical values, use 3, 4, and 5 words as the size of the convolution kernel, and each filter height has multiple convolution kernels. This is because the convolutional neural network learns the parameters in the convolution kernel. Each filter has its own focus, so that multiple convolution kernels can learn multiple different information, and can learn complementary features from the same window. A convolution kernel can only obtain a constant value through a convolution operation. A plurality of constants convolved from the same filter size are combined to form a feature vector. Finally, all vectors of different sizes are combined into a single vector, which is used as the input of the last layer of softmax for the prediction of the end-user portrait. The overall process is shown in Fig. 1.

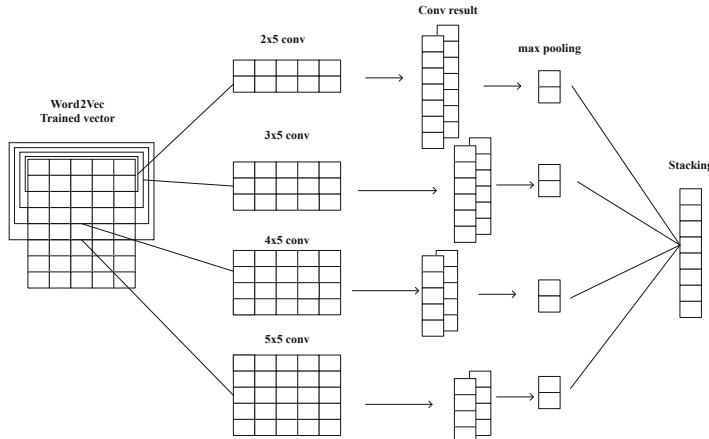


Fig. 1. User microblog calculation user portrait process.

Use two models to make preliminary predictions of basic user attributes. The first model uses Word2Vec model to train the text by adopting the CBOW model, and finally gets a vector representation of each word. Next, each microblog is vectorized. Assuming that the average number of microblogs is w_c , the text is represented in a two-dimensional vector by word stacking. The text with less

than wc is filled with 0, and the text with more than wc is cropped when processing. Then for each text input to the Text-CNN network, use three types of convolution kernels to operate, two for each type of convolution kernels, pool the results of the convolution through max-pooling, and finally pass to the full connection layer using the sigmoid function to predict the basic user attributes.

The second model uses the LDA model to model the text, it uses a 2-layer neural network to train, and finally predicts the basic user attributes. The two models are then connected through a two-layer stacking method. The first layer divides the data set into five folds, then cross-trains the two models. The output value is finally connected through the second layer LR, and output users final basic properties. After this part of the training, the model can make preliminary predictions of the user's basic attributes based on the user's Weibo. Each basic attribute is trained using the above model. After training a total of three stacking models, all the user's basic attributes can be predicted. Starting from any user, the user portrait construction process is shown in Algorithm 1.

Algorithm 1. User portrait construction algorithm

Input: User posts Weibo T, user relationship R, user age LA, user geographic location LG, user education LE.

Output: Text-CNN network parameters of user age, geographic location, and education level are WTA, WTG, and WTE, respectively, and NN parameters of user age, geographic location, and education level are WNA, WNG, and WNE, respectively. The stacking network parameters are WSA, WSG, and WSE.

- 1: Model Word T with Word2Vec to get Word2Vec vector Fw
 - 2: Using LDA model to model T, the LDA vector is Fl
 - 3: Use Fw to train the user's age, geographic location, and education level classifier respectively, and use LA, LG, and LE as labels to train, and obtain network parameters WTA, WTG, and WTE
 - 4: Use Fl to train the user's age, geographic location, and education level classifier respectively, and use LA, LG, and LE as labels to train, and obtain the network parameters WNA, WNG, and WNE
 - 5: Use the vectors obtained by the Text-CNN and NN networks as inputs, and use LA, LG, and LE as labels to train the stacking classifier to obtain the network parameters WSA, WSG, and WSE
 - 6: Returns the network parameters WTA, WTG, WTE, WNA, WNG, WNE, WNA, WNG, WNE
-

2.2 User Financial Portrait Generation Based on User Financial Characteristics

The user's financial characteristic data sources mainly include income and expenditure information, historical amount transactions, consumer behavior records, and credit records. First, the user's various types of information are embedding coded, and the discrete data is vectorized. The main method is to normalize the

data by dividing the data range. After that, the user's credit defaults and the frequency index are used to construct the portrait of the user.

The training methods are mainly divided into two categories. One is based on the SVM method, which is a linear classifier with the largest interval defined in the feature space. Using kernel techniques, it is constructed into a substantially nonlinear classifier. Using the maximum interval learning strategy, the problem is formalized as a problem for solving convex quadratic programming, which is also equivalent to the minimization problem of the regularized hinge loss function. The learning algorithm used is an optimization algorithm for solving convex quadratic programming.

Assuming that the hyperparameters defining the separation are w and b , the expression of the solution target is shown in Eq. (1).

$$\min \theta(w) = \min \max L(w, b, \alpha) = p^* \quad (1)$$

Another learning method is to use a KNN model, whose input is the feature vector of the instance, which corresponds to the points in the feature space, and whose output is the category of the instance. The KNN algorithm is a training data set for a given user's financial characteristics, in which the credit characteristics of each user have been determined and divided into 3 different levels. During classification, the new instance is predicted by means of majority vote, etc. according to the categories of its k nearest neighbor training instances. The label prediction result of data i is shown in Eq. (2).

$$L_i = \min_j \left(\sum_{d=0}^v (v_i(d) - v_j(d))^2 \right) \quad (2)$$

At the same time, as in the previous part, use the two-layer stacking model to train the fusion of multiple classifiers, using the divided training set and test set, and finally train the final classifier through LR to train each stacking model. Lastly, use a 2 layer stacking model to collect all results from revenue and expenditure information, historical amount information, and credit record information through LR, and the final classification result is used as the user's financial characteristic portrait.

2.3 PageRank-Based User Portrait Correction

After constructing the user's financial portrait in the first two parts, construct a directed graph of social relations based on the user's social relationship, and then calculate the PR value of the node through the PageRank algorithm. Then use the PR value to modify the basic attributes of the user, and finally get the ultimate portrait of the user. The calculation of the PR value is shown in Eq. 3, where $\text{PR}(A)$ is the PR value of user A, $\text{PR}(T_i)$ is the PR value of user T_i , and T_i is a user related to A, such as friends, relatives, $C(T_i)$ is the total number of users associated with T_i , that is, the number of points T_i points to, and d is the damping coefficient.

$$PR(A) = (1 - d) + d \sum_{i=1}^n \frac{PR(T_i)}{C(T_i)} \quad (3)$$

When the previous construction of the current user portraits is completed, use the PR value calculated by PageRank to modify the portrait attributes. Assuming that the original user portrait vector result is R and the corrected user portrait result is RN, the users ultimate portrait can be obtained via RN. The calculation of RN is shown in Eq. 4.

$$RN(k) = e^{I \cdot L(k)^T} * R(k) \quad (4)$$

The PageRank method is used to modify the basic user attributes obtained by the first part of the users social media, including the users age, geographic location, and education level. At the same time, the second part of the user's credit index is modified, and finally forms the ultimate portrait.

3 Experimental Results and Analysis

3.1 Dataset

The financial data set contains 52,956 Weibo accounts of financial users, covering a total of 1875 users, and using the income and expenditure information, historical amount transactions, and credit records of 1875 users as user financial data information, the numbers are 19871, 25469 and 9467 respectively. Among them, there are 6,194 social relationships among 1,875 users. As predicted, there are 1,875 user credit indexes in total. After all the data is preprocessed and cleaned, it is divided into training sets, cross-validation sets, and test sets. Use 50% of data as the training set, 10% as the cross validation set, and 40% as the test set.

3.2 Experimental Results

The experimental evaluation index mainly uses the results of the accuracy rate, recall rate and F_1 value to evaluate the proposed model. The true positive example is defined as TP, the false negative example is FN, and the false positive example is FP.

First, analyze user portraits on user's social network Weibo. For all Weibo texts sent by users, remove stop words and punctuation, convert all text to simplified Chinese, and do segmentation. Then train through the method of Algorithm 1, and finally predict the three types of information of the user's age, geographical location and education level in the test set, and calculate the accuracy rate, recall rate and F_1 value of each type of information, as shown in Table 1, Table 2 and Table 3.

From the overall results of the three tables, only the LDA with BP network has the lowest comprehensive index, because LDA cannot well model the topic of Weibos short text, and it is difficult to learn the number of topics and the

distribution of topic and words. It is better to only keep the Word2Vec with text-CNN network, because the Word2Vec method can extract the words of the text, and the granularity is smaller. At the same time, the text-CNN network is used to extract the features of the local text through different windows, thus have knowledge of the context, so the performance is better than pure BP networks. When the two models are fused using simple weighted summation, the effect is not much improved compared to the text-CNN method, because the linear summation alone cannot be used to further integrate the advantages of the two models. Only when the two weak classifiers are combined using the stacking model with LR, the final performance is improved, indicating that the model fusion method is more complete. At the same time, the results obtained on the three basic attributes are relatively stable, indicating that the algorithm has good versatility.

Table 1. Average accuracy rate, recall rate and F_1 value of user age prediction

	Precision	Recall	F_1 Measure
BP Network	0.8009	0.7789	0.7897
text-CNN	0.8277	0.8001	0.8137
BP and text-CNN	0.8301	0.8137	0.8218
Stacking	0.8793	0.8503	0.8646

Table 2. The average accuracy rate, recall rate and F_1 value of the user's geographical location prediction

	Precision	Recall	F_1 Measure
BP Network	0.8014	0.7950	0.7982
text-CNN	0.8198	0.8207	0.8202
BP and text-CNN	0.8241	0.8332	0.8286
Stacking	0.8564	0.8693	0.8628

Table 3. Average accuracy rate, recall rate and F_1 value of users' education level prediction

	Precision	Recall	F_1 Measure
BP Network	0.8146	0.8123	0.8134
text-CNN	0.8312	0.8221	0.8266
BP and text-CNN	0.8573	0.8378	0.8474
Stacking	0.8694	0.8575	0.8634

Then it comes to the prediction result of the user's credit rating based on the user's financial data information. The model is also evaluated using the accuracy rate, recall rate, and F_1 value, and the three credit ratings are predicted respectively as low, medium, and high. The results obtained are shown in Table 4, Table 5, and Table 6, respectively.

From the results of the three tables, only using KNN to predict the user's credit rating has a poor performance. The main reason is that the KNN algorithm only considers the Euclidean distance between the input vector and other vectors. It determines the result by simply discriminating several points around it. The labeling of points will lack some accuracy, resulting in points that are farther away from radiating to the current point, and the fitting ability is limited. SVM takes into account the distances of all points and fits the hyperplane to make the overall distance interval the largest, so that the distances between different classes are separated, and the overall classification performance is better. Using linear fitting to weight the two models at the same time, the results obtained are almost no improvement, because the two models are both based on distance classification, not completely irrelevant classifiers, so the effect is not obvious. But it can be seen later that the overall effect is slightly higher when combined with the stacking model, because the folding of the stacking model divides the

Table 4. Accuracy, recall and F_1 values for low credit rating predictions

	Precision	Recall	F_1 Measure
BP Network	0.8573	0.8435	0.8503
text-CNN	0.8789	0.8646	0.8717
BP and text-CNN	0.8802	0.8636	0.8718
Stacking	0.8961	0.8893	0.8927

Table 5. Accuracy, recall and F_1 values of the middle credit rating prediction

	Precision	Recall	F_1 Measure
BP Network	0.8678	0.8417	0.8545
text-CNN	0.8877	0.8661	0.8768
BP and text-CNN	0.8977	0.8603	0.8786
Stacking	0.9001	0.8877	0.8939

Table 6. Accuracy, recall and F_1 values for high credit rating predictions

	Precision	Recall	F_1 Measure
BP Network	0.8479	0.8567	0.8523
text-CNN	0.8809	0.8584	0.8695
BP and text-CNN	0.8878	0.8641	0.8758
Stacking	0.9136	0.8759	0.8944

data set, makes the two models have a higher degree of irrelevance, and the overall performance has improved.

4 Conclusion

This paper proposes a multi-model fusion-based user portrait method for financial credit reporting, which mainly contains three steps, the base social attribute prediction of user, the credit rating and results correlation. Experiments demonstrate the effectiveness of the proposed method. However, the method proposed has a huge network size, which requires training a large number of networks, making the overall training time longer. In the future, network merge or random dropout will be used to speed up the training speed and improve the overall performance.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant (No. 61532006, No. 61772083, No. 61802028), and Science and Technology Major Project of Guangxi (GuikeAA18118054).

References

1. Lin, Y., Xie, X.: User portrait of diversified groups in microblog based on social identity theory. *Inf. Stud. Theor. Appl.* **41**(03), 142–148 (2018)
2. Liu, Y., Wu, X., Xie, B.: Research on information recommendation based on dynamic user portrait. *Comput. Syst. Appl.* **27**(6), 236–239 (2018)
3. Li, H., Lin, H., Yang, L., et al.: Two-level stacking algorithm framework for building user portrait. *Comput. Sci.* **45**(1), 157–161 (2018)
4. Zhu, J., Chen, G., Shi, Y., Xue, Z.: Abnormal behavior detection based on user profile. *Inf. Stud. Theor. Appl.* **50**(10), 2310–2315 (2017)
5. He, X., Huang, M., Song, F.: User behavior profiling scheme for network security. *Commun. Technol.* **50**(04), 789–794 (2017)
6. Que, Y., Zhang, W.: Study for the construction method of group profile based on network structure and text content. *Libr. Inf. Serv.* **63**(22), 21–30 (2019)
7. Chen, Q., Ling, M., Sun, L., Jia, Y.: User profile method based on ensemble learning framework. *J. Zhejiang Sci.-Tech. Univ. (Nat. Sci. Edn.)* **43**(01), 86–93 (2020)
8. Zhang, Z., Feng, X., Qian, T.: User profiling based on multimodal fusion technology. *Acta Scientiarum Naturalium Universitatis Pekinensis* **56**(01), 105–111 (2020)



E-Product Recommendation Algorithm Based on Knowledge Graph and Collaborative Filtering

Jianghai Lv¹, Yawen Li², Junping Du^{1(✉)}, and Lei Shi¹

¹ Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia, School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

junpingdu@126.com

² School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. The recommendation system is of great significance for screening effective information and improving the efficiency of information acquisition. Traditional recommendation systems face problems such as sparse data and cold starts. Based on the combination of external scoring and item connotation knowledge, an e-product recommendation model RKGCF based on cyclic knowledge graph and collaborative filtering is proposed. After fully considering the correlation between items, users, and ratings, Top-K recommendations are made using collaborative filtering based on items and users. To reveal the semantics between entities and relationships and understand user interests, external data and user preference data of items are added to the knowledge graph to extract the dependency relationship between entities and construct interactive information between users and items. Multiple sets of different negative samples are used to train the model, and the real e-commerce data are used for testing. The experimental results demonstrate that the model has significantly improved the accuracy of the recommendation effect.

Keywords: Knowledge graph · Collaborative filtering · LSTM

1 Introduction

The main problem in the research of recommendation systems is how to find the content that each user is interested in from overloaded information and push these contents to the user. Researchers consider using knowledge maps to improve the description of users and items in content-based recommendation systems to improve the recommendation effect. Auxiliary information can enrich the description of users and items, enhance the mining ability of recommendation algorithms, thereby effectively solving the problems of sparsity and cold start, and improve the accuracy, diversity and interpretability of recommendation results.

On the basis of collaborative filtering recommendation, Liao et al. [1] proposed a hybrid method that uniformly mixes the results of collaborative filtering based on user attribute information with traditional collaborative filtering into the final result. Wang et al. [2] proposed a hybrid model combining collaborative filtering recommendation with content-based recommendation, which effectively solved the problem of recommending new items. Yan [3] proposed an integrated recommendation model based on two-stage deep learning. Xu et al. [4] based on the basic user information and movie meta-information, which could generate the top-N recommendation list with the highest similarity.

As a new type of auxiliary data source, knowledge graph has attracted scholars' attention. Jia et al. [5] proposed the use of network embedding for feature extraction of tourism knowledge atlas, which makes the feature extraction more sufficient. Xing et al. [6] carried out research and application of knowledge map construction technology for specific fields. Wu et al. [7] integrates the semantic information of items into collaborative filtering recommendations by calculating the semantic similarity between items. Sang et al. [8] proposed a model that combines biomedical knowledge maps, map embeddings with deep learning methods based on literature discovery.

In order to effectively recommend, this paper proposes an efficient recommendation model: RKGCF (Recommendation Algorithm Based On Knowledge Graph and Collaborative Filtering) based on the basic idea of hybrid recommendation, combined with deep learning, collaborative filtering and knowledge map on the basis of recurrent network. Cyclic knowledge graph embedding are used to learn the semantic representation of entities and the paths between entities. External scores are added to the knowledge map as learning weights to better express the user's preference. A penalty factor is added to the similarity calculation to eliminate the influence of popular items and inactive users on the results. Experimental results show that the proposed framework effectively improves the accuracy of the recommendation.

2 The Proposed Method

This chapter will introduce the specific internal structure of RKGCF in blocks. First, the e-commerce recommendation model of recurrent knowledge graph is used to automatically learn the semantic representation of entities and paths between entities to characterize the user's preference for items. Based on the knowledge graph, it is combined with recurrent neural network to learn and form a recurrent knowledge graph. Considering the length of the relationship sequence between e-commerce entities, this paper uses cyclic knowledge graphs to learn the semantics of entity relationships more conveniently. It can model sequences of different lengths. It is particularly suitable for modeling paths and capturing entities and entity pairs. The whole path between the semantic capabilities is better. For multiple paths and possible connection entities of different lengths, the net-work can capture all possible relationships.

After joining a batch of recursive recurrent neural networks, the paths of the same entity pair can be linked, that is, the association between entities

is completed. Then, the semantics of the path of the entity pair are modeled, and the path is seamlessly integrated into the recommendation model, so that each entity and relationship can learn the corresponding low-dimensional vector. While maintaining the original structure or semantic information, it is also convenient to link the paths of the same semantic entity, and then integrate these paths into the recommendation to improve the accuracy of the recommendation. Therefore, a good set of entity vectors can fully and completely represent the interrelationships between entities. Using cyclic knowledge graph feature learning can easily introduce data features into various recommendation system algorithms.

Then collaborative filtering recommendations can be added, including item-based collaborative filtering and user-based collaborative filtering. For collaborative filtering recommendations, the system will perform a nearest neighbor search and calculate the related similarity to get the recommendation results. The cyclic knowledge map can learn the connotation knowledge in the recommendation relationship, and collaborative filtering can make good use of external ratings. The proposed method combines the connotation knowledge and external ratings to effectively improve the efficiency of recommendation.

2.1 Recurrent Knowledge Graph (RKG)

The cyclic knowledge graph includes the semantic path, LSTM network layer and output layer. The e-commerce knowledge graph triplet is obtained based on the feature of the e-commerce knowledge graph, which indicates that the user has purchased the product. Blue indicates an e-commerce entity, and green indicates a user entity. The arrow describes the relationship between the person and the product, which means that the person has purchased the product. Knowledge graphs are formed by connecting similar multiple triples to each other.

The RKGCF model contains a batch of LSTM structures, and each LSTM learns the semantic representation of a specified path. The path length of the entity pair (u_i, v_i) is dynamic. For any path p_l with length T , it can be expressed as: r_1, r_2, r_T , the path is $p_l = e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} e_2 \rightarrow \dots \rightarrow e_T$, $e_0 = u_i$, $e_T = v_j$. LSTM encodes a path by learning the semantic representation of each entity and a single representation of the entire path. In order to make full use of the entity relationships in the knowledge graph, firstly, the paths with different semantics between the entities are mined, and then these paths are seamlessly integrated into the recurrent network batch processing for effective recommendation. In order to improve the efficiency of the model, length-constrained enumerated paths are used, that is, only paths shorter than the threshold are used.

Therefore, it is possible to mine different semantic paths between entities according to the content in the knowledge map. After extracting the relationship path, the user's preference relationship can be inferred based on the key path for easy recommendation.

2.2 LSTM Network Layer

In the previous section, the user-item entity pair in the model is considered as a sequence, and the elements in the sequence are the entities in the path, and the path is encoded using the LSTM composed of the embedding layer and the attention gate hidden layer. The architecture contains a batch of LSTMs that encode paths by learning the semantic representation of each entity and a single representation of the entire path. Learn a distributed representation p_{lt} for each entity e_t in p_l at the embedding layer, which maps e_t to a low-dimensional vector and captures the entity's semantics, then provides this new representation as input to the hidden layer to learn coding a single representation of the entire path. Note that in order to learn the path representation, the gated hidden layer considers the embedding of the entities in the path and the order of these entities, and uses a stream-based method to encode the sequence of the start entity to the end entity of the path, and finally obtains the entire path representation h_{lT} . Let a_{lt} be the attention gate at step t , which is a scalar value between [0, 1]. The hidden state at time t can be expressed as:

$$h_{lt} = (1 - a_{lt}) \cdot h_{l(t-1)} + a_{lt} \cdot h'_{lt} \quad (1)$$

The attention gate a_{lt} balances the input contributions of the previous hidden state $h_{l(t-1)}$ and the current candidate hidden state h'_{lt} . By fully considering the input of the current time step, the current candidate hidden state is further given:

$$h'_{lt} = \sigma(W \cdot h_{l(t-1)} + H \cdot p_{lt} + b) \quad (2)$$

W , H are the linear transformation parameters of the previous step and the current step, b is the bias term, and σ is the sigmoid activation function. Finally, based on the input observations of the current time step and the information of adjacent observations in two directions, a model of the attention gate is established:

$$a_{lt} = \sigma(M^T \cdot (\overrightarrow{h_{lt}} : \overleftarrow{h_{lt}}) + b') \quad (3)$$

σ is a sigmoid activation function used to control the range of attention gates between [0, 1]; M is a weight vector, and b' is an offset term of the attention layer; ":" represents a connection between quantities. $\overrightarrow{h_{lt}}$ summarizes the path from start to step t , and $\overleftarrow{h_{lt}}$ summarizes the path from end to step t , given by:

$$\overrightarrow{h_{lt}} = \sigma(\overrightarrow{M} \cdot \overleftarrow{p_{lt}} + \overrightarrow{H} \cdot \overrightarrow{h_{lt-1}} + \overrightarrow{b}) \quad (4)$$

$$\overleftarrow{h_{lt}} = \sigma(\overleftarrow{M} \cdot \overleftarrow{p_{lt}} + \overleftarrow{H} \cdot \overleftarrow{h_{lt-1}} + \overleftarrow{b}) \quad (5)$$

The qualified paths between u_i and v_i are merged into the corresponding attention-gated network at the same time, and the entity relationships of all u_i and v_i are obtained. Because there are multiple paths connected between u_i and v_i , different paths will have different degrees of influence when modeling

the relationship between them. Therefore, to distinguish the most important characteristics of different vectors through pooling operations, the max-pooling layer can be formalized as:

$$h[j] = \max_{1 \leq i \leq S} h_{iT_i}[j] \quad (6)$$

If there are s paths between u_i and v_i , after learning through LSTM, the final hidden state is $h_{1T_1}, h_{2T_2}, \dots, h_{ST_S}$, where TS is the last step. The most significant feature on all paths is obtained through the pooling layer. Then the fully connected layer is used to further quantify the relationship (proximity) between u_i and v_i . After completing the model training, sort the items by the proximity score, and recommend the top K items with the highest score to u_i .

2.3 Collaborative Filtering (CF)

Both user-based collaborative filtering and item-based collaborative filtering algorithms are used. Two different Top-K recommendation results are obtained respectively, and the final Top-K results can be obtained by merging with the results obtained from the previous circular knowledge map.

First treat each user or item as a vector, and then calculate the similarity of all other users or items with other ones. After having the similarity between two pairs, the system also recommends to the user. The circular knowledge map can learn the connotation knowledge in the recommendation relationship, collaborative filtering can make good use of external scores, and the proposed method combines the connotation knowledge and external scores to effectively improve the efficiency of recommendation.

In user-based collaborative filtering, when the similarity between a user and other users is calculated based on the user's historical behavior, the calculation is as follows:

$$W_{uv} = \frac{\sum_{i \in N(u) \cap N(v)} \frac{1}{\lg 1 + |N(i)|}}{|N(u)| \cdot |N(v)|} \quad (7)$$

$N(u)$, $N(v)$ respectively represent the set of products with positive feedback through u and v . If users have taken similar behaviors on unpopular items, they can better express the similarity between the two. So we add a penalty factor $\frac{1}{\lg 1 + |N(i)|}$, used to penalize the impact of similarity of popular products in the common product list between users.

In item-based collaborative filtering, when calculating the similarity between items, use the following formula:

$$W_{ij} = \frac{\sum_{i \in N(i) \cap N(j)} \frac{1}{\lg 1 + |N(u)|}}{|N(i)| \cdot |N(j)|} \quad (8)$$

$N(i)$ and $N(j)$ represent the number of users who like product i and user j , respectively. Because the contribution of active users to the item similarity calculation will be smaller than that of inactive users, a penalty factor $\frac{1}{\lg 1 + |N(u)|}$ is also added to reduce the impact.

3 Knowledge Graph and Collaborative Filtering Fusion Algorithm

The cyclic knowledge graph embedding adopts a new recursive network architecture. The architecture contains a batch of recursive networks for semantic modeling of paths linking the same entity pairs. These paths are seamlessly integrated into recommendations. And choose a good recommendation path for Top-K recommendation. At the same time, collaborative filtering is used to collect user behaviors to obtain their explicit or implicit information about items, and Top-K recommendations are given based on item-based collaborative filtering and user-based collaborative filtering, respectively. The model finally fuses the Top-K recommendation results given by the three methods to obtain the final recommendation list. The fusion of the two parts of the recommendation results can better improve the recommendation efficiency.

We propose two different result fusion algorithms for the situation of entity connotation knowledge and external scoring. According to the recommendation list obtained from the cyclic knowledge map and the recommendation list obtained by the collaborative filtering algorithm, the same or top-ranked results can be extracted through the fusion algorithm to fuse and obtain new recommendation results. The obtained recommendation results can improve the effectiveness of the recommendation on the one hand and explain the source of each recommendation result on the other hand, so that the results of collaborative filtering and circular knowledge map can make up for each other.

The first fusion method is referred to as the Loop Extraction Fusion (LEF). Collaborative filtering based on users, collaborative filtering based on items, and collections of items generated based on cyclic knowledge maps are traversed, and the items in the three sets are sequentially placed in the Top-K recommendation set T , and in the process of being placed in the recommendation set T . To ensure that the placed objects do not exist in T , that is, to ensure the uniqueness of the objects in the recommendation set T .

The second fusion method is referred to as Loop Comparison Fusion (LCF). Collaborative filtering based on users, collaborative filtering based on items, and collections of items generated based on cyclic knowledge maps are traversed to determine whether the current object in each set exists in the other two sets. If it exists, the current object in the current set is placed in the Top-K recommendation set T , in the process of putting in the recommendation set T , it is also necessary to ensure that the placed objects do not exist in T , and the uniqueness of the objects in the recommendation set T is guaranteed.

4 Experimental Results and Analysis

4.1 Experimental Data Set

In order to test the validity of the model, the real purchase information and review information of users purchased in the JD data set are used for verification.

The data set contains information about the products and reviews purchased by the user, as well as data such as the manufacturer, type, and price of the product. Each user has purchased at least two items.

The total number of times a user purchases a product is 6000, the number of users is 3321, the number of product reviews is 2447, the number of product manufacturers is 1419, the type of product is 7 and the number of relationship types is 5. In the experimental data set, each user corresponds to multiple purchased products, and the data of the user's product rating less than or equal to the threshold r is regarded as negative feedback, and then train the model. $r = 0, 1, 2, 3, 4$ to obtain the impact of different external scores on connotative knowledge. Among them, 0 means that the user has not rated the product.

To verify the performance of the recommendation system, precision is used to evaluate the recommendation ability of the model. Precision describes the proportion of users who have actually purchased the product in the final recommendation list given in the recommendation system.

4.2 Comparative Analysis of Experimental Parameters

The LEF fusion method and the LCF fusion method were respectively used to recommend Top-1, Top-5, Top-10, and Top-15 on the JD data set. UserCF and itemCF were added on the basis of RKGCF, and added at the same time. userCF and itemCF (hereinafter collectively referred to as CF), and the changes in adding different ratings to RKGCF based on the simultaneous addition of userCF and itemCF. All the data below RKGCF (r_0, r_1, r_2, r_3, r_4) represent the impact of different external scores on connotative knowledge, where r_0 means that no external score is added. For example, RKGCF (r_2) indicates that data with a movie score of 2 or less is used as negative feedback, and then the model is trained.

Table 1. Precision of fusion algorithm (LEF).

	Precision-1	Precision-5	Precision-10	Precision-15
RKGCF+user	0.214	0.218	0.232	0.208
RKGCF+item	0.235	0.224	0.222	0.232
RKGCFr4	0.273	0.210	0.332	0.290
RKGCFr3	0.333	0.330	0.371	0.331
RKGCFr2	0.328	0.309	0.272	0.291
RKGCFr1	0.346	0.338	0.387	0.372
RKGCFr0	0.324	0.314	0.338	0.334

It can be seen from Table 1 that the recommendation performance of LEF is excellent. When the K value recommended by Top-K is relatively large, the

model can maintain relatively good performance. At the same time, the recommendation result of adding CF will be significantly better than adding user-CF alone. Or itemCF; when the score is less than 1 for negative feedback, the comprehensive recommendation result is slightly better than other scores.

4.3 Comparative Analysis of Experimental Methods

Table 2 compares the proposed model with MostPop, BPRMF, LIBMF, NCF, HeteRS, HeteRec, GraphLF, CKE, RKGCF and other methods to prove that the model has good performance. They are introduced as follows. CF: The recommendation algorithm based on collaborative filtering is one of the most popular recommendation algorithms. BPRMF: A Bayesian posterior optimized personalized score ranking algorithm based on matrix factorization. It does not optimize the user's rating of the item itself, but only uses the score to optimize the user's ranking of the item. NCF: Neural Collaborative Filtering Algorithm, a recommendation method based on neural networks. It is mainly used to solve the problem of collaborative filtering based on implicit feedback. RKG: Use knowledge graph embedding and a set of recursive network structures to automatically learn the paths and semantic relationships between entities, so as to better describe the preference information of items to users.

Table 2. Comparison results of 5 experimental algorithms.

	Precision-1	Precision-5	Precision-10	Precision-15
CF	0.105	0.092	0.112	0.131
BPRMF	0.265	0.289	0.333	0.307
NCF	0.333	0.355	0.382	0.384
RKG	0.331	0.389	0.377	0.349
RKGCF	0.448	0.438	0.417	0.474

It can be seen from Fig. 1 that the current mainstream collaborative filtering-based algorithms have low recommendation performance, rely on the quality of the data set, and have low utilization of other auxiliary information in the data, and the semantic information of the knowledge map can be sufficiently effective Improve the accuracy of the recommendation system. In summary, the recommendation algorithm based on the RKGCF model is far superior to other models on different Top-1, Top-5, Top-10, and Top-15, and it can be maintained when the value of K is large. Higher recommendation accuracy.

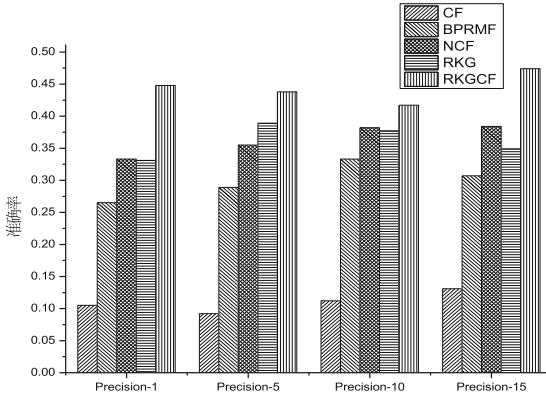


Fig. 1. IDCNN-ATT-CRF entity name identification network structure.

5 Conclusion

This paper proposes a hybrid recommendation model based on cyclic knowledge graph embedding. This model can either discover the user's existing interests through collaborative filtering, or use the knowledge graph to mine the user's potential interests. The two results are fused to obtain a personalized Recommended results. The model can automatically learn path relationships between entities and infer preference relationships. At the same time, an external score is added to the knowledge map as a learning weight to better express the user's preference. Finally, different fusion methods are used to fuse the connotative knowledge with the external scoring results to obtain the optimal fusion recommendation result. The results show that the proposed framework achieves accurate recommendation results.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant (No. 61902037, No. 61532006, No. 61772083, No. 61802028), and Science and Technology Major Project of Guangxi (GuikeAA18118054).

References

1. Liao, X., Liu, C., Xu, X., et al.: Research on collaborative filtering hybrid recommendation system based on user attributes. *Comput. Digital Eng.* **47**(6), 79–85 (2019)
2. Wang, R., Zongda, W., Jiang, Y., et al.: An integrated recommendation model based on two-stage deep learning. *J. Comput. Res. Dev.* **56**(8), 1661–1669 (2019)
3. Yan, W., Jie, T.: Personalized recommendation algorithm for papers based on deep learning. *J. Chin. Inf. Process.* **32**(4) (2018)
4. Xu, M., Ding, Y., Chen, F., et al.: Personalized movie recommendation algorithm based on joint neural network. *Mod. Film Technol.* **11**, 25–30 (2019)
5. Jia, Z., Gu, T., Bin, C., et al.: Attraction recommendation for feature learning of tourism knowledge atlas. *J. Intell. Syst.* **14**(03), 430–437 (2019)

6. Xing, L.: Research and Application of Knowledge Graph Construction Technology for Specific Fields (2018)
7. Wu, X., Chen, Q., Liu, H., et al.: Collaborative filtering recommendation algorithm based on representation learning of knowledge graph. *Comput. Eng.* **44**, 226–232 (2018)
8. Sang, S.: Knowledge graph-based bidirectional recurrent neural network method for literature-based discovery. In: IEEE International Conference on Bioinformatics and Biomedicine. IEEE (2018)



Financial Sentiment Analysis Based on Pre-training and TextCNN

Xunpu Yuan¹, Yawen Li², Zhe Xue^{1(✉)}, and Feifei Kou¹

¹ Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia, School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

xuezhe@bupt.edu.cn

² School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. Since the research of sentiment analysis is mostly concentrated in the field of sentiment analysis on Weibo, and there is less research on sentiment analysis of financial text, this thesis proposes a financial sentiment analysis model based on pre-training and TextCNN. First, the pre-trained model is used to initially extract the emotional features of the text. It can extract text features well, and can extract information between words at arbitrary intervals when processing text sequences. Then use the improved TextCNN to construct a sentiment analysis network to further extract the sentiment features of the text, effectively identify the sentiment of the text, and complete the sentiment analysis of financial text. This thesis conducts experiments on a balanced corpus data set based on financial texts, and compares it with other classic sentiment analysis algorithms. Experimental results show that the proposed method works best in the field of financial text sentiment analysis.

Keywords: Pre-trained model · Sentiment analysis · Financial text

1 Introduction

Because people pay attention to artificial intelligence technology and apply it to all aspects of society, text sentiment analysis has also attracted the attention of many researchers. Most sentiment analysis research is based on microblog texts sent by users. There are also sentiment analysis on product reviews, but less analysis of financial texts.

Nowadays, economic development is closely related to people's lives. People need to know the development of the financial field, so they need to analyze financial data, and emotional analysis of financial text data is an important part of it, which has very important research value and application value. The sentiment analysis of financial texts, on the one hand, allows relevant departments to understand the social and economic development situation, market prospects

and industry development trends. On the other hand, it can let investors or investment institutions understand the financial situation of each company and the development of the stock market, so as to provide them with certain help.

2 Related Work

For sentiment analysis technology, the predecessors have done a lot of research work. As for the text sentiment analysis technology, there are sentiment classifications based on the number and difference of sentiment words. However, this method requires manual creation of an emotional dictionary, which requires a lot of manpower. There are other methods, such as literature [1] proposes a new method for extracting features from visual and text modalities, and then implemented sentiment analysis by providing the extracted features to a multi-core learning classifier. Literature [2] proposed a text sentiment analysis method combining the text representation of latent Dirichlet distribution and convolutional neural network. Literature [3] proposed a multi-modal learning algorithm for sentiment analysis through the relationship between discourses and dependencies. Literature [4] proposed a bidirectional GRU network model for text sentiment analysis based on attention mechanism. Literature [5] proposed a SER method based on tandem CNN and RNN for sentiment analysis. Literature [6] proposed a divide-and-conquer method. This method first divided sentences into different types, and then performed sentiment analysis on each type of sentence. Literature [7] proposed a CNN model with multiple functions. First analyze the characteristics of Weibo, and then introduce user-based features and content-based features into the CNN model to analyze the sentiment of Weibo events. Literature [8] uses support vector machines for sentiment analysis, specifically by finding support vectors to achieve sentiment classification. Literature [9] proposed a sentiment analysis algorithm. First, the last two layers of convolution in the CNN network were replaced with bidirectional LSTMs, and then sentiment classification was performed by using maximum pooling and fully connected layer output. Literature [10] uses Bi-LSTM to achieve sentiment classification. Specifically, it uses forward LSTM and backward LSTM to obtain context features in two directions, and then completes text sentiment classification. However, these methods can be improved in text semantic extraction to further improve the effect.

3 Implementation of Financial Sentiment Analysis Model Based on Pre-training and TextCNN (PTEAM)

The architecture of the financial sentiment analysis model based on pre-training and TextCNN is shown in Fig. 1. The model proposed in this thesis combines the pre-trained model with the TextCNN network to optimize text feature extraction, which can better extract the emotional features in financial text. First, use a pre-trained model to obtain a preliminary text semantic representation vector.

Then the obtained financial text semantic representation vector is input into the TextCNN network, and multiple dimensions of convolution kernels of different sizes are used to perform multi-dimensional feature extraction on financial text to achieve further extraction of emotional features in the text. Finally, after extracting the features, instead of using max pooling to select the maximum value of the obtained feature vectors, the obtained feature vectors are put into the same dimension and combined, so that the information captured before can not be lost. Finally, a classification layer is added to complete the emotional classification of financial text. Precision.

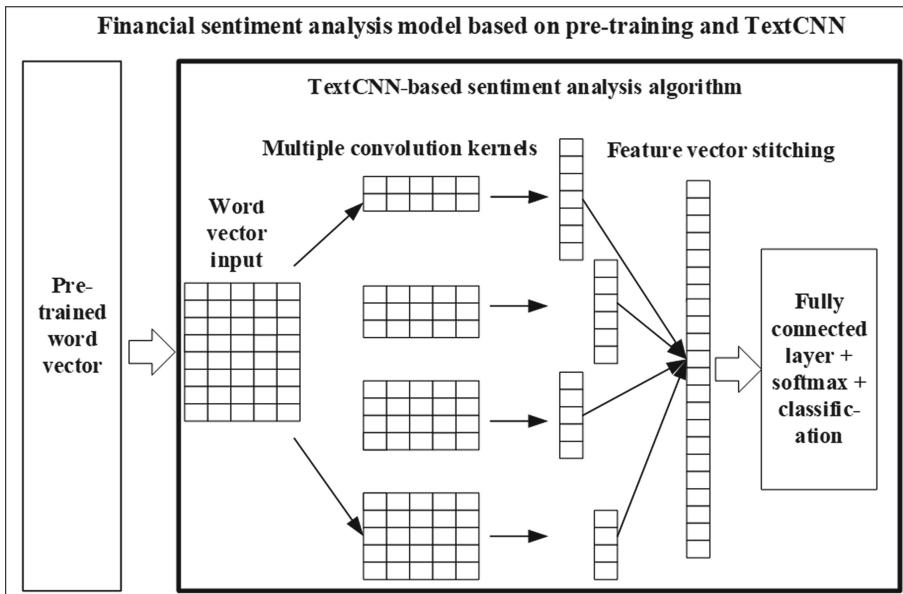


Fig. 1. Financial sentiment analysis model based on pre-training and TextCNN

3.1 Pre-trained Model BERT

Financial text semantic representation is obtained through a pre-trained model BERT. The BERT model incorporates a two-way Transformer and uses its encoding part for feature extraction. Different from previous models, it is to input the entire text sequence into the model once, rather than sequentially. Then learn the features of the entire text through two-way training, and each training can use all the feature information of the text. Unlike the one-way model, which can only utilize feature information in one direction at a time, BERT can learn deeper features than previous one-way models and perform better in experimental results. The overall architecture of the pre-trained model

BERT is shown in Fig. 2. First, a word vector input needs to be constructed. The word vector input is made up of three parts, including the word vector itself, segment representation and position representation. The word vector itself is a financial text word vector obtained by the word embedding technology. There is a lot of semantic information in it, and a CLS mark and a SEP mark are added. The CLS flag represents the beginning of the sentence, that is, E1, and can be used as the entire sentence vector, and can also be used for subsequent classification tasks. The SEP flag represents the end of the sentence, that is, En. Segment representation is used to distinguish two sentences, and position representation is a learned vector to distinguish between different positions.

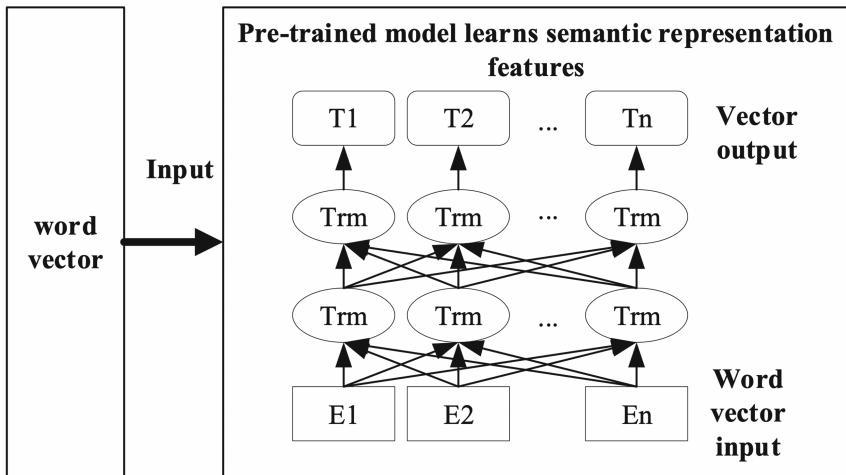


Fig. 2. The overall architecture of BERT

Then the fused word vector is input into the model for training. Specifically, some words in the input need to be masked first. Generally, 15% of the words in the sentence are masked, and then the context is used to predict the masked words. Through continuous input and training, you can get the output vector corresponding to each word. Finally, the financial text data is input into the BERT model to obtain the financial text word vector.

3.2 Improved TextCNN Algorithm

First, financial text data is input into a BERT pre-trained model to obtain financial text word vectors. Then treat each financial corpus as a sentence, and construct the financial text word vector into a matrix. The size of each matrix is fixed, the number of matrix columns is the dimension of the word vector, and the number of rows is the number of words that take the longest financial text. Then the obtained matrix is input into a TextCNN model, and semantic features in

financial text are extracted through a number of convolution kernels of different sizes. However, after extracting the features, instead of using max pooling to select the maximum value of the obtained feature vectors, the obtained feature vectors are put into the same dimension and combined, so that the information captured before can not be lost. Finally, a classification layer is added to output the emotional categories of the financial text. The specific activation function is shown in Eq. (1), the loss function used is shown in Eq. (2).

$$\sigma = \alpha(z)_i = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (1)$$

where $\alpha(z)_i$ is the softmax value, z_i refers to the current element, the activation value is between 0 and 1, and the sum is 1.

$$L = - \sum_{j=1}^N T_j \log \sigma \quad (2)$$

where σ_j is the class prediction value vector obtained through the softmax function, and T_j is the real value vector. L is cross entropy. The smaller the cross entropy, the more similar the two are. Generally, only one of T_j is 1, which is the probability of true classification, and the others are 0.

Use the above-mentioned sentiment analysis algorithm to perform sentiment analysis on financial texts. The sentiment categories are mainly divided into two categories: bullish and bearish. Bullish means that the financial text is positive. For example, the economic development is good and the stock market is good. And bearish means that the financial text is negative, contrary to the previous bullish.

4 Experiments

4.1 Dataset

The data set for sentiment analysis of financial texts is based on market evaluations made manually, and the sentiment tendencies are divided into two categories, bullish and bearish, and labeled. Build a balanced corpus based on 5000 market bullish news and 5000 market bearish news, as shown in Table 1.

Table 1. Experimental data of sentiment analysis

Financial text Dataset	Total	bullish	bearish
Training set	8000	4000	4000
Test set	2000	1000	1000
Total	10000	5000	5000

4.2 Experimental Evaluation and Analysis

The effectiveness of the model is measured through three different evaluation indicators: Precision is the proportion of true cases that are predicted to be positive (including true cases and false positive cases FP); The recall is the ratio of true cases to all positive cases (including true cases TP and false counter-examples FN) in this type of sample; The F1 is defined as the harmonic average of the precision and the recall. The calculation formula is as follows.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

In order to evaluate the sentiment analysis effect of the proposed model PTEAM on financial text data, the same data set is used, and several classic text sentiment analysis methods are selected for comparison experiments: SVM, it implements classification by finding support vectors; CRNN, which uses bidirectional LSTM to replace the last two layers of convolution in the CNN network for text feature extraction, and then uses maximum pooling and fully connected layer output for sentiment analysis; Bi-LSTM, using bidirectional LSTM to obtain context features in two directions, and then stitching vectors to achieve emotion classification; Using BERT and the original TextCNN to realize sentiment classification, thus verifying the effectiveness of this thesis to improve TextCNN.

As shown in Table 2, we can see that the PTEAM model proposed in this thesis performs best, and all indicators are improved compared to other algorithms. Precision and recall are a pair of contradictory values. Precision can measure the accuracy of model classification, while recall is a measure of whether the model can find all samples in that category. Therefore, the effect of the algorithm is best only when both values are relatively large. F1 is the balance parameter of the two. The overall effect of the SVM model is not good. The CRNN model structure uses bidirectional LSTM, which improves the network structure of CNN and improves the model. The overall performance is better than SVM. The Bi-LSTM model captures features in two directions, and uses the advantages of the LSTM structure to capture long-distance features, so the effect is improved. The model proposed in this thesis is the largest in all indicators. From the analysis of the algorithm structure, BERT is not limited by the distance between words in the sequence when processing the text sequence, and the relationship between any two words can be captured. However, other models have defects in this aspect: the SVM model cannot capture this relationship; the CRNN model only adds

a bidirectional LSTM unit to the last two layers of the model, and the capture features are not comprehensive enough; the Bi-LSTM LSTM unit improves the structure of the RNN, But when the distance between words is getting farther and farther, it is difficult to capture this feature; using BERT and TextCNN, you can better capture the connection between words, and feature vector stitching at the last layer of TextCNN, you can use More extracted information. Make the experimental effect further improved. The above experimental results show that the proposed method is better than previous models in dealing with sentiment analysis of financial texts, and proves the effectiveness and feasibility of the proposed method.

Table 2. Comparison of five algorithms in precision, recall, and F1

Algorithm	Precision	Recall	F1
SVM	82.36	81.73	82.04
CRNN	85.04	84.19	84.61
Bi-LSTM	86.83	84.92	85.86
BERT+TextCNN	87.18	86.75	86.96
PTEAM	88.12	88.28	88.20

As shown in Fig. 3, the precision of the five algorithms is compared, and the precision and average of the two categories of bullish and bearish are compared. It can be seen that the model proposed in this thesis is the best.

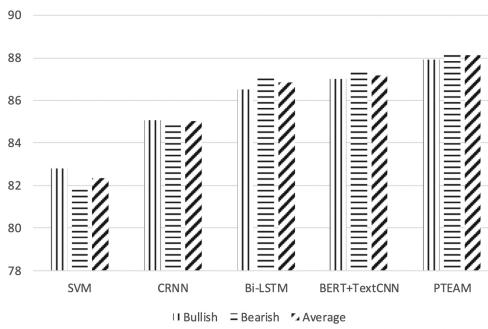


Fig. 3. Comparison of five algorithms in precision

As shown in Fig. 4, Fig. 5, the comparison of the five algorithms on recall and F1, respectively compared the two categories of indicators of bullish and bearish classification, we can see that the model proposed in this thesis is the best, reached about 88%, which is about 6% points higher on average than the

worst SVM algorithm. Because this model uses BERT for semantic extraction, it can better capture the connection between text contexts, and can ignore the distance limit, and then use TextCNN for further context semantic information extraction, which can better capture the text sentimental semantic information for emotional classification to divide the text into bullish and bearish.

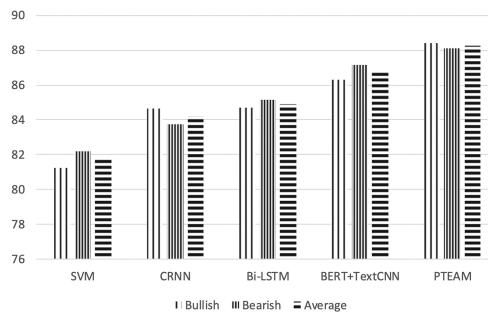


Fig. 4. Comparison of five algorithms in recall

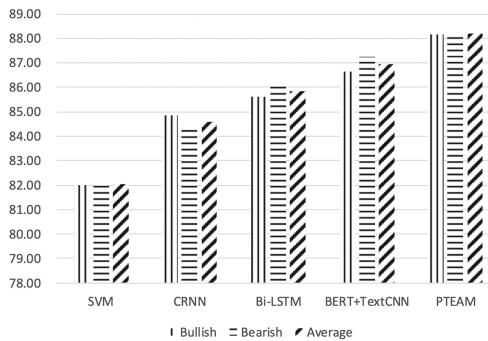


Fig. 5. Comparison of five algorithms in F1

5 Conclusions

The financial sentiment analysis model based on pre-training and TextCNN proposed in this thesis combines pre-trained models with TextCNN. First, the word vector is obtained through the word embedding model, and then a pre-trained model is introduced to further adjust the financial text word vector to obtain a word vector that can better express the semantics of financial text. Then use

TextCNN network to extract deeper information between financial text contexts. And improve the last layer of the TextCNN network, instead of using max pooling to take the maximum value of the convolutional feature vector, but stitching the feature vectors to get all the feature information. Finally, a fully connected softmax layer is added for sentiment classification, so that as much of the semantic information of financial text as possible can be used. The experimental results prove that its effect is better than the classic SVM model and LSTM model.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant (No. 61902037, No. 61532006, No. 61772083, No. 61802028), and Science and Technology Major Project of Guangxi (GuikeAA18118054).

References

1. Poria, S., Chaturvedi, I., Cambria, E., et al.: Convolutional MKL based multimodal emotion recognition and sentiment analysis. In: 2016 IEEE 16th International Conference on Data Mining (ICDM), pp. 439–448. IEEE (2016)
2. Luo, L.: Network text sentiment analysis method combining LDA text representation and GRU-CNN. *Pers. Ubiquit. Comput.* **23**(3–4), 405–412 (2019)
3. Lian, Z., Tao, J., Liu, B., et al.: Conversational Emotion Analysis via Attention Mechanisms. arXiv preprint [arXiv:1910.11263](https://arxiv.org/abs/1910.11263) (2019)
4. Hu, C., Xue, W.: Bi-GRU text emotion analysis model based on attention mechanism. *World Sci. Res. J.* **5**(9), 296–301 (2019)
5. Lim, W., Jang, D., Lee, T.: Speech emotion recognition using convolutional and recurrent neural networks. In: 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), pp. 1–4. IEEE (2016)
6. Chen, T., Xu, R., He, Y., et al.: Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Exp. Syst. Appl.* **72**, 221–230 (2017)
7. Shi, S., Zhao, M., Guan, J.U.N., et al.: Multi-features group emotion analysis based on CNN for Weibo events. *DEStech Trans. Comput. Sci. Eng.* (2017)
8. Xiang, L., Yibulayin, T., Kahaerjiang, A., et al.: Emotion analysis of active learning based on SVM in Uyghur language. *J. Xinjiang Univ.* (2015)
9. Fu, X., Ch, E., Aickelin, U., et al.: CRNN: A Joint Neural Network for Redundancy Detection. Social Science Electronic Publishing (2017)
10. Gui, X.X., et al.: Sentiment analysis of comment texts based on BiLSTM. *IEEE Access* **7**, 51522–51532 (2019)



Formation Terminal Sliding Mode Control for Multiple Mobile Robots

Lisen Gong¹ and Wuxi Shi^{2(✉)}

¹ Weifang Engineering Vocational College, Shandong 262500, China
qzgls@163.com

² School of Electrical Engineering and Automation, Tiangong University,
Tianjin 300387, China
shiwuxi@163.com

Abstract. Based on the leader-follower framework, a formation terminal sliding mode control method is presented for multiple-mobile robots. In this method, a virtual robot is introduced, the error dynamics between the follower and the virtual robot is established in the global coordinates, and the reasonable assignment of the error dynamics is used to design the formation terminal sliding mode controller. By using Lyapunov theory, it is proved that the tracking errors converge to the origin or the neighborhood of the origin in finite time. The experiments results show the effectiveness of the proposed method.

Keywords: Mobile robots · Formation controller · Leader-follower · Terminal sliding mode control

1 Introduction

In recent years, the mobile robot control technology has been developed a lot, however, due to the information acquisition ability, individual work ability and system stability, it is difficult to finish a given task in the case of complex and varied dynamic environment and complicated tasks by a single robot. Recently, the multi-robot formation system has attracted considerable attention. The multi-robot formation system is composed of a series of robots interacting with each other. By using partial information exchange, the multiple robots in the system compete with each other and cooperate with each other to perform a large number of complex tasks that are difficult for a single robot to accomplish. At present, the multi-robot formation methods mainly include: leader-follower method [1–3], behavior-based method [4], artificial potential field method [5], virtual structure method [6] and graph theory method [7]. Among them, the behavior-based method has better parallelism, so that the information exchanged between the robots is less. However, it is difficult to analyze its behavior mathematically. The artificial potential field method is computationally small, but it is difficult to design the controller to meet the local minimum potential function. Virtual structure method is easy to describe cooperative behavior in formation, however,

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 57–65, 2021.

https://doi.org/10.1007/978-981-15-8458-9_7

formation shape needs to maintain the same rigid structure, lack of flexibility and adaptability. Graph theory is easier to describe arbitrary formations, however, mainly limited to simulation research.

Since the leader-follower method has a distribution structure, multiple robots formation problems can be simplified to the trajectory tracking control problem of several robots. The controller design in this method mainly include robust control [8], model predictive control [9], feedback linearization [10], backstepping method [11, 12], sliding mode control [13] and so on. In [8], a robust controller for fixed-wing aircraft formation was designed to avoid obstacles and has a good formation effect. However, the design process was complicated, and this scheme was difficult to verify by physical objects. In [9], probability density function was used to model the task space, and multiple mobile robots were coordinated to cover an event, however, the computation burden online was large. In [10], the formation model in polar coordinates was designed, and the feedback controller was constructed. The calculation was simple and the position guarantees that the position error converges to zero, but the attitude error converges to a constant value affecting the formation effect. In [11, 12], the controller was established for the formation stability problem of multi-robot with uncertain dynamics. In [13], the sliding mode controller was designed for the cooperation of multiple robots, however, the tracking errors in finite time was not considered. In [14], a continuous fractional order control method based on fast terminal sliding mode controller was proposed, and the tracking error can converge to origin in a limited time.

In this paper, a formation control method is proposed for multi-mobile robots. In this method, the direction of motion is guided by the leader, and the reference trajectory is generated by the virtual robot. The tracking error system is constructed by the virtual robot and the actual follower, and the formation controller is designed. By using Lyapunov theory and fast terminal sliding mode control, it is proved that the tracking error converge to the origin or the neighborhood of the origin in finite time. Experiments show the effectiveness of the proposed method.

2 System Description

In this paper, we consider a differential-driving mobile robot whose kinematics is given by

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 \\ \sin \theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} \quad (1)$$

where (x, y) and θ are the position and the orientation of the mobile robot with respect to an inertial coordination frame, and $v(t)$ and $w(t)$ are the linear velocity and angular velocity, respectively.

For a leader-follower formation system, the main trajectory is usually determined by the leader, and the reference trajectory of the follower is determined by the virtual robot generated by the leader and the structural parameters. In this

paper, a simple triangular formation in Fig. 1 is used, in which the coordinates of the leader, the follower, and the virtual robot are $[x_L, y_L, \theta_L]^T$, $[x_F, y_F, \theta_F]^T$, and $[x_V, y_V, \theta_V]^T$, respectively. The line velocity and angular velocity of the leader are v_L and w_L , and those of the follower are v_F and w_F . λ_{L-F}^d and ϕ_{L-F}^d are the distance and the orientation of the virtual robot respect to the leader.

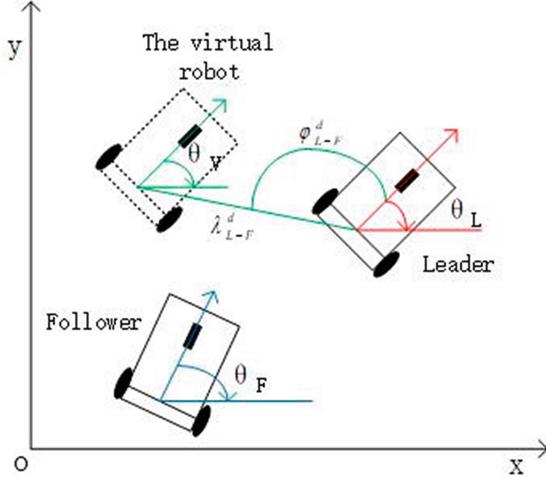


Fig. 1. Simple triangle formation of three robot

Assumption 1 [15]: The orientation of the virtual robot is the same as that of the leader, that is $\theta_V = \theta_L$.

Let the leader as a reference, from Fig. 1 we have

$$\begin{aligned} x_V &= x_L + \lambda_{L-F}^d \cos(\phi_{L-F}^d + \theta_L) \\ y_V &= y_L + \lambda_{L-F}^d \sin(\phi_{L-F}^d + \theta_L) \\ \theta_V &= \theta_L \end{aligned} \quad (2)$$

Based on (1) and using (2), we have

$$\begin{aligned} \dot{x}_V &= v_L \cos \theta_L - w_L \lambda_{L-F}^d \sin(\phi_{L-F}^d + \theta_L) \\ \dot{y}_V &= v_L \sin \theta_L + w_L \lambda_{L-F}^d \cos(\phi_{L-F}^d + \theta_L) \\ \dot{\theta}_V &= w_L \end{aligned} \quad (3)$$

Define the errors between the virtual robot and the follower as $x_e = x_V - x_F$, $y_e = y_V - y_F$, $\theta_e = \theta_V - \theta_F$, using (1) and (3), we obtain the tracking errors dynamics as

$$\begin{aligned} \dot{x}_e &= v_L \cos \theta_L - w_L \lambda_{L-F}^d \sin(\phi_{L-F}^d + \theta_L) - v_F \cos \theta_F \\ \dot{y}_e &= v_L \sin \theta_L + w_L \lambda_{L-F}^d \cos(\phi_{L-F}^d + \theta_L) - v_F \sin \theta_F \\ \dot{\theta}_e &= w_L - w_F \end{aligned} \quad (4)$$

We know that the formation problem of multi-robot system is to make the follower track the virtual robot, thus, the objective of this paper is to design a controller $(v_F, w_F)^T$ such that the tracking errors x_e, y_e, θ_e converge to zero in a limited time.

3 Controller Design

In order to design a fast terminal sliding mode controller, the following lemmas are needed:

Lemma 1. [16] Suppose a_1, a_2, \dots, a_n and $0 < p < 2$ are all positive numbers, then the following inequality holds:

$$(a_1^2 + a_2^2 + \dots + a_n^2)^p \leq (a_1^p + a_2^p + \dots + a_n^p)^2 \quad (5)$$

Lemma 2. [16] If a Lyapunov function $V(x)$ satisfies the following inequality

$$\dot{V}(x) \leq -\alpha V(x) - \beta V^\gamma(x) \quad (6)$$

where $\alpha > 0, \beta > 0, 0 < \gamma < 1$, then the settling time given by

$$T \leq \frac{1}{\alpha(1-\gamma)} \ln \frac{\alpha V^{1-\gamma}(x_0) + \beta}{\beta} \quad (7)$$

Let

$$v_F \cos \theta_F = v_L \cos \theta_L + c_1 x_e + h_1 |x_e|^r \text{sign}(x_e) := A_x \quad (8)$$

$$v_F \sin \theta_F = v_L \sin \theta_L + c_2 y_e + h_2 |y_e|^r \text{sign}(y_e) := A_y \quad (9)$$

where $0 < r < 1, h_1 > 0, h_2 > 0, c_1 > a_1, c_2 > a_2$ and a_1, a_2 are positive constants designed later.

Using (8) and (9) we obtain

$$v_F = \sqrt{A_x^2 + A_y^2} \quad (10)$$

$$\theta_F = \arctan \frac{A_y}{A_x} \quad (11)$$

From (11) yields

$$w_F = \dot{\theta}_F = \frac{(c_2 + h_2 r y_e^{r-1}) \dot{y}_e A_x - (c_1 + h_1 r x_e^{r-1}) \dot{x}_e A_y}{A_x^2 + A_y^2} \quad (12)$$

Theorem 1. For the error dynamics (4), if $|\theta_e(0)| < \pi$ is selected, and the control laws are designed as (10) and (12), then, (1) If the leader's trajectory is linear, the tracking errors converge to zero in limited time, and (2) if the leader's trajectory is curve, the tracking errors converge to the neighborhood of the origin in limited time.

Proof. Substituting (8) and (9) into (4), we have

$$\dot{x}_e = -c_1 x_e - w_L \lambda_{L-F}^d \sin(\varphi_{L-F}^d + \theta_L) - h_1 |x_e|^r \text{sign}(x_e) \quad (13)$$

$$\dot{y}_e = -c_2 y_e + w_L \lambda_{L-F}^d \cos(\varphi_{L-F}^d + \theta_L) - h_2 |y_e|^r \text{sign}(y_e) \quad (14)$$

Consider the following Lyapunov function candidate

$$V = \frac{1}{2} x_e^2 + \frac{1}{2} y_e^2 \quad (15)$$

Differentiating (15) with respect to time yields

$$\dot{V} = x_e \dot{x}_e + y_e \dot{y}_e \quad (16)$$

Substituting (13) and (14) into (16), one has

$$\begin{aligned} \dot{V} &= -c_1 x_e^2 - c_2 y_e^2 - h_1 |x_e|^{r+1} - h_2 |y_e|^{r+1} - w_L \lambda_{L-F}^d x_e \sin(\varphi_{L-F}^d + \theta_L) \\ &\quad + w_L \lambda_{L-F}^d y_e \cos(\varphi_{L-F}^d + \theta_L) \\ &\leq -c_1 x_e^2 - c_2 y_e^2 - h_1 |x_e|^{r+1} - h_2 |y_e|^{r+1} + |x_e| w_L \lambda_{L-F}^d + |y_e| w_L \lambda_{L-F}^d \end{aligned} \quad (17)$$

After some manipulations, (17) can be further rewritten as follows

$$\begin{aligned} \dot{V} &\leq -(c_1 - a_1)x_e^2 - (c_2 - a_2)y_e^2 - h_1 |x_e|^{r+1} - h_2 |y_e|^{r+1} + \left(\frac{1}{4a_1} + \frac{1}{4a_2} \right) (w_L \lambda_{L-F}^d)^2 \\ &\quad - (\sqrt{a_1} \cdot |x_e| - \frac{1}{2\sqrt{a_1}} w_L \lambda_{L-F}^d)^2 - (\sqrt{a_2} \cdot |y_e| - \frac{1}{2\sqrt{a_2}} w_L \lambda_{L-F}^d)^2 \\ &\leq -(c_1 - a_1)x_e^2 - (c_2 - a_2)y_e^2 - h_1 |x_e|^{r+1} - h_2 |y_e|^{r+1} + \zeta \\ &\leq -c(x_e^2 + y_e^2) - h(|x_e|^{r+1} + |y_e|^{r+1}) + \zeta \end{aligned} \quad (18)$$

where $c = \min\{c_1 - a_1, c_2 - a_2\}$, $h = \min\{h_1, h_2\}$, $\zeta = (\frac{1}{4a_1} + \frac{1}{4a_2})(w_L \lambda_{L-F}^d)^2$.

According to Lemma 1, we can get $(x_e^2 + y_e^2)^{\frac{r+1}{2}} \leq |x_e|^{r+1} + |y_e|^{r+1}$, then (18) can be expressed as

$$\dot{V} \leq -\alpha_1 V - \beta_1 V^\gamma + \zeta \quad (19)$$

where $\alpha_1 = 2c$, $\beta_1 = h2^{\frac{1+r}{2}}$, $\gamma = \frac{r+1}{2}$.

(1). If the leader's trajectory is linear, that is $w_L = 0$, in this case, (19) can be written as

$$\dot{V} \leq -\alpha_1 V - \beta_1 V^\gamma \quad (20)$$

Since $0 < r < 1$, we have $0 < \gamma < 1$, according to Lemma 2, the tracking error x_e, y_e can converge to zero in a finite time.

From (11) we get

$$\sin(\theta_F - \theta_L) = \frac{(-c_1 x_e - h_1 |x_e|^r \text{sign}(x_e)) \sin \theta_F}{v_L} + \frac{(c_2 y_e + h_2 |y_e|^r \text{sign}(y_e)) \cos \theta_F}{v_L} \quad (21)$$

Since x_e, y_e converge to zero in finite time, from (22) we obtain that $\sin(\theta_F - \theta_L)$ can be reached in finite time, it follow that $\theta_F - \theta_L$ converge to zero in finite time, and then θ_e can converge to zero in finite time.

(2). If the trajectory of the leader is a curve, that is $w_L \neq 0$, in this case, (19) becomes as

$$\dot{V} \leq -(\alpha_1 - \frac{2\zeta}{x_e^2 + y_e^2})V - \beta_1 V^\gamma \quad (22)$$

or

$$\dot{V} \leq -\alpha_1 V - (\beta_1 - \frac{\zeta}{V^\gamma})V^\gamma \quad (23)$$

From (22), if we can keep $\alpha_1 - \frac{2\zeta}{x_e^2 + y_e^2} > \bar{\alpha}_1 > 0$, where $\bar{\alpha}_1$ is a constant, then

$$\dot{V} \leq -\bar{\alpha}_1 V - \beta_1 V^\gamma \quad (24)$$

According to Lemma 2, the tracking errors x_e and y_e converge to the regions

$$|x_e| \leq \sqrt{\frac{2\zeta}{\alpha_1 - \bar{\alpha}_1}}, |y_e| \leq \sqrt{\frac{2\zeta}{\alpha_1 - \bar{\alpha}_1}} \quad (25)$$

in infinite time.

By using similar way for (23) and Lemma 2, we can obtain that the tracking errors x_e and y_e converge to the regions

$$|x_e| \leq 2 \sqrt{\frac{\zeta}{\beta_1 - \bar{\beta}_1}}, |y_e| \leq 2 \sqrt{\frac{\zeta}{\beta_1 - \bar{\beta}_1}} \quad (26)$$

in infinite time, where $\bar{\beta}_1$ satisfy $\beta_1 - \frac{\zeta}{V^\gamma} > \bar{\beta}_1 > 0$.

From (25) and (26), we have that the regions

$$|x_e| \leq \min\left\{\sqrt{\frac{2\zeta}{\alpha_1 - \bar{\alpha}_1}}, 2 \sqrt{\frac{\zeta}{\beta_1 - \bar{\beta}_1}}\right\}, |y_e| \leq \min\left\{\sqrt{\frac{2\zeta}{\alpha_1 - \bar{\alpha}_1}}, 2 \sqrt{\frac{\zeta}{\beta_1 - \bar{\beta}_1}}\right\} \quad (27)$$

can be reached infinite time. From (21), θ_e can be reached to the neighborhood of the origin in finite time.

4 Experiment

In order to verify the effectiveness of the proposed method, we use a three-mobile robot formation system as shown in Fig. 2. The main part of robot is a rectangular plastic base plate. The driving wheel and the directional wheel form a solid triangle at the bottom of the base plate. There are also motors and encoders at the bottom of the car. Power supply module, wireless module, control module and voltage conversion module are installed above the car body. DC motor is used in the driving part of wheeled robot, and the duty cycle of pulse width modulation (PWM) is used to control the voltage of the motor, so as to control the speed of the motor. Wireless transceiver uses ZigBee module to realize the coordinated control of multi-mobile robots. The robot and robot interact with each other to transmit their own posture, position and other information, and then adjust their own behavior. In the experiment, the leader's ZigBee module

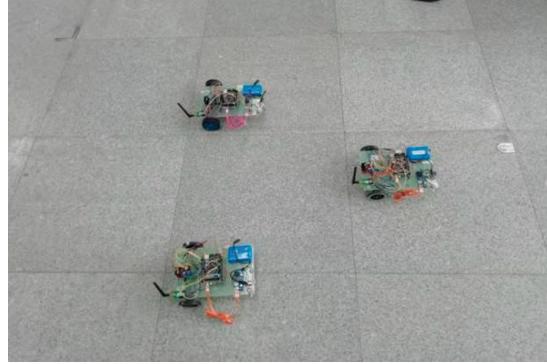


Fig. 2. Experimental platform

has the function of real-time sending its own position and attitude information, while the follower's ZigBee module is responsible for receiving the leader's information. Through the algorithm designed in this paper, the desired position and attitude can be obtained, and then the speed can be adjusted. The position and velocity information of the three robots are collected through serial port.

In this paper, two formation experiments are provided, $a_1 = a_2 = 0.01$.

Experiment 1: Three mobile robots move in a straight line from the starting position and form a triangular formation. Leader's initial linear velocity and angular velocity are set to $v_L = 0.1 \text{ m/s}$, $w_L = 0 \text{ rad/s}$, the initial posture is $(2, 2\pi/4)$, and initial posture of two followers are $(1, 0, \pi/4)$ and $(0, 1, \pi/4)$. $\lambda_{L-F1}^d = 1, \varphi_{L-F1}^d = 5\pi/6$ and $\lambda_{L-F2}^d = 1, \varphi_{L-F2}^d = -5\pi/6$.

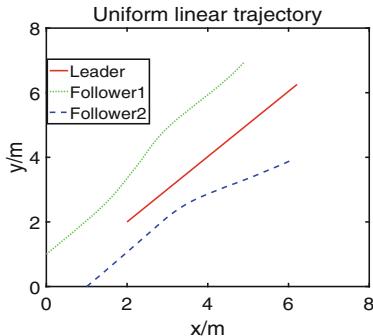


Fig. 3. Trajectory of linear motion

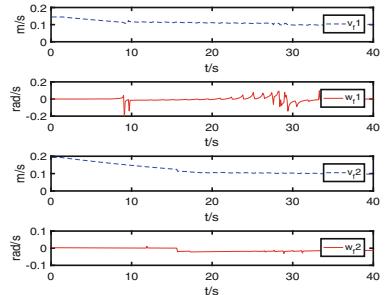


Fig. 4. Velocity curve of followers for experiment 1

The controller parameters of followers 1 and 2 are: $c_1 = c_2 = c_3 = c_4 = 0.03, h_1 = h_2 = 0.005, h_3 = h_4 = 0.001, r_1 = 0.08, r_2 = 0.05$.

The trajectories of three mobile robots are shown in Fig. 3. The velocity curves of the two followers are shown in Fig. 4, in which the blue dotted line and the red solid line are the follower's linear velocity and angular velocity, respectively.

Experiment 2: Three mobile robots move along curves from their initial positions and form linear formation. Leader's initial linear velocity and angular velocity are set to $v_L = 0.1 \text{ m/s}$, $w_L = 0.1 \text{ rad/s}$, the initial posture is $(2, 2, 0)$, initial posture of two followers are $(1, 2, 0)$ and $(1, 1, 0)$. $\lambda_{L-F1}^d = 0.5$, $\varphi_{L-F1}^d = \pi/2$ and $\lambda_{L-F2}^d = 0.5$, $\varphi_{L-F2}^d = -\pi/2$.

The controller parameters of followers 1 and 2 are: $c_1 = c_2 = 0.002$, $c_3 = c_4 = 0.03$, $h_1 = h_2 = h_3 = h_4 = 0.002$, $r_1 = 0.08$, $r_2 = 0.05$.

The trajectories of three mobile robots are shown in Fig. 5 and the velocity curves of two followers are shown in Fig. 6.

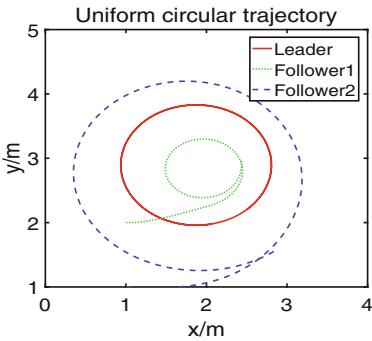


Fig. 5. Trajectory of circular motion

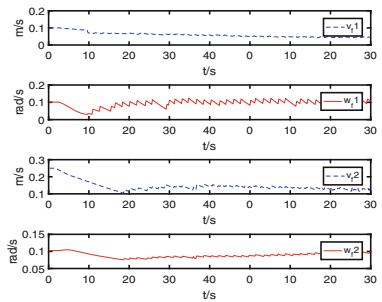


Fig. 6. Velocity curve of followers for experiment 2

It can be seen that the two followers start to move from the given initial position, and can follow the leader to form a triangle and a linear formation structure, respectively. It is pointed out that due to the influence of hardware factors such as the transmission friction of the motor, the battery voltage and the signal delay, the robot is disturbed in the process of movement, and the velocity curve is not very smooth.

5 Conclusion

A formation terminal sliding mode control method has been presented for multi-mobile robots. In this method, the pose parameters of the virtual robot are determined by the leader and expected formation, and the formation problem of the leader-follower method is transformed into the trajectory tracking problem of the virtual robot by the follower. Then, the formation controller is designed according to the characteristics of the kinematics model and the fast terminal sliding mode, and the tracking error can converge to the origin or the neighborhood of the origin in finite time.

References

1. Xiao, H.Z., Li, Z.J., Chen, P.: Formation control of leader-follower mobile robots' systems using model predictive control based on neural-dynamic optimization. *IEEE Trans. Ind. Electr.* **63**(9), 5752–5762 (2016)
2. Miao, Z.Q., Liu, Y.H., Wang, Y.N., et al.: Distributed estimation and control for leader-following formations of nonholonomic mobile robots. *IEEE Trans. Autom. Sci. Eng.* **15**(4), 1946–1954 (2018)
3. Jia, Y.N., Wang, L.: Leader-Follower flocking of multiple robotic fish. *IEEE/ASME Trans. Mechatron.* **20**(3), 1372–1383 (2015)
4. Lin, J.L., Hwang, K.S., Wang, Y.L.: A simple scheme for formation control based on weighted behavior learning. *IEEE Trans. Neural Netw. Learn. Syst.* **25**(6), 1033–1044 (2014)
5. Ranjith, R.N., Laxmidhar, B., Vinod, K., et al.: Multisatellite formation control for remote sensing applications using artificial potential field and adaptive fuzzy sliding model control. *IEEE Syst. J.* **9**(2), 508–518 (2015)
6. Zhou, D.J., Wang, Z.J., Mac, S.: Agile coordination and assistive collision avoidance for quadrotor swarms using virtual structures. *IEEE Trans. Robot.* **34**(4), 916–923 (2018)
7. Zhang, L., Lu, Y., Xu, S., et al.: Multiple UAVs cooperative formation forming control based on back-stepping-like approach. *J. Syst. Eng. Electron.* **29**(4), 816–822 (2018)
8. Yang, A., Naeen, W., George, W., et al.: Stability analysis and implementation of a decentralized formation control strategy for unmanned vehicles. *IEEE Trans. Control Syst. Technol.* **22**(2), 706–720 (2014)
9. Fatemen, M., Ali, D., Mohammad, B.M.: Distributed receding horizon coverage control for multiple mobile robots. *IEEE Syst. J.* **10**(1), 198–207 (2016)
10. Das, A.K., Fierro, R., Kumar, V., et al.: A vision-based formation control framework. *IEEE Trans. Robot. Autom.* **8**(5), 813–825 (2002)
11. Cai, X.Y., Marcio, D.Q.: Adaptive rigidity-based formation control of uncertain multi-robotic vehicles. In: 2014 American Control Conference, pp. 293–298 (2014)
12. Cai, X.Y., de Queiroz, M.: Adaptive rigidity based formation control for multi-robotic vehicles with dynamics. *IEEE Trans. Control Syst. Technol.* **23**(1), 389–396 (2015)
13. Gian, P.I., Gianluca, D.F., Antonella, F., et al.: A supervisory sliding mode control approach for cooperative robotic system of systems. *IEEE Syst. J.* **9**(1), 263–272 (2015)
14. Wang, Y.Y., Gu, L.Y., Xu, Y.H., et al.: Practical tracking control of robot manipulators with continuous fractional-order nonsingular terminal sliding mode. *IEEE Trans. Ind. Electron.* **63**(10), 6194–6204 (2016)
15. Shao, J., Xie, G., Wang, L.: Leader-following formation control of multiple mobile vehicles. *IET Control Theor. Appl.* **1**(2), 545–552 (2007)
16. Yu, S., Yu, X., Shirinzadeh, B., Man, Z.: Continuous finite-time control for robotic manipulators with terminal sliding mode. *Automatica* **41**, 1957–1964 (2005)



Mining Correlation Features of User Financial Behavior Based on Attention Mechanism and Dual Channel

Pengchao Cheng¹, Yawen Li², and Zhe Xue^{1(✉)}

¹ Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia, School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

xuezhe@bupt.edu.cn

² School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. For the financial field, the amount of data is large and the data forms are complicated. Traditional data processing methods are unsuitable for current big data. Therefore, we propose a financial big data feature mining method based on attention mechanism and dual channel (NDUA) to understand the deep features of financial big data and mine the association features of customers' financial behavior. First, the pre-trained model Bert is used to vectorize textual data in financial data. Next, a dual-tower model is proposed to predict user behavior, and multiple user behaviors are used as labels. Then, we perform model training and use the trained three sets of models to extract the last fully connected layer as the user behavior feature of the item, and for each user to cluster the user behavior feature representations of all the operated items to further obtain the association of user behavior. Experiments demonstrate that the proposed method can effectively understand the deep features of financial big data and mine the association features of customers' financial behavior.

Keywords: Association features · Behavior prediction · Attention mechanism · Pre-trained model

1 Introduction

With the development of information technology, data in various fields are exploding. For the financial field, the amount of data is large and the data forms are complicated, but the data processing, analysis and features mining are all indispensable operations. The traditional method for processing text data into text features is to use the TF-IDF algorithm to establish the SVM space [1]. In principle, the feature of text is represented based on the bag of words model. It can pass a small number of high frequent vocabulary to represent text, while it only considers the dictionary position of the text and does not consider the

order of each word in the sentence, which reduces the semantics in the vectors. With the development of deep learning, the famous word2vec [2] is developed. Word2vec uses a large amount of text data to pre-train words in text based on language models and obtained a fixed-length word vector.

The Factorization Machines (FM) [3] is a traditional method for predicting user behavior. FM not only considers the first-order features but also can obtain any order of feature interactions. However, due to its complexity, only first-order and second-order interaction features are applied in practice. Field-aware Factorization Machines (FFM) [4] is an upgraded model of FM. In FFM, each dimension feature, for each field, will learn a hidden vector. The hidden vector is related to both the feature and the field. FNN [5] is to solve the problem that FM and FFM only consider first-order and second-order features. Since the low-order features and high-order combined features are of the same importance, the Wide&Deep [6] model is proposed. The wide part uses low-order features to memorize the frequent correlation of combined features in user data, and the deep part uses deep neural networks to learn high-order combinations between features to achieve generalization of data. However, complex artificial feature engineering and experienced engineers are required in the Wide&Deep method. DeepFM [7] solves the complex artificial feature operation in Wide&Deep, which truly realizes end-to-end user behavior prediction.

In this paper, we propose a method of financial big data feature mining (user behavior association feature mining) based on natural language understanding and deep learning to perform deep feature understanding of financial big data and mine customer financial action association features. First, for textual information in financial big data, the pre-trained Bert [8] model is used to obtain the embedding of the textual data as one of the model inputs. Next, the twin tower model is used to predict user behavior. In light of multi-task learning, using multiple user behaviors as labels to obtain the training model. Then, we use the trained three groups of models to extract the last fully connected layer as the user behavior characteristics of the item. For each user, the user behavior feature representations of all operated items are clustered to further obtain the associations of user behaviors. Finally, experiments demonstrate that the proposed method can understand the deep characteristics of financial big data and mine the customer's financial behavior association features.

2 The Proposed User Behavior Prediction Algorithm Based on Attention Mechanism and Dual Channel (NDUA)

The framework of user behavior prediction method based on attention mechanism and dual channel is shown in Fig. 1. The framework consists of two parts. The left side shows FM is used to obtain the low-order cross-combined features of discrete features and continuous features to memorize the high-frequency cross-over features in the data. On the right side, the features of FM are trained

through the Attention module to obtain which features are effective for the construction of higher-order features. The obtained feature vectors are connected to the text vector as the input of the highway network to obtain higher-order combinations of features. After combining low-order features and high-order features, extraction, connect the results as the input of the shallow fully connected network, and finally get the prediction result of user behavior.

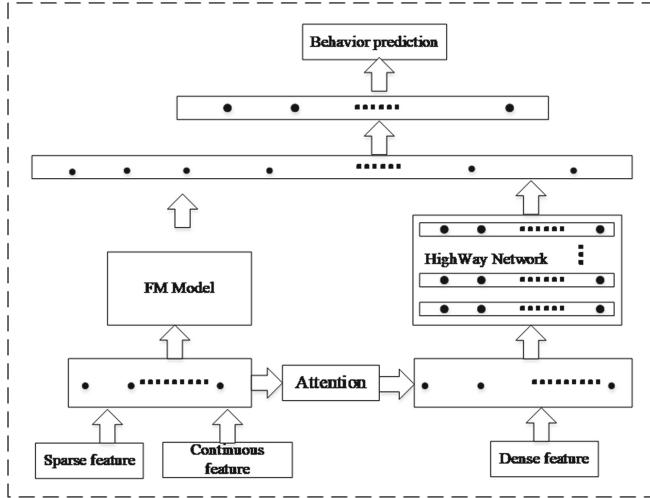


Fig. 1. The framework of NDUA

2.1 Low-Order Feature Extraction Based on Factorization Machine

FM mainly uses the inner product of pairwise features as the coefficient of second-order feature crossover, thereby solving the problem of sparse data features such as user id, item id that may have millions of classification features after one-hot encoding. It not only considers the first-order features, but also may obtain any kind of feature interaction. However, due to its complexity, we usually only apply the first-order features and second-order interaction features. Because it is indistinguishable from each one-dimensional feature after one-hot, field information is introduced into FM. For features that belong to the same field, only one-dimensional features work at a time.

We only use the second-order feature combination, Eq. (1) is a general form of the factor decomposition machine FM.

$$y = \theta_0 + \sum_{i=1}^n \theta_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \theta_{ij} x_i x_j \quad (1)$$

Where θ_0 represents a bias parameter, θ_i represents a parameter of a first-order feature, θ_{ij} represents a parameter of a second-order combined feature, x_i represents the i -th feature, and x_j represents the j -th feature.

Equation (2) indicates that there is a hidden vector for each feature of the FM, and the inner product of the pairwise hidden vectors is used to represent the coefficient where the pairwise features cross.

$$y = \theta_0 + \sum_{i=1}^n \theta_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j \quad (2)$$

Where v_i and v_j respectively represent the i -th feature, the hidden vector corresponding to the j -th feature, and other representations have the same meaning as in Eq. (1).

Equation (3) represents a simple expression after the optimization of the quadratic term of FM. Compared to Eq. (2), its complexity is reduced from the original $O(n^2)$ to the current $O(n)$.

$$\begin{aligned} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j &= \frac{1}{2} \sum_{f=1}^k \left(\sum_{i=1}^n v_{if} x_i \sum_{j=1}^n v_{jf} x_j - \sum_{i=1}^n v_{if} x_i \right) \\ &= \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{if} x_i \right) - \sum_{i=1}^n v_{if}^2 x_i^2 \right) \end{aligned} \quad (3)$$

Where v_{ij} and v_{jf} respectively represent the i -th dimension of the f -th feature and the f -th dimension of the j -th feature. The other parameters are the same as those in Eq. (1).

2.2 Feature Extraction of Deep Combination Based on Highway Network

The inspiration of highway network comes from solving the problem of RNN, the proposed LSTM structure is to join the “gate” structure. The problem that highway network mainly solves is that the depth of the network is deepened, and the gradient at the deep layer may be 0, which makes the network parameters at the shallow layer update slowly or not update and the network is difficult to train. Highway network uses a principle similar to the gate mechanism of LSTM to solve the problem of vanishing gradients, thereby speeding up network convergence.

$$y = f(x, w) \cdot T(x, u) + x \cdot (1 - T(x, u)) \quad (4)$$

Where $T(x, u)$ represents the conversion gate, and its specific function is to control how much new information is retained as the final output, and $1 - T(x, u)$ represents how much original information is retained as the final output. The dimensions of x , y , and $T(x, u)$ must be the same. $f(x, w)$ represents the non-linear transformation of the input vector x , and u , w represent the model parameters.

2.3 Semantic Feature Extraction Based on Attention Mechanism

Attention mechanism mimics the internal process of biological observation behavior, that is, a mechanism that aligns internal experience and external sensation to increase the observation precision of some areas. For example, when human vision processes an image, it quickly scans the global image to obtain the target area that needs to be focused on, that is, the focus of attention. Then invest more attention resources in this area to get more detailed information about the target that needs attention, and suppress other useless information. The attention mechanism is first proposed in the image field. It is inspired by the fact that humans focus on one part instead of the whole when observing things. At present the attention mechanism is also widely used in NLP applications.

In this paper, the highway network module is used to obtain the high-order combination results of features, and the final user behavior prediction is performed through low-order combination features and high-order combination features. However, in the process of behavior prediction, not all data features have a positive impact on the prediction result. For the one-dimensional feature and the two-dimensional cross feature, the degree of influence on the prediction of user behavior is different, so we use attention mechanism to obtain one-dimensional and two-dimensional features which are strong features for predicting user behavior.

There are three basic elements of the general model of the attention mechanism: query, key, and value. Generally, keys and values are the same. In the attention calculation, a weight function is used to calculate the weight between the query and all keys, which is assigned to each value as a weight, and the weighted sum of all values is returned as the fusion context feature vector.

When understanding the attention mechanism, the calculation of representation is the core of the attention mechanism. The difference between different attention mechanisms is the weight function. The three commonly used weight functions are inner product, multiplication and stitching. The calculation formula is shown in Eq. (5).

$$\text{score}(\text{query}, \text{key}) = \begin{cases} \text{query}^T \text{key} \\ \text{query}^T W_s \text{key} \\ v^T \tanh(W_s[\text{query}; \text{key}]) \end{cases} \quad (5)$$

Where v and W_s are model parameters, and $[;]$ is a vector splicing.

The extra vectors are used to find the most important features of all first- and second-order combined features, and finally a weighted sum of all vectors is used as an input to the highway network.

3 User Behavior Association Feature Mining

There are potential associations between user behavior clicks, favorites, purchases and other behaviors. The user behavior feature of the item obtained in 2 parts is used as the input of user behavior association feature mining. For all items, the behaviors of all users are clustered to get the association between the behaviors.

The key point in the clustering algorithm is the similarity calculation method between data. Commonly used similarity calculation methods include cosine similarity, Euclidean distance, and KL divergence. For example, the formula for calculating the similarity of the text using the cosine similarity is as follows Eq. (6):

$$s(d_i, d_j) = \frac{\sum_{l=1}^n d_{il}d_{jl}}{\sqrt{\sum_{l=1}^n d_{il}^2} \times \sqrt{\sum_{l=1}^n d_{jl}^2}} \quad (6)$$

Where $s(d_i, d_j)$ represents the calculation of the similarity between the i -th text d_i and j -th text d_j , and d_{il} and d_{jl} represent the l -th component and the j -th component of the text feature vector.

4 Experiments

4.1 Experimental Dataset

The experimental data are the operations of Tiantian Fund users in 2018. User operations include viewing funds, collecting funds, and purchasing funds. Viewing fund operations means that the user clicks the fund on the interface to query basic information. Collecting funds means that the user spends a certain amount of time on getting to know the fund and considering it as the fund for next purchase. Buying a fund means that the user has arrived to purchase. A detailed description of the financial behavior data set of Tiantian Fund users is shown in Table 1.

Table 1. Tiantian fund dataset

	View	Collect	Purchase
Train	8768	6122	4218
Valid	1500	1500	1500
Test	1000	1000	1000

4.2 Evaluation Metric

For the behavior prediction of user behavior data, the accuracy index, recall rate, and F1 are generally used as indicators for evaluating the prediction result. In actual scenarios, we also use AUC to indicate the effectiveness of the model.

4.3 Performance Analysis

User behavior prediction experiment process: The three user behaviors in the data set are used to train the corresponding three models to predict the user's viewing behavior, collection behavior, and purchase behavior of items. In order

Table 2. AUC of NDUA algorithm and comparison algorithm in predicting three user's behavior

Algorithm	AUC		
	View	Collect	Purchase
FM	0.753	0.722	0.711
Wide&Deep	0.768	0.741	0.735
DeepFM	0.775	0.754	0.744
NDUA	0.781	0.763	0.752

to verify the effectiveness of NDUA, the shallow model FM, the deep models Wide&Deep and DeepFM which are currently popular in the industry are used as comparative experiments.

It can be seen from Table 2 that NDUA is superior to the traditional FM algorithm, the Wide&Deep algorithm, and the DeepFM algorithm based on deep learning in the AUC. When predicting user's viewing behavior, the NDUA algorithm is 0.86% higher on AUC than DeepFM, indicating that the NDUA algorithm can rank the data viewed by the user to the head of the list. When predicting user's collecting behavior, the NDUA is 0.9% higher in AUC than the Wide&Deep algorithm. The low-order feature extraction of the Wide&Deep algorithm is performed by artificial feature engineering. In contrast, NDUA uses FM to automatically perform low-order features. Extraction is better than artificial feature engineering in effect. When predicting user's purchasing behavior, NDUA is 1.73% higher than the Wide&Deep in precision and 1% higher than the Wide&Deep in recall. The NDUA uses the FM to explicitly extract the first-order and second-order features of the data. The NDUA algorithm automatically extracts low-level features to memorize the features.

It can be seen from Table 3(a) and Fig. 2(a) that the NDUA algorithm is superior to the comparison algorithm in precision, recall and F1. The NDUA algorithm is 1.73% higher than the Wide&Deep in precision and 1% higher than the Wide&Deep in recall. NDUA is 2.39% higher than the Wide&Deep in precision and 0.73% higher than the DeepFM in recall. NDUA has similar double-tower structure as Wide&Deep algorithm and DeepFM algorithm. The extraction of low-order features of Wide&Deep algorithm mainly depends on artificial feature engineering, which has strong subjectivity and requires strong prior knowledge. DeepFM algorithm uses FM to reduce order feature generation, but it does not consider the importance of the feature to the user's viewing behavior prediction. Compared to these methods, NDUA solves the problem of feature importance by introducing attention mechanism.

It can be seen from Table 3(b) and Fig. 2(b) that NDUA is superior to the comparison algorithm in precision, recall and F1. The NDUA is 1.73% higher than the Wide&Deep in precision and 1% higher than the Wide&Deep in recall. The NDUA uses the FM to explicitly extract the first-order and second-order features of the data. The Wide module of the Wide&Deep algorithm requires

Table 3. Precision and Recall of NDU A and comparison algorithm in predicting users' different behaviors

(a) viewing behavior			(b) collecting behavior		
Algorithm	Precision	Recall	Algorithm	Precision	Recall
FM	0.734	0.726	FM	0.766	0.758
Wide&Deep	0.752	0.741	Wide&Deep	0.774	0.768
DeepFM	0.61	0.753	DeepFM	0.781	0.771
NDUA	0.768	0.759	NDUA	0.787	0.781

(c) purchasing behavio		
Algorithm	Precision	Recall
FM	0.724	0.715
Wide&Deep	0.741	0.728
DeepFM	0.75	0.742
NDUA	0.758	0.753

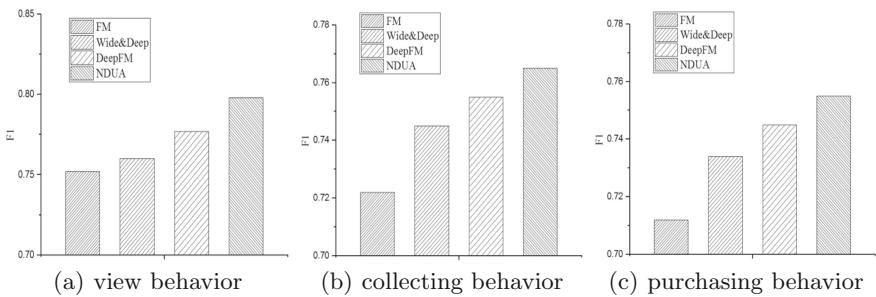


Fig. 2. F1 of NDU A and comparison algorithms in predicting users' different behaviors

artificial feature engineering to perform low-level feature crossing. The NDU A algorithm automatically extracts low-level features to memorize the features. And it implements end-to-end prediction of user behavior.

It can be seen from Table 3(c) and Fig. 2(c) that the NDU A algorithm is superior to the comparison algorithm. The NDU A algorithm is 0.6% higher than DeepFM in precision and 0.7% higher than DeepFM in recall. Both the NDU A algorithm and the DeepFM algorithm combine data low-order features and high-order cross-features to predict user purchasing behavior. The NDU A algorithm uses the attention mechanism to select the features in the FM, and uses the features that are positively related to the prediction of the user's purchasing behavior to conduct behavior prediction. Compared with the DeepFM algorithm, which only uses all the features of the FM with equal weight, the NDU A could more accurately predict the user's purchasing behavior.

5 Conclusions

We propose a financial big data feature mining method based on attention mechanism and dual channel (NDUA) to understand the deep features of financial big data and mine the association features of customers' financial behavior. First, for textual information in financial big data, the pre-trained model Bert is used to vectorize textual data in financial data. Next, a dual-tower model is proposed to predict user behavior, and multiple user behaviors are used as labels to perform model training. Then, we use the trained three sets of models to extract the last fully connected layer as the user behavior feature of the item, and the user behavior feature representations of all the operated items are clustered to further obtain the association of user behavior.

Acknowledgement. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant (No. 61902037, No. 61532006, No. 61772083, No. 61802028), and Science and Technology Major Project of Guangxi (GuikeAA18118054).

References

1. Kai, Y.L.: Design and implementation of weighted SVM forecasting model. *Electron. Prod.* **369**(04), 33–36 (2019)
2. Yao, Y., Xia, L., Xiao, P.L., et al.: Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *Int. J. Geograph. Inf. Sci.* **31**(4), 1–24 (2016)
3. Krisztian, B., Tomás, H.: Factorization machines for blog feedback prediction. *CORES* **79–85**, 2019 (2019)
4. Yu, C.J., Yong, Z., Wei S.C., Chih J.L.: Field-aware factorization machines for CTR prediction. *RecSys* **2016**, 43–50 (2016)
5. Wei, N.Z., Tian, M.D., Jun, W.: Deep Learning over Multi-field Categorical Data: A Case Study on User Response Prediction. *CoRR* abs/1601.02376 (2016)
6. Yan, L., Bin, G., Nuo, L., Jing, Z., et al.: DeepStore: an interaction-aware wide&deep model for store site recommendation with attentional spatial embeddings. *IEEE Internet Things J.* **6**(4), 7319–7333 (2019)
7. Hui, F.G., Rui, M.T., Yun, M.Y., Zhen, G.L., Xiu, Q.H.: DeepFM: a factorization-machine based neural network for CTR prediction. *IJCAI* **1725–1731**, 2017 (2017)
8. Jacob, D., Ming, W.C., Kenton, L., Kristina, T.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *CoRR* abs/1810.04805 (2018)



Current Control of Magneto-Rheological Clutch for Robotics Application

Jiangtao Ran, Guangzeng Chen, Chenguang Bai, and Yunjiang Lou^(✉)

Harbin Institute of Technology Shenzhen, Shenzhen 518000, China
louyj@hit.edu.cn

Abstract. The magneto-rheological clutch (MRC) has been used in brakes and clutches. In the field of robotics, the MRC is applied to dynamically change the torque through the current, which requires much higher precision than these traditional use. It is observed that using the simple PI controller results in significant overshoot and ripples in the current control of the MRC and thus degenerate the torque control performance of the MRC. This paper focuses on how to eliminate those to accurately control the MRC current. To solve the huge overshoot, a fuzzy PI control rule is established for the circuit model. It can dynamically change the control parameters and reduce the overshoot by 96%. For the noise in the circuit model, this paper proposes a combination of an analog filter and a digital filter, which reduces the ripple current of the circuit model by 80%.

Keywords: MRC control · Fuzzy PI control · Capacitor filter · Kalman filter

1 Introduction

The MRC is a kind of clutch with the material called magnetorheological fluid (MRF). It has been proved that it has many advantages in safety, compliance, high torque density, high-bandwidth, high-fidelity force control and impact-tolerant. By controlling the input current of the MRC, the phases of MRF inside the MRC can be changed from a free-flowing liquid state to a solid-like state within several milliseconds [1]. In traditional applications of MRC, the control circuit only needs to control the clutch current on and off, because it does not require the precision of control [2].

Dynamically and precisely controlling the output torque of the MRC is critical for robots in tracing the desired trajectory and safe physical interaction, etc. Note that the MRC is a single input single output system where the output torque can only be controlled by alternating the input current of the MRC.

J. Ran—This work was supported in part by the NSFC-Shenzhen Robotics Basic Research Center Program U1713202 and in part by the Shenzhen Science and Technology Program under Grant JCYJ20180508152226630.

Therefore, to precisely and dynamically control the input current of the MRC is the key to precise and dynamic torque control. As a common method in motor control, current control is widely used through the PI control algorithm [3], model predictive control (MPC) [4], optimized commutation angle [5]. These control algorithms have shown a good effect in motor, but can't perform well in MRC. The structure of both is significantly different in two points. One is that MRC doesn't have back electromotive force (EMF). The other is that, at the same power, the inductance and resistor of MRC can be ten times larger than that of the motor. It is observed that the PI controller will produce huge overshoot and ripple current in MRC current.

Facing the problems of huge overshoot and ripple current in the control, this paper mainly focuses on the current control of MRC. To reduce the overshoot, this paper establishes a fuzzy PI control rule that can dynamically adjust PI parameters. A combined filter proposed aims to reduce the ripple current by reducing the noise in the circuit model. The capacitor filter, as an analog filter, can swiftly eliminate the system noise generating by the circuit structure. Based on that, the Kalman filter, a kind of digital filter, can calculate the most accurate estimate value by observing the feedback current.

The paper is organized as follows. The circuit model is established in Sect. 2. The problem of PI control is shown in Sect. 3 and the overshoot is solved by a fuzzy PI control. The combined filter is proposed in Sect. 4 can effectively reduce the ripple current caused by system noise and measurement noise while increasing the system bandwidth. Finally, a conclusion is drawn in Sect. 5.

2 Structure and Circuit Model of MRC

2.1 MRC Structure

The structure of MRC is shown in Fig. 1. It consists of a rotor made of pure iron, aluminum stator, coils, a discs supporter made of aluminum, many stator discs, and rotor discs that are made of silicon iron and immersed in MRF [1].

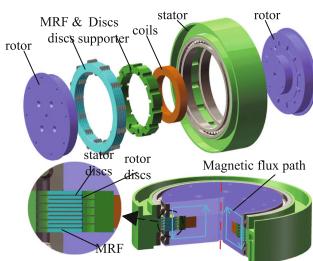


Fig. 1. MRC physical structure

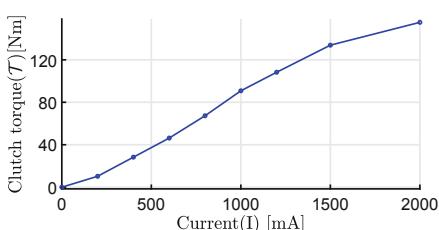


Fig. 2. Relationship between input current and output torque in MRC

When current is applied to the coil, magnetic fluxes are generated, passing along the path through MRF. Then MRF changes rapidly from liquid to semisolid state. Hence the stators and the rotors immersed in the MRF are engaged under magnetic forces. In general, the MRC shows current-controllable brake torque between the stator and rotor. According to Fig. 2, the input current and output torque of the coil can establish a linear functional relationship.

The PI controller will produce huge overshoot and ripple current in controlling that of MRC. For its resistance and inductance are much larger than the motor. The equivalent resistance in the MRC coil is $4.8\ \Omega$, equivalent inductance is 4.4 mH and maximum power is 20 W . Refer to the parameter of DC motor with part number 651606 produced by Maxon Motor. Its rated power is 50 W , internal resistance is $447\text{ m}\Omega$ and inductance is $243\text{ }\mu\text{H}$. The resistance of MRC is 10.7 times that of this DC motor, and the inductance is 18.1 times. Meanwhile, the MRC coil does not generate back EMF during operation.

2.2 Control Circuit Model

Control circuit is designed as Fig. 3(a). MRC can be modeled as an inductor and resistor marked in red in Fig. 3(a) and Fig. 3(b). To reach the high-precision current control, not only the MRC but also the current controlling circuit and current sensing circuit need to be taken into account, for they will bring great noise into the current sensing of MRC. The high-frequency MOS switch will generate a high-frequency system noise in the controlling circuit, which will affect current sensing through the feedback circuit. Current sensing is realized by the ADC chip, which also generates the measuring noise while working. These noises must be considered and filtered in the control.

Therefore, the simulation circuit is shown in Fig. 3(b) was established based on the actual control circuit of MRC. The necessary simulating parameters are shown as Table 1, where the Power Voltage is U_S , Power Capacitor is C , Equivalent Inductance is L , Equivalent Resistance is R_L , Current Sensing Resistance is R_{sen} , Forward Resistance is R_{DSon} , Forward Voltage Drop is U_{FV} , PWM

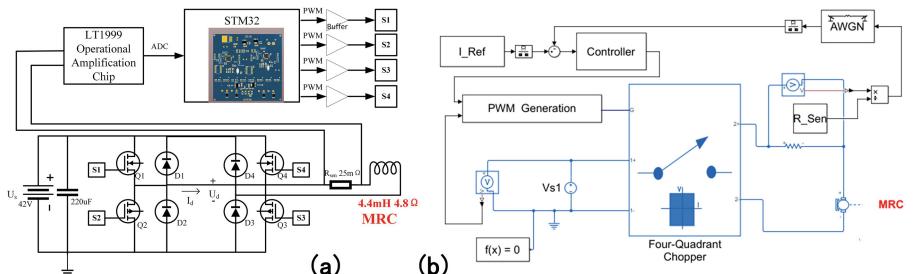


Fig. 3. (a) Control circuit and current sensing model design (b) Control circuit and current sensing model simulation structure

Table 1. MRC simulation circuit parameter

Symbol	Value	Unit	Symbol	Value	Unit
U_S	42	V	C	220	μF
L	4.4	mH	R_L	4.8	Ω
R_{sen}	25	$\text{m}\Omega$	R_{DSon}	56	$\text{m}\Omega$
U_{FV}	0.7	V	f_s	20	kHz
f_{sc}	10	kHz	f	100	Hz
m	2.2	kg	Φ	128	mm
T_{max}	156	Nm	P_{max}	19	W

Control Frequency is f_s , Current Sensing Frequency is f_{sc} , Reference Current Frequency is f , MRC Mass is m , MRC Diameter is Φ , MRC Max Torque is T_{max} , MRC Max Power is P_{max} .

The noise is also considered to better simulate the actual circuit working state that controls the MRC current. According to the definition of signal to noise ratio (SNR). In the Eq. (1), U_{source} is the amplitude of source voltage and U_{noise} is the amplitude of noise voltage. According to the datasheet of LT1999, the SNR of the feedback signal is 46 dB, which is also called measurement noise.

$$SNR = 20 \lg \frac{U_{source}}{U_{noise}} \quad (1)$$

3 Fuzzy PI Control

Due to PI control cannot change its parameters during the control, the best tuning should take the balance between the dynamic and static indicators of the control. However, when controlling objects with complex structures that are difficult to build mathematical models, a set of fixed-parameter PI controls cannot balance both indicators. According to the actual effect and engineering experience, this needs to generate a changing parameter to realize a better control effect. Therefore, a method that can dynamically change the PI parameters is needed to adjust the control effect.

Fuzzy PI control, shown in Fig. 4, can adjust the control parameters in real time according to the effect of the controller. When a step response occurs, it can increase the P-value based on original PI parameters and decrease the I-value to obtain a higher response speed and reduce the overshoot. When the signal tracking is stable, decrease the P and increase the I to improve stability, and reduce the steady-state error. The input and output domains, shown in Table 2, are divided into 7 equal parts. They are $NB, NM, NS, ZO, PS, PM, PB$. According to the above design ideas to establish the fuzzy PI rules as the Table 3.

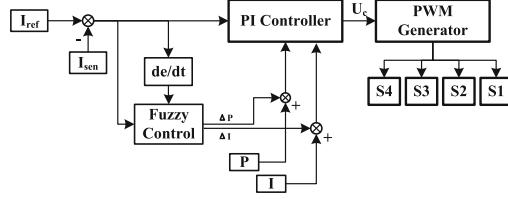


Fig. 4. Structure of fuzzy control

Through the simulation, the random step response of fuzzy PI control is shown as Fig. 5(a) and the frequency response is shown as Fig. 5(b).

Table 2. Fuzzy input and output ranges

Input	Range	Output	Range
e Error	$[-6, 6]$	ΔP	$[-50, 50]$
ec Error rate	$[-6, 6] \times 10^{-4}$	ΔI	$[-5, 5] \times 10^{-2}$

Table 3. ΔK_I and ΔK_P fuzzy logic rules

e	NB	NM	NS	ZO	PS	PM	PB	e	NB	NM	NS	ZO	PS	PM	PB
ec	NB	NB	NM	NM	NS	NS	ZO	NB	PB	PB	PB	PM	PS	PS	ZO
	NM	NB	NM	NM	NS	NS	ZO	ZO	NM	PB	PB	PM	PM	PS	ZO
	NS	NB	NM	NS	ZO	ZO	PS	PS	NS	PB	PM	PM	PS	PS	ZO
	ZO	NM	NM	NS	ZO	PS	PM	PM	ZO	PM	PM	PS	PS	PS	ZO
	PS	NM	NS	ZO	PS	PS	PM	PB	PS	PM	PS	PS	ZO	ZO	NS
	PM	ZO	ZO	PS	PS	PM	PM	PB	PM	PS	PS	ZO	ZO	NS	NM
	PB	ZO	ZO	PS	PM	PM	PM	PB	PB	PS	PS	ZO	NM	NS	NB

In Fig. 5(a), the input reference current is a random square wave signal with an amplitude from -3 to 3 and the frequency of 500 Hz. Figure 5(b) is a frequency response of PI control and fuzzy PI control. Calculating curves in Fig. 5(a) and Fig. 5(b), the overshoot of PI control is $\sigma = 16\% \text{--} 60\%$, the system bandwidth is 770 Hz and the ripple current is $I_{rms} = 200$ mA. This result is the best parameter tuning in PI control. According to the fuzzy PI control, its overshoot is reduced to $\sigma = 12\%$ and its system bandwidth is increased to 810 Hz. Although fuzzy PI control can dynamically change the PI parameters, its drawback is that the accuracy will be reduced to some extent, which is shown as the ripple current

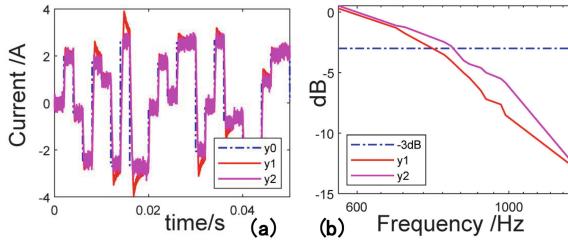


Fig. 5. (a) PI control and fuzzy PI random step response (b) PI control and fuzzy PI frequency response y_0 is input reference, y_1 is PI control, y_2 is fuzzy PI

slightly increased to $I_{rms} = 260$ mA. Therefore, how to reduce the ripple current will be study in follow.

4 MRC Ripple Current Filter

The ripple current in the fuzzy PI mentioned above mainly comes from the system noise and measurement noise in the circuit. The filter is a common method that can effectively suppress noise, which includes analog filter and digital filter. This section uses a combination of an analog filter and a digital filter to eliminate ripple current. Capacitor filter can quickly eliminate a lot of system noise. It does not cause delay to the measurement circuit and prepares for the digital filter. As a digital filter, the Kalman filter can estimate the most accurate current.

4.1 Capacitor Filter in Circuit

In the ideal condition, the four groups MOS and freewheeling diode of the H-bridge circuit in Fig. 3(a) can be regarded as a high-frequency switch in the model. It is considered that controlling the duty cycle of the switch DU_S in Fig. 6(a) can achieve the equivalent effect of the H-bridge circuit. The circuit shown in Fig. 6(a) establishes a transfer function with DU_S as the input and U_O as the output, written with Eq. (2). According to the Eq. (2), The Bode diagram is shown in Fig. 7. If the frequency of reference input 200 Hz, according to the blue curve in Fig. 7, the SNR of system noise is $-45\text{ dB}/-72\text{ dB} = 27\text{ dB}$. The higher SNR has a lower system noise. The measurement noise mentioned above by Eq. (1) is 46 dB.

Adding the capacitor filter to the Fig. 6(a), where C is $10\mu\text{F}$, $R_C = 5\Omega$ is the equivalent series resistance of C at a switching frequency of 20 kHz. The circuit is transformed into Fig. 6(b). The transfer function of the Fig. 6(b) is shown in Eq. (3) and $Z(s)$ is $R_{sen}/(R_C + \frac{1}{Cs})$. From the first order to the second, the transfer function can increase the low-frequency gain to reduce the system noise. According to the above calculation, the SNR of system noise in

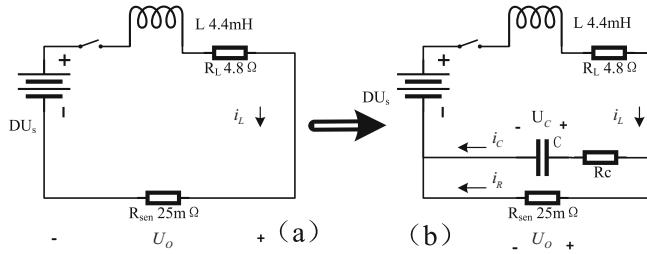


Fig. 6. (a) Primary circuit equivalent conversion (b) Adding a capacitor filter to the circuit

the Fig. 6(b) is $-30 \text{ dB} / -72 \text{ dB} = 42 \text{ dB}$. The improvement decreases the system noise from 27 dB to 42 dB, a reduction of 5.6 times.

$$\frac{U_O(s)}{U_{in}(s)} = \frac{R_{sen}}{Ls + R_L + R_{sen}} = \frac{1}{0.192s + 177} \quad (2)$$

$$\frac{U_O(s)}{U_{in}(s)} = \frac{Z(s)}{Ls + R_L + Z(s)} = \frac{62500000s + 3125}{12000000s^2 + 62500603s + 3125} \quad (3)$$

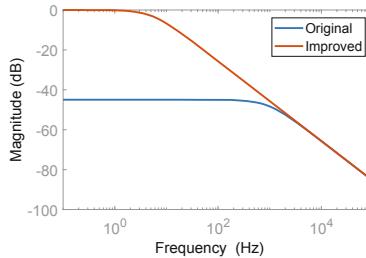


Fig. 7. Bode diagram of primary circuit and improved circuit with a capacitor filter

In Fig. 6(a), the maximum current is 2 A, and the peak current can reach 5 A. 5 A current can make the voltage generated by the R_{sen} to $U_O = I_{max}R_{sen} = 75 \text{ mV}$.

In Fig. 6(b), the voltage across capacitor $U_C \leq U_O$. Considering the worst case is $U_C = U_O$ and the current changes from +5 A to -5 A. In this situation, $i_C = C \frac{dU_C}{dt} = 1.5 \mu\text{A}$. Compared with 5 A, i_C can be ignored. In summary, the circuit in Fig. 6(a) and the circuit in Fig. 6(b) can be considered equivalent.

4.2 Establish Kalman Filter Equation

The Kalman filter equation can be established for the circuit in Fig. 6(b). The input value is set to the voltage that can control the duty cycle $DU_s(t)$.

The output is defined as the current flowing through the inductor $i_L(t) \approx i_R(t)$, the state variable $\begin{bmatrix} i_L(t) \\ U_O(t) \end{bmatrix}$

$$\begin{bmatrix} \frac{i_L(t)}{dt} \\ \frac{U_O(t)}{dt} \end{bmatrix} = \begin{bmatrix} \frac{-R_L}{L} & \frac{-1}{L} \\ \frac{1}{C} - \frac{R_C R_L}{L} & \frac{-(\frac{R_C}{L} + \frac{1}{CR_{sen}})}{1 + \frac{R_C}{R_{sen}}} \end{bmatrix} \begin{bmatrix} i_L(t) \\ U_O(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{R_C} \\ \frac{1}{1 + \frac{R_C}{R_{sen}}} \end{bmatrix} DUS(t) \quad (4)$$

$$i_L t = [1 \ 0] \begin{bmatrix} i_L(t) \\ U_O(t) \end{bmatrix} \quad (5)$$

Adding noise terms to Eq. (4) (5) can be expressed as the standard form shown as the Eq. (6).

$$\begin{aligned} X_k &= AX_{k-1} + BU_k + \omega_k \\ Z_k &= HX_k + \nu_k \end{aligned} \quad (6)$$

Kalman equation

$$\begin{aligned} \hat{x}'_k &= A\hat{x}_{k-1} + BU_k \\ P_k' &= AP_{k-1}A^T + Q \\ \hat{Z}_k &= Z_k - H\hat{x}'_k \\ K_k &= P_k'H^T (HP_k'H^T + R)^{-1} \\ \hat{x}_k &= \hat{x}'_k + K_k\hat{Z}_k \\ P_k &= (I - K_kH)P_k' \end{aligned} \quad (7)$$

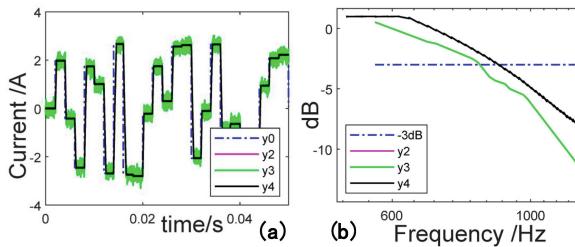


Fig. 8. (a) Step response (b) Frequency response y_0 is reference input, y_2 is fuzzy PI, y_3 is fuzzy PI with Kalman filter, y_4 is fuzzy PI with capacitor filter and Kalman filter

where x_k is true value, \hat{x}_k is Kalman estimated value, P_k is Kalman estimated error covariance matrix, \hat{x}'_k is prime predicted value, P_k' is prime predicted error covariance matrix, K_k is Kalman gain, \hat{Z}_k is measurement margin. ω_k is subject to normally distributed noise with mean value 0, covariance matrix Q . $\omega_k \sim N(0, Q)$ ν_k is subject to normally distributed noise with mean value 0, covariance matrix R . $\nu_k \sim N(0, R)$. Equation (7) are the Kalman filter derived from Eq. (6).

Table 4. Performance indexes of different control algorithms

	PI control	Fuzzy PI	Kalman
Overshoot	16%–60%	12%	1%–2%
Cripple current	± 200 mA	± 260 mA	± 40 mA
System bandwidth	770 Hz	810 Hz	825 Hz

Bring Eq. (7) into the simulation circuit. The simulation results are shown in Fig. 8(a) and Fig. 8(b). If only the capacitor filter were applied, the measurement noise in the circuit couldn't be eliminated. SNR is an indicator of noise suppression, and the smaller noise has bigger the SNR. If only the fuzzy PI control were applied with the Kalman filter, because of the measurement noise 46 dB much smaller than the system noise 27 dB, the estimated value of the Kalman filter would more biased to the observed value. This would lead to the Kalman filter has no effect as shown as y_3 in Fig. 8. y_4 is the result of a Kalman filter with a fuzzy PI and capacitor filter, which decrease the ripple current significantly. At the same time, its system bandwidth is higher than before. By calculating, these system performance indicators are shown as Table 4.

5 Conclusion

The work of this paper mainly focuses on the current control of MRC. In the process of converting the physical circuit into the Simulink model, the parameters of the physical circuit are fully considered. However, due to the electrical characteristics of MRC, large inductance, the traditional PI control with the best tuning parameters which will lead to a huge overshoot. This paper uses the fuzzy control algorithm to reduce the overshoot in the circuit. Besides this problem, it is also necessary to solve the problem of excessive ripple current to further improve the control accuracy. To solve that, the paper combines the capacitor filter and Kalman filter to better observe feedback current on the basis of reducing system noise. Finally, it can be concluded that the Kalman filter with fuzzy PI control and capacitor filter will improve the effectiveness of traditional PI control drastically. This reduces the original overshoot by 96%, the ripple current by 80%, and increases the system bandwidth to 825Hz.

References

- Chen, G., Lou, Y., Shang, T.: Mathematic modeling and optimal design of a magneto-rheological clutch for the compliant actuator in physical robot interactions. *IEEE Robot. Autom. Lett.* **4**(4), 3625–3632 (2019)
- Kasemi, B., Muthalif, A.G.A., Rashid, M.M., Rahman, M.: Optimizing dynamic range of magnetorheological fluid dampers: modeling and simulation. In: 2011 4th International Conference on Mechatronics (ICOM), pp. 1–4, May 2011

3. Hua, C., Wu, C., Chuang, C.: A digital predictive current control with improved sampled inductor current for cascaded inverters. *IEEE Trans. Ind. Electron.* **56**(5), 1718–1726 (2009)
4. Cortes, P., Rodriguez, J., Quevedo, D.E., Silva, C.: Predictive current control strategy with imposed load current spectrum. *IEEE Trans. Power Electron.* **23**(2), 612–618 (2008)
5. Sun, J., Chen, X., Wang, Z., He, J.: Current spike and efficiency optimization by using dynamic model of open-loop voltage mode single-phase BLDC cooling fan motor. In: 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), pp. 1209–1214, December 2017



Research on Open-pit Mine Virtual Environment Construction and Working Vehicle Modeling Simulation Based on PreScan

Yixin Guo¹, Guizhen Yu¹, Bin Zhou^{2(✉)}, Yucan Wang³, and Guoqiang Liu¹

¹ School of Transportation Science and Engineering, Beijing Advanced Innovation Center for Big Data and Brain Computing, Beijing Key Laboratory for Cooperative Vehicle Infrastructure Systems and Safety Control, Beihang University (BUAA), Beijing 100191, China

² School of Electronic Information Engineering, Beihang University (BUAA), Beijing 100191, China
zhoubinhappy00@126.com

³ School of Computer and Communication Engineering, University of Science and Technology Beijing, No. 30 Xueyuan Road, Haidian District, Beijing 100083, China

Abstract. Unmanned driving is the core and difficulty in realizing the intelligent and unmanned mining. For the urban block environment, there are already many available simulation systems that can meet the needs of authenticity and accuracy of urban block buildings simulation. But for the mining area environment, there is no simulation software that can provide matching scenario construction and vehicle modeling. The purpose of this article is to verify the unmanned driving algorithm of mining vehicle in the simulation system before it is used on the actual mining truck. Mining vehicle models and simulated driving environment of the open-pit mine area is built in this article. A more realistic solution for setting up a mining scenario and mining vehicle model simulation system is also provided. Through the content of this paper, the construction of unmanned driving simulation environment and the modeling and simulation of operating vehicles in open-pit mining area are realized.

Keywords: Open-pit mine · Unmanned driving · Open-pit mining environment simulation · Mining vehicle simulation

1 Introduction

China is rich in mineral resources with 159 proven minerals and more than 100,000 mining mountains. Mining and transportation of mines have attracted much attention due to environmental, safety, and efficiency issues. To increase productivity, reduce costs, and cope with labor shortages, intelligent and

unmanned technologies are urgently necessary to enhance industrial capabilities [1]. With the construction of smart mining, unmanned driving has become a new way of transportation in mining areas. The obvious sign of smart mining is unmanned, that means unmanned operation in mining, unmanned operation in digging, unmanned operation in the dangerous place, unmanned operation of the large equipment, and unmanned operation of the entire mine [2]. The operation of all aspects of the entire mine is expected to use the intelligent robot and intelligent equipment.

Autonomous driving simulation test platform mainly includes virtual scenario construction, vehicle dynamics simulation, sensor simulation [3]. Autonomous driving virtual scenario needs to be as similar as possible to actual scenes. It uses mathematical modeling to restore and generalize the real world digitally so that the simulation platform can be authentic and reliable, and the test results could be highly credible. With the help of computer simulation technology, the virtual simulation test has the advantages of faster test speed, lower cost and no safety risk compared with the actual vehicle test [4]. Due to the different traffic scenes and some variable factors, how to construct a traffic participant model that simulates real traffic scenes has become a key issue of the automatic driving test [5]. There are some relatively mature simulation systems which can meet the researchers claim about authenticity and accuracy in building urban block virtual scenes [6].

The research of this paper considers how to build the platform to verify the test quality of the real vehicle automatic driving algorithm as soon as possible when the simulated environment is as close as possible to the real mining area environment. The algorithms which is used on real vehicles can be tested on virtual vehicle models so that the problems can be revealed at the simulation level, and then they can be corrected before real vehicle testing. Besides, they can also be used to test the algorithm focus on solving problems such as the weather that are impossible or difficult to obtain, and environmental conditions that hard to meet in reality. It helps achieve comprehensive test verification of the test algorithm and decrease the problem of algorithm misfits caused by special weather conditions.

2 Related Work

In this paper, a simulation platform system for open-pit mining based on PreScan, SketchUp 3D and Matlab/simulink is researched. Through modeling and mapping of special scenarios in open-pit mining area, 3D modeling and dynamic construction of unmanned mining trucks, 3D model building of electric excavator and the matching movement of loading and unloading, finally realize the open-pit mining environment construction and vehicle modeling and simulation, and provide simulation conditions for the test and verification of unmanned mining trucks algorithm.

Simcenter PreScan is an advanced simulation software developed in the field of vehicle simulation, which is relatively perfect in function, vivid in visual display, and can be used in conjunction with a variety of software to realize auxiliary driving. Because of its simulation capability and various interfaces, it can be

combined with a variety of drawing formats and software. It provides users freedom with scenario-setting and allows users to customize their demands in vehicle models to achieve the feasibility and applicability of vehicle control algorithms in various road environments and weather conditions.

In PreScan, combined with SketchUp, arbitrary material can be used arbitrarily. Not only static material can be added, but also dynamic material can be added. In terms of vehicle dynamics, with the help of Model Preparation Tool, we can configure key parameters for the dynamic vehicle model that has set the motion joint points to achieve the performance effect, we can quickly build vehicle dynamics models. In sensor modeling, PreScan provides a variety of sensor models with adjustable parameters. It can set up the sensor model according to the actual needs, and output the data through Matlab/Simulink. The system structure process of building a mining simulation platform is shown in Fig. 1.

The main contents of this paper are divided into two parts: in the first part we construct the basic model of the open-pit mining area which is based on the environmental data and make detailed corrections and improve the reality of the environmental simulation. In the second part, the 3D model of mining truck is drawn up according to the data of the real mining vehicle type and model size, and the node components are set up for vehicle tires, dump box of trucks, electric shovel body and shovel arm to realize loading and unloading movement. The dynamic model and functional test of mining trucks are realized and make the control algorithm for vehicle parking and speed control testing come true.

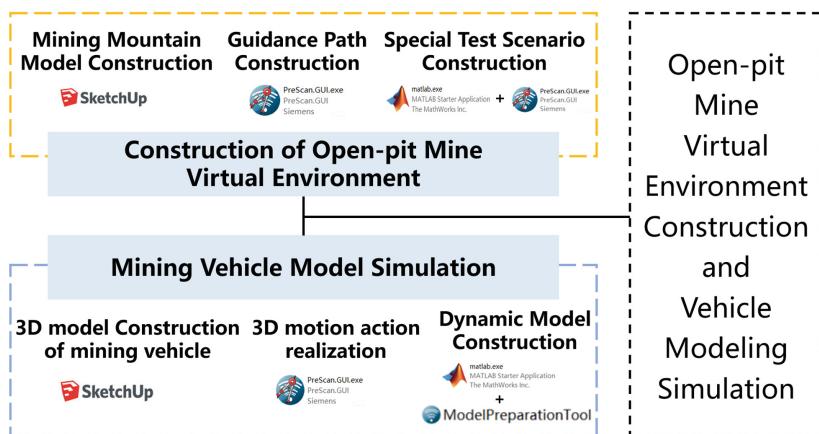


Fig. 1. System structure

3 Mining Environment Construction

PreScan and Panosim are relatively mature in the construction of urban blocks. The road models have been integrated into the software such as straight roads,

turns, ramps, and crossroads. Researchers can directly use the building environment and many other environmental building materials. But for the open-pit mining environment, there is almost no simulation environment to use. In this paper, we use SketchUp to build the mining area environment and draw the guidance path and build variable speed vehicle tracking scenario to achieve the simulation environment of the open-pit mine.

3.1 Mountain Environment Construction in the Mining Area

The mining area is mostly a mountain or a huge mine. It is located in a remote location and the climate is bad. The shape and trend of each pit or mountain are different. There are no specific locations for the vehicle running area and some stopping points for the special mining vehicles. According to the characteristics of the mining area, the drawing of the mining area map requires the collection of geographic information and mapping (Fig. 2).

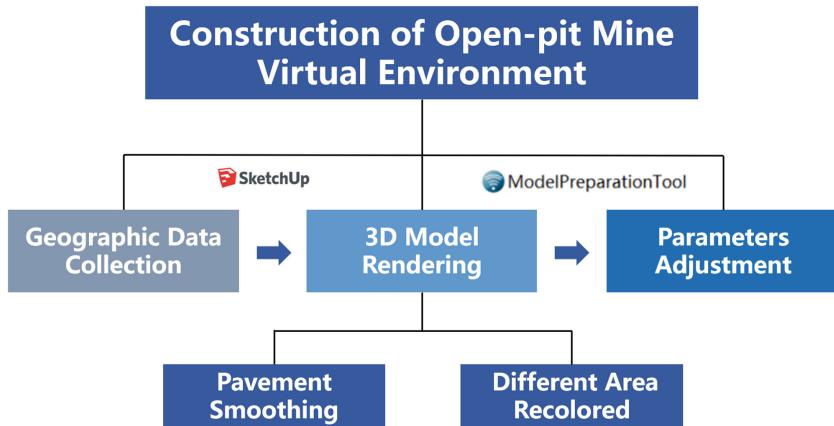


Fig. 2. Virtual mining environment construct structure

Due to the generalization of data in the rendering process some join positions are not smooth enough. Because of the high slope the vehicle is easy to cause rollover phenomenon when it reaches the location. To solve this problem, we need to continuously smooth and optimize model building. Besides, to improve the visual effect, it is necessary to distinguish the location of vehicles and the location of vehicle parking points with the color of the mountain to make the model more lifelike. The gray part is the vehicle driving area and the vehicle working area at the bottom of the mine pit. The Yellow part is the mountain body. The virtual environment can be seen in Fig. 3. The first white graph is the original drawing, the second graph is the colored map, and the third chapter is the real mining area photograph.

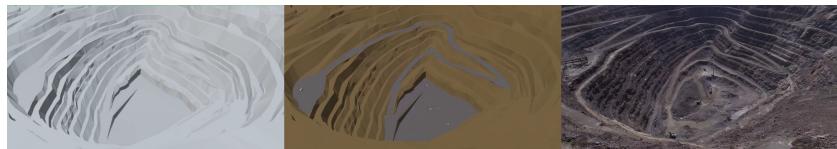


Fig. 3. Mining area simulation diagram

3.2 Vehicle Guidance Path and Special Driving Scene Construction Based on the Mountain Environment

After the completion of construct of the mining area, all the moving objects in PreScan should be set up with the initial starting point to make the tire and pavement match tightly. Because of only the matching path can be added to the own road model in PreScan, the path matching of the imported elements cannot be carried out directly while the mountain environment model is an import element. Therefore, to ensure that vehicles not to rollover when they are started at the beginning, it is necessary to use PreScan's own road to create path to guide vehicles to achieve the strict fit between the tire and the road when the vehicle starts.

Vehicle guidance path construction. Due to the mountain environment model is an irregular shape model and the height of each point in the vehicle area is not strictly measurable. It is not appropriate to set up the vehicle starting height only according to the elevation information of the mountain environmental model data. We use the PreScan Ramp road to set up the starting road which height is near to the mountain model road. The vehicle can be tightly fitted with the help of its own road track. Secondly, the slope of the road needs to be reasonably set up and vehicle can be driven into the mountain environment model road with a certain smooth slope. Because the mountain environment model is detectable for vehicles the vehicle will continue to run on the road surface of the mountain model road when the vehicle tires are guided by the smooth path to the mountain model (Fig. 4).



Fig. 4. Mining truck

In order to facilitate the algorithm-testing, we add 7 vehicle starting roads in the mining-environment model (Fig. 5). With the help of the guidance path, the vehicle can start smoothly and drive on the mining environment model according to the control trajectory under the control of the algorithm.

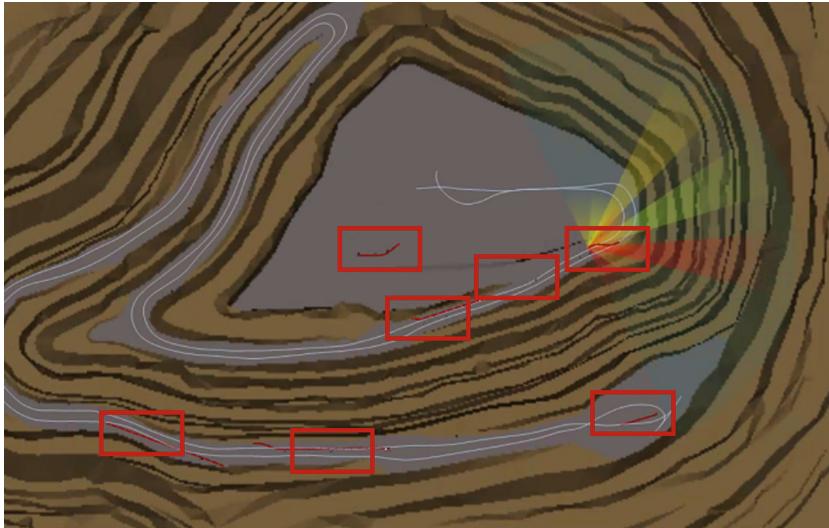


Fig. 5. Starting roads diagram

Special test scenario construction. According to the test demand, we can also add the guidance path to guide the obstacle vehicles and carry out different vehicle speed plans to create different test scenarios such as car-following. In this paper we build the speed-changing vehicle as an obstacle car and achieve the car-following test through Matlab/Simulink (Fig. 6) in the car-following scenario (Fig. 6).

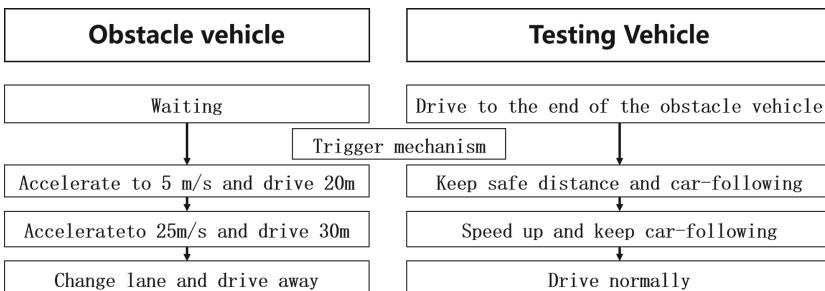


Fig. 6. Construction process of car-following

In the testing process, the first test scheme was to wait for the scheduled time to realize the scenario that the obstacle vehicle just started when the testing vehicle approached. The waiting time is calculated by the range from the start point to the obstacle vehicle and speed of the test vehicle. Number 1 scheme cannot meet the multi-start car-following test, so add trigger mechanism to realize that the obstacle vehicle can start to move and running under the variable speed when any vehicle reaches the 5 m behind the obstacle vehicle. The process is realized through Simulink (Fig. 7). The rear vehicle is the test vehicle controlled by the control algorithm and the front vehicle is the obstacle vehicle with different driving speeds.

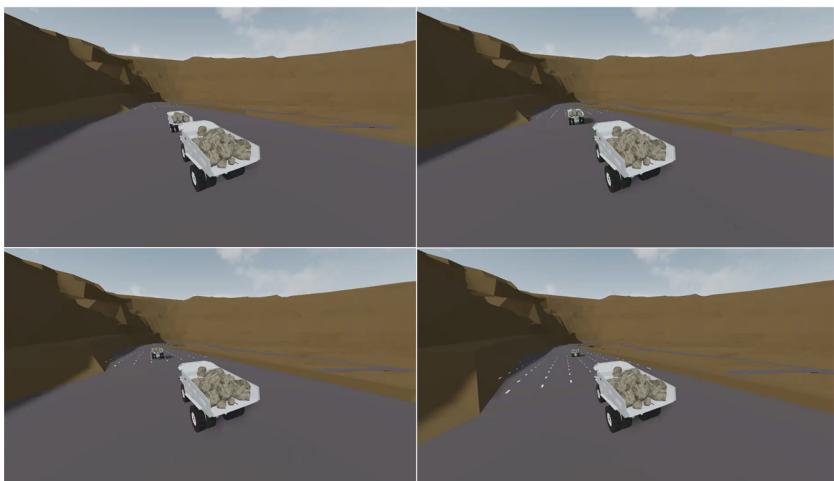


Fig. 7. Car-following scenario

4 Vehicle Model Simulation

4.1 Construction of Three-Dimensional Model and Functional Dynamic Model of Mining Truck

Mine trucks have their special characteristics as special vehicles. For example, the size of the mining trucks is 2 times larger than ordinary trucks and a single tire is 3 m higher than normal trucks. The mining trucks types and brands are different in different mining areas. According to the data of different type and size of mine trucks the 3D model of mine trucks can be built.

In order to achieve the lifting and lowering of the cargo bucket with the action of dumping the mineral material, it is necessary not only to set up the rotary joint for the cargo bucket, but also need to add the joint to realize the translation and rotation of mineral material. So that visual authenticity of the ore unloading can be realized (Fig. 8).



Fig. 8. Unloading simulation

The vehicle dynamics model is generally built by the vehicle dynamical simulation software such as Matlab/Simulink that can simulate the vehicle control and real-time simulation. The simulation terminal receives the control signals such as throttle and brake that is given by the control module of the autopilot system and generates and improves the vehicle pose and chassis parameters at the same time. The simulation terminal also directly controls each module of the vehicle models. The automatic driving control module works out the vehicle position and speed through the vehicle dynamics simulation data and sends out reasonable control signals.

In this paper the method of dynamic construction is based on PreScan and Matlab/Simulink. The purpose is to use the mine truck model to carry out the functional verification of the control algorithm. Besides, it is so difficult to obtain the kinetic data of the mining truck that the precise and accurate dynamic model of mine truck is temporarily difficult to simulate. Therefore, we proposed a fast method to build a functional mining truck dynamics model.

According to the three-dimensional model of the mining truck which was already built and the Model Preparation Tool software we can build the 3D or 2D dynamics model. The dynamic model of the mining truck can be built quickly in Model Preparation Tool and the dynamic model in this paper is built by using the data of van (Fig. 9).

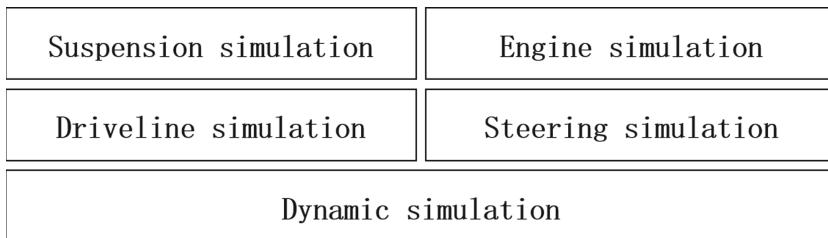


Fig. 9. Construction process of mining truck dynamics model

4.2 Three-Dimensional Model Construction of Mining Electric Shovel

The mining electric shovel has its own special characteristics too as a special vehicle. For example, it has track instead of tires and the movable shovel arm and the 360-degree horizontal rotating cab are used. The models and brands of the excavators are different in different mining areas and the three-dimensional model of the mining electric shovel can be built according to the required size.

In order to achieve the combined action of the shovel arm bucket, it is necessary to build a linkage timing scheme for the three joints of the shovel arm body connection point and the middle point of the shovel arm and the shaft of the bucket. According to the loading operation with the mining truck, it is necessary to build the timing scheme through Simulink. Finally, the loading operation of lowering arm, digging material, lifting arm, body rotation and lifting arm - dumping body rotation can be realized.

In this paper, the ore material is drawn as a large ore to simulate the loading scenario to improve authenticity and make it more realistic. The joints of each ore are set up so that each ore can be set as a single individual for speed setting and movement timing. The speed programming of each block is done through the free-falling motion equation and the action time is allocated according to the loading plan (Fig. 10).

5 Analysis of Experimental Results

In the experiment, a mining truck model is added to the mining mountain environment and the mining truck model is controlled by the controller. Compared and verified the simulation data of virtual vehicle and the GPS data obtained by the real vehicle, results are shown in Fig. 11. The comparison between the vehicle trajectory in the simulation environment and the real vehicle GPS trajectory also can be seen in Fig. 12. The following Table 1 gives a summary of all the average difference of each data.

From the curves and tables, we can see that the data of trajectory of the vehicle model simulated in the virtual environment of the open-pit mine is

basically the same with the real vehicle's running trajectory in the real mining area. The results verified the function of the simulation system and the feasibility of the test for the real vehicle control algorithm.

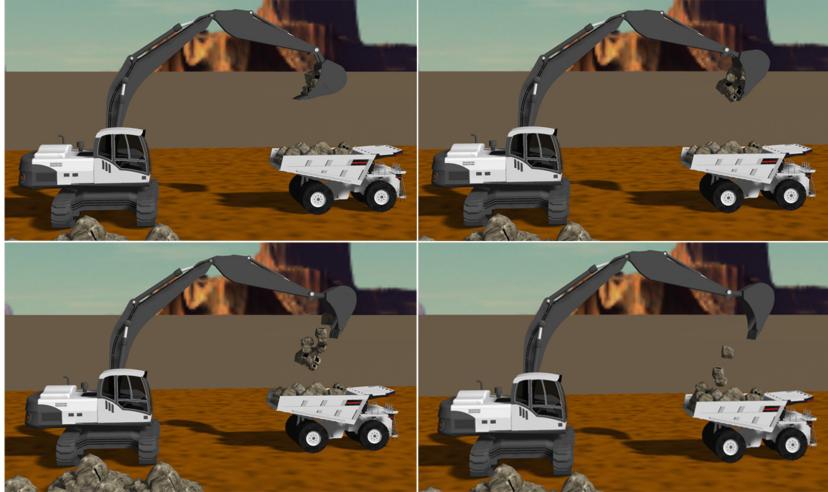


Fig. 10. Loading simulation

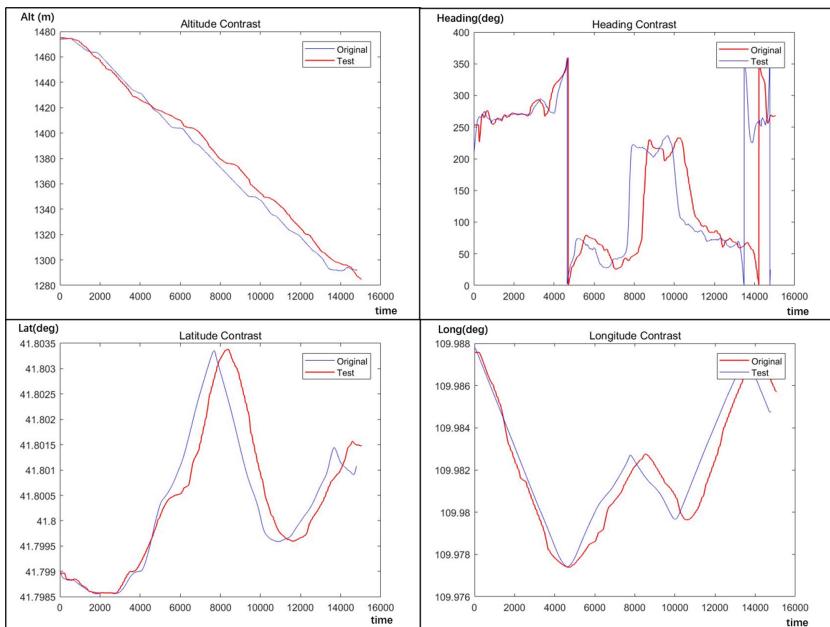


Fig. 11. Different data contrast

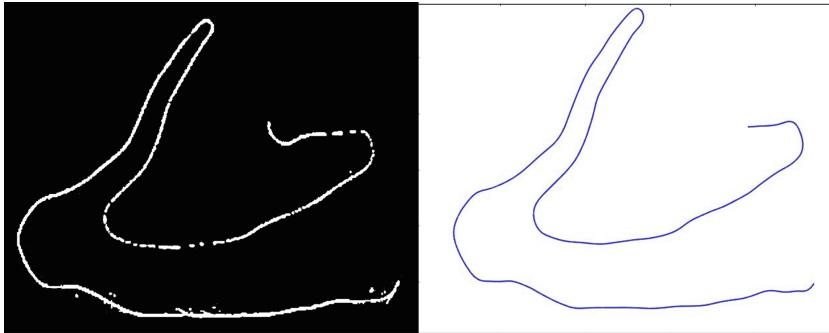


Fig. 12. Vehicle trajectory contrast

Table 1. Average difference between two methods

Data type	Average
Altitude (deg)	4.826613
Latitude (deg)	0.000036
Longitude (deg)	-0.00034
Heading (deg)	-7.75766

6 Conclusions

In this paper, a simulation platform for open-pit mining based on PreScan and Sketchup and Matlab/Simulink is researched. Specifically, the 3D simulation environment of the mining area and the 3D shell of the vehicle models and mining trucks and mining electric shovels are established by the sports joint addition. Besides, a rapid dynamic construction method for mining trucks are presented. According to the method of constructing the simulation platform of the open-pit mine, the algorithm of the unmanned vehicle control in the open-pit area can be tested. The simulation of the special scenarios that car-following is carried out to achieve the multi-scene test of the driverless algorithm. The security of algorithm can be improved by test in the virtual mining simulation as far as possible. The methods of building special scenario simulation platform which is proposed in this paper is of great practical significance for constructing other special scenarios simulation platform.

Acknowledgments. This work is partially supported by the Beijing Natural Science Foundation under Grant L191002. The authors would also like to thank the insightful and constructive comments from anonymous reviewers.

References

1. Donglin, L., Xiangyang, L., Lei, L., et al.: An overview of autonomous open-pit mine Haulage system. *Electric Drive Locomotives* (02), 1–8 (2019). <https://doi.org/10.13890/j.issn.1000-128x.2019.02.001>
2. Elgharbawy, M., Scherhauser, I., Oberholzer, K., et al.: Adaptive functional testing for autonomous trucks. *Int. J. Transp. Sci. Technol.* **8**(02), 202–218 (2019)
3. Yu, C., Shitao, C., Tangyike, Z., et al.: Autonomous vehicle testing and validation platform: integrated simulation system with hardware in the loop. In: 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, Suzhou, China, 26–30 June 2018 (2018)
4. Kalra, N., Paddock, S.M.: Driving to safety: how many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transp. Res. Part A Policy Pract.* **94**, 182–193 (2016)
5. Wei, Z., Xinhui, L., Xueyi, W., et al.: Research status of autonomous vehicle simulation technology. *Auto. Electric Parts* (08), 13–15 (2019). <https://doi.org/10.13273/j.cnki.qedq.2019.08.007>
6. Zhu, B., Zhang, P., Zhao, J., et al.: Review of scenario-based virtual validation methods for automated vehicles. *China J. Highway Transport* **32**(6) (2019)



Research on Hardware-in-the-Loop Test System for the Main Control Unit of Unmanned Vehicles

Haoyuan Ni¹, Guizhen Yu^{1(✉)}, Bin Zhou², and Guoqiang Liu¹

¹ School of Transportation Science and Engineering, Beijing Advanced Innovation Center for Big Data and Brain Computing, Beijing Key Laboratory for Cooperative Vehicle Infrastructure Systems and Safety Control, Beihang University, Beijing 100191, China

[✉] yugz@buaa.edu.cn

² School of Electronic Information Engineering, Beihang University, Beijing 100191, China

Abstract. At present, there are more researches on the automatic driving technology of unmanned vehicles. It needs a main control unit, which enables the vehicle to automatically drive. The main control unit supports the access of various sensors and controllers. It can process data of each device and control the movement state of vehicles. Therefore, the main control unit needs to be tested repeatedly to ensure its functionality and stability. Based on Prescan simulation software and Matlab/Simulink software, this paper studies a hardware-in-the-loop (HIL) test system for testing the main control unit of unmanned vehicles according to the communication protocol and communication method of the real sensor and controller hardware devices. At the same time, this paper studies a simulation method of the main control unit, which can effectively verify the function of the test system and conduct the feasibility test of the communication protocol and communication method used.

Keywords: Automatic driving technology · Unmanned vehicles · Main control unit · Prescan · Matlab/Simulink · Hardware-in-the-loop (HIL) test system

1 Introduction

With the rapid development of automatic driving technology, more and more unmanned vehicles begin to use the automatic driving system. The so-called main control unit refers to the system that integrates the functions of main automatic driving technology, such as decision planning and execution control. Generally, it is an Industrial Personal Computer (IPC), which is the “brain” in the automatic driving system. The device can obtain the data information of the connected sensors and controllers, and combine the automatic driving

control algorithm programs developed under various Integrated Development Environment (IDE) platforms in the device, and perform data fusion, analysis and calculation, and output vehicle control signals to the behavior control actuator. In terms of behavior control actuator, the drive-by-wire control system is the ultimate development trend, which is to use electronic units to replace mechanical units. However, from small vehicles to heavy vehicles, using actual vehicles to test the main control unit is more expensive. And there is much software that can simulate unmanned vehicles, such as Udacity, CarSim, Prescan, etc. However, the interface effect of Udacity is relatively cartoon and functions are relatively simple. CarSim is mainly the vehicle simulation, to improve the speed of simulation, the models are relatively simple. Simcenter Prescan is a physically-based simulation platform based on Global Positioning System (GPS) and other sensor technologies for the development of the Advanced Driving Assistance System (ADAS). All models added in the “Prescan GUI” can be automatically generated in Matlab/Simulink for collaborative simulation. The advantages of selecting Prescan are summarized as follows: (1) The software is powerful and the sensor system is rich. (2) Collaborative simulation with Matlab/Simulink is very suitable for data processing and interaction with external devices. (3) Through Prescan’s 3D VisViewer visual interface and Simulink, the driving situation and real-time data information during the simulation process can be observed and analyzed.

At present, there are many simulation studies on intelligent driving technology by using Prescan and Matlab/Simulink. In the domestic aspect, Gu et al. [1] analyzed the braking efficiency of the Autonomous Emergency Braking(AEB) system under different working conditions through the co-simulation of PreScan, Carsim and Matlab. Wu et al. [2] verified the algorithm of collision avoidance strategy at no signal intersections by simulating the scene in Prescan. Zhang et al. [3] proposed a performance test method for lane departure warning system based on Prescan, the LDWS camera was positioned in front of the computer which was displaying the virtual environment to capture images including lane markings and other information. Ruhai et al. [4] proposed realistic lane-change warning rules. Through Matlab/Simulink and Prescan co-simulation, the lane-change warning rules were verified.

In foreign countries, Hendriks et al. [5] used Prescan and VeHIL to study the active collision avoidance system. Prescan was a software-in-the-loop (SIL) environment, and VeHIL was its corresponding hardware-in-the-loop (HIL). Molemaar et al. [6] described the need for physics-based simulation and introduced a new spectral HDR camera model in the ADAS simulation platform Prescan in order to support the development process of ADAS by means of simulation. Nacu et al. [7] used virtual reality technology and the Prescan-Simulink simulation environment to realize the research of automatic electric vehicles. Zaki et al. [8] conducted simulation tests on Prescan to verify the vehicle traffic sign recognition algorithm model. Kim et al. [9] used sensing sensors in Prescan and to conduct hardware-in-the-loop tests on connected self-driving vehicles in real-time traffic.

In the current research, the use of Prescan software for the simulation of the integrated navigation system and drive-by-wire control system in the automatic driving system, and the HIL test systems for the main control unit is still few. Without combining the hardware device of the actual vehicle, or testing only some functions that do not constitute a closed-loop, the function of the simulation system cannot be well-reflected.

2 Research Contents

In view of the above contents, this paper studies a HIL test system for the main control unit of unmanned vehicles based on Prescan and Matlab/Simulink. It can test the functionality and stability of vehicle tracking control of the main control unit and the communication protocol and communication method of the devices used in reality. The problem of the main control unit will be exposed on the simulation level, and it can save a lot of manpower and material resources and time consumption.

The system structure is shown in Fig. 1. The simulation system is a computer host equipped with Prescan and Matlab/Simulink, which contains the simulation of the sensor and the controller. The simulation sensor is the integrated navigation system and the simulation controller is the drive-by-wire control system. Through serial communication and Can-Bus communication, the communication between the integrated navigation system and the drive-by-wire control system and the main control unit can be realized. Specifically, the simulation of the drive-by-wire control system requires the vehicle dynamics model, and Can-Bus communication is implemented by Simulink and C# Program. The main control unit can be an actual vehicle's device or an IPC equipped with Prescan and Matlab/Simulink. The experimental process and results are reflected in the simulation animation, Simulink real-time display frame, or UI interface developed in the main control unit or recorded log files.

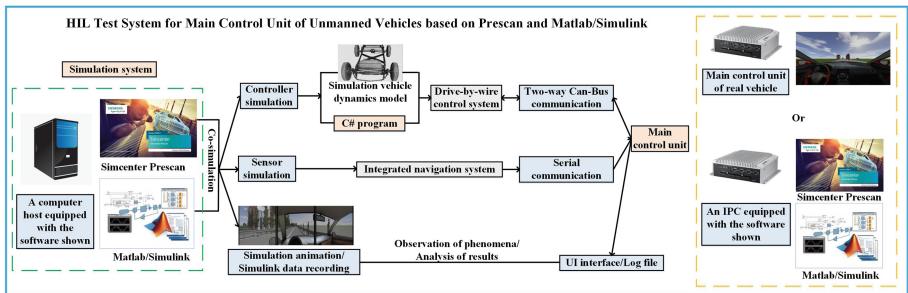


Fig. 1. System structure

The main research contents and the innovations of this paper are divided into 3 parts. (1) Simulate the integrated navigation system, and transmit all

kinds of data to the main control unit according to the communication protocol and communication mode. (2) Simulate the drive-by-wire control system, and realize the transmission of vehicle control signals to control the simulated vehicle driving and receive the feedback information of the simulated vehicle to realize the closed-loop simulation process. (3) Study the simulation method of the main control unit of the unmanned vehicle, which can verify the function of the test system and the feasibility of the communication protocol and communication method used in the test system without the main control unit of the actual vehicle.

3 Integrated Navigation System and Drive-by-Wire Control System Simulation

The main control unit needs to be able to access many real sensors and controllers in the actual vehicle test. While in the HIL test system, the hardware devices need to be simulated. In the process of vehicle simulation, the simulation system continuously outputs data to the main control unit and receives data from it.

3.1 Integrated Navigation System Simulation and Data Simulation

The integrated navigation system can provide information such as latitude, longitude, altitude, velocity, attitude angle (Roll, Pitch, Yaw), heading angle, and so on. It may include Global Positioning System (GPS), Inertial Navigation System (INS), Global Navigation Satellite System (GNSS), and so on. Each simulated vehicle will generate a “SELF_Demux” module in Matlab/Simulink, which can output real-time data such as latitude, longitude, altitude, velocity, heading angle, attitude angle, and so on, and all the information can be regarded as a true value because the simulation is ideal. This paper selects an integrated navigation system as the simulation object, taking “INS” message as an example.

The message contains the header, INS status, POS status, latitude, longitude, altitude, velocity, roll angle, pitch angle, heading angle, update time, CRC check value, etc., separated by commas, and the message data type is “String”. The simulation involves converting the Double-type data output by the module to String-type data, and the Char-type data such as punctuation marks and “#” need to be simulated, and the 8 bit CRC check value needs to be calculated. Compared with the Matlab language, it is easier to write in C language, and the execution of C language is faster and more efficient. The “S-function” module in Simulink can be selected for this purpose. It can be used to write programs in C language in the platform of Matlab. The application flow of “S-function” is shown in Fig. 2.

3.2 Drive-by-Wire Control System Simulation and Data Simulation

On the one hand, the drive-by-wire control system receives data from the main control unit to control the vehicle, on the other hand, the data information

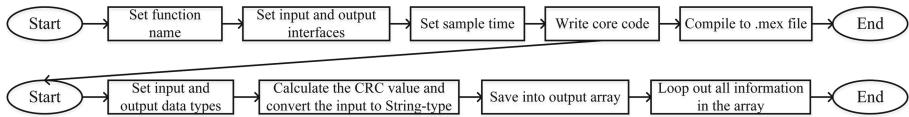


Fig. 2. The application flow of the “S-function” module

of the vehicle is fed back to the main control unit, which is a two-way data transmission process. The type of vehicle can be selected and added in Prescan GUI, and the vehicle dynamics module generated in Simulink provides interfaces to control the simulated vehicle. The main input interfaces are “Steer (deg)”, “Throttle (%)”, “Brake Pres (bar)” and “Auto_GearShift”. The “Steer” input can be steering wheel angle or the front wheel angle multiplied by the gear ratio. The “Throttle” input can be the accelerator pedal range. The “Brake Pres” input can be the pressure value converted from the brake pedal range. The “Auto_GearShift” input can be the vehicle gear value, such as 1 for forward gear. The received data are processed and connected to the input interface, and the state of the simulated vehicle will be completely controlled by the main control unit. At the same time, the module provides the output interface called “Perf Params”, which can output real-time data of the simulated vehicle and it can be used to simulate feedback data of the vehicle. Therefore, the simulation and data simulation of the drive-by-wire control system can be realized.

4 Realization of Communication Mode Based on Simulink and C# Program

The simulation according to the communication protocol and communication method of the actual device can make the main control unit “think” the real sensor and controller are connected to it, and the actual vehicle is being controlled to run, which makes the simulation system closer to reality, and the conclusions drawn from the test system are highly reliable.

4.1 Integrated Navigation System Communication

The communication mode of the integrated navigation system is serial communication. The serial data transmission program is written.

The serial port number, baud rate, etc. are needed to set in the serial data transmission program, so that the data simulated in the “S-function” module can be sent to the main control unit through serial communication. Various types of serial lines such as RS232 can be used to connect the simulation system to the main control unit. According to the research in this paper, due to the small data volume and low frequency of the simulated integrated navigation system, the non-blocking mode can be used. This mode can increase the running speed of the simulation project without affecting the real-time simulation.

4.2 Drive-by-Wire Control System Communication

In the practical application, the drive-by-wire control system mostly uses Can-Bus communication mode to transmit data, while the communication module (Can Transmit) in Simulink only supports the recognition of Can devices from several foreign companies, which are higher priced than the domestic Can equipment and are seldom used in China. This paper studies a general method, which is to transfer the simulated data in Simulink to the local C# program, and use the Can device to transfer the Can data to the main control unit. In this paper, a domestic Can device is taken as an example. One end of the device is the USB line. When the simulation system has no Can hardware interface, the USB line can be used.

Communication can be implemented as follows. The UDP (User Datagram Protocol) network sending program is written, and the target IP address is set as the local loopback address “127.0.0.1”, and the data are packed, and the local UDP communication method is completed to transmit the data. In general, the domestic Can device’s information contains some development programs for it under different IDEs which can realize Can data transmission and reception by the device. For instance, the selected device’s information contains a development program written in C# language. In this C# program, a UDP communication program is written to receive the vehicle feedback data transmitted by Simulink, and the required data for different Can frame IDs are extracted according to the communication protocol. At the same time, the program can be used to receive the Can data transmitted from the main control unit to the drive-by-wire control system. After receiving, the data can be transmitted to Simulink through the local UDP communication method. The UDP network receiving program is written to receive and analyze data to control the simulated vehicle.

This method studied in this paper is not constrained by the “Can Transmit” module and it does not require the Can hardware interface of the simulation system.

5 The Main Control Unit Simulation and Experimental Test

In this paper, a simulation method of the main control unit is studied. It is simulated by an IPC equipped with Prescan software and Matlab/Simulink software. The effectiveness of the simulation system can be tested repeatedly without the actual vehicle’s main control unit, and the feasibility of the communication protocol and communication function used can be verified.

Step 1: The first step is to import the same GPS trajectory in the Prescan of the simulation system and the Prescan of the IPC. The “Drive model” of the simulation vehicle in IPC is set to “Path Follower”, and a “Path Follower” module will be generated in Simulink. The function of this module is to present the function of controlling vehicle tracking. Depending on the GPS trajectory and desired speed, etc., the simulated vehicle is controlled to complete the tracking

driving. The required inputs are the vehicle's position information, the real-time speed, etc. The output is the control values of the steering wheel angle, throttle range, and brake pressure. Because the input and the output of the module are similar to the main control unit and it has functions such as controlling vehicle tracking, this method can be regarded as the simulation of the main control unit of the actual vehicle.

Step 2: Through the integrated navigation system simulation method studied in this paper, the input information required by the "Path Follower" module is sent from the simulation system to the IPC through serial communication. Serial data are received in the IPC's Simulink, and the "S-function" module for data analysis is written, and the data are connected to the "Path Follower" module after analysis.

Step 3: By using the research of the drive-by-wire control system simulation and Can-Bus communication method in this paper, the vehicle control value calculated by the "Path Follower" module according to the input of the simulation system is transmitted to the simulation system in real-time to control the movement of the simulated vehicle. At the same time, it receives the simulated vehicle's status information feedback and completes the closed HIL simulation. The simulated vehicle is driving along the GPS track in the whole test by the control of the IPC. Part of the simulation process of an experimental test is shown in Fig. 3.

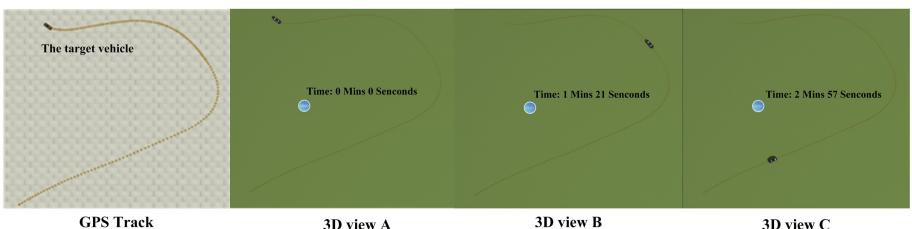
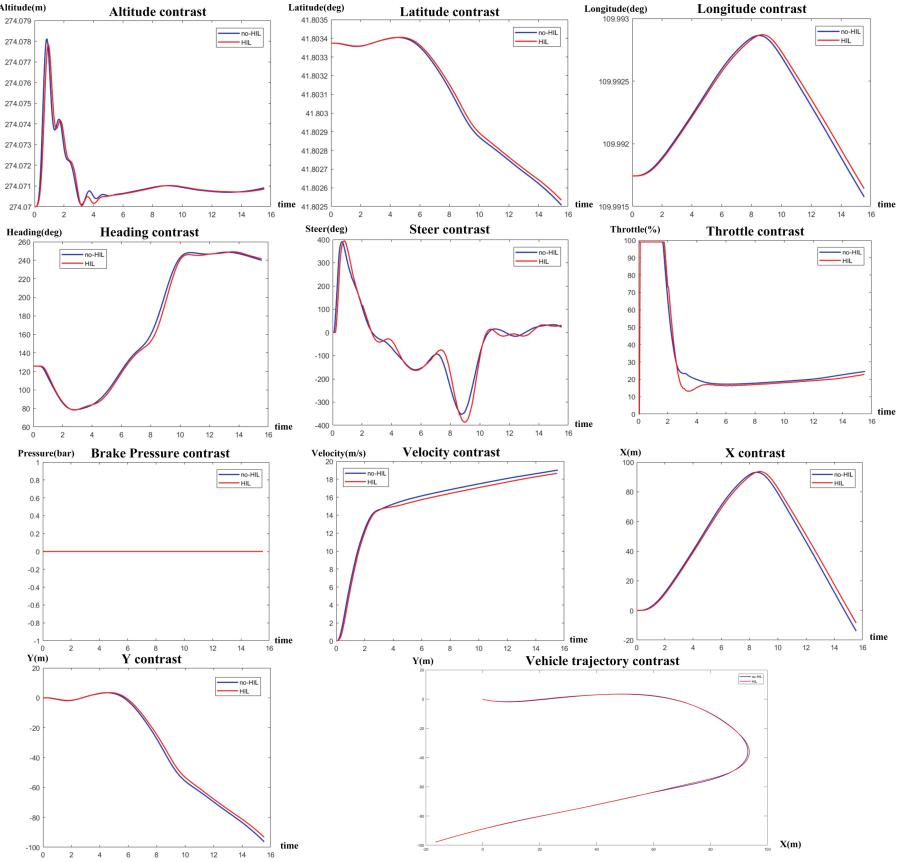


Fig. 3. Part of an experiment

The simulated vehicle is not controlled by the external main control unit, and the "Path Follower" module is added to the simulation system to test the vehicle tracking function under the same GPS track. This no-HIL test is compared with the HIL test studied in this paper to test whether the system function is normal when the simulation system is involved in the interaction with the outside world and the reasonability of the main control unit simulation method. In an experimental simulation animation, the vehicle running status of no-HIL and HIL is consistent. The data comparisons in an experiment are shown in Fig. 4. The X and Y values refer to the position of the vehicle in the Prescan world coordinate system, and the blue line is no-HIL, the red line is HIL.

The average difference between the no-HIL method and the HIL method in the vehicle tracking process is shown in Table 1.

**Fig. 4.** Data comparisons of an experiment**Table 1.** The average difference between two methods

Data type	The average difference
Altitude (deg)	0.0001314
Latitude (deg)	0.0000136
Longitude (deg)	0.0000307
Heading (deg)	2.3693955
Velocity (m/s)	0.3792413
Steer (-500~+500 deg)	18.536966
Throttle (0~100%)	2.5194463
Brake pressure (0~150)	0
X (m)	2.5531712
Y (m)	1.5149769

Through the graphs in Fig. 4 and the data in Table 1, the driving trajectory of the simulated vehicle in the two methods is consistent, and the trajectory in the “Vehicle trajectory contrast” graph is the same as the GPS track in Fig. 3, and the average difference between the vehicle position data, the execution control data and the vehicle status data of the two methods is low, which verifies the function of the HIL test system and the feasibility of the communication protocol and communication method used. Therefore, the HIL test system studied in this paper can be applied to test the function of vehicle tracking of the actual vehicle’s main control unit and the communication protocol and communication method used in reality.

6 Conclusions

In this paper, a HIL test system for the main control unit of unmanned vehicles based on Prescan and Matlab/Simulink is studied. According to the communication protocol and the communication method, the simulation of the integrated navigation system and the drive-by-wire control system is studied, which covers the main hardware devices of the automatic driving system. In addition, this paper studies a simulation method of the main control unit, which can test the test system studied. The simulation vehicle and driving environment are built, and the GPS trajectory is imported, and the simulation system is linked to the main control unit to control the vehicle, and the HIL test for the main control unit of the unmanned vehicle can be completed. In the stage of the rapid development of automatic driving technology, the HIL test system can save a lot of resources and has great practical significance.

Acknowledgements. This work is partially supported by the Beijing Natural Science Foundation under Grant #L191002. The authors would also like to thank the insightful and constructive comments from anonymous reviewers.

References

1. Gu, L., Sun, T.: Simulation comparison of AEB system with different control strategies in PreScan. Agric. Equip. Veh. Eng. **56**, 76–79 (2018). <https://doi.org/10.3969/j.issn.1673-3142.2018.10.019>
2. Wu, Z., Yang, J., Huang, L.: Study on the collision avoidance strategy at unsignalized intersection based on PreScan simulation. Procedia Soc. Behav. Sci. **96**, 1315–1321 (2013). <https://doi.org/10.1016/j.sbspro.2013.08.149>
3. Zhang, Q., Chen, D., Li, Y., Li, K.: Research on performance test method of lane departure warning system with PreScan. Springer, Heidelberg (2015)
4. Ruhai, G.E., Yating, C., Yongdong, X., Ming, X.U.: Warning rules of lane change based on Prescan simulation. J. Chongqing Jiaotong Univ. **35**, 172–176 (2016). <https://doi.org/10.3969/j.issn.1674-0696.2016.01.33>
5. Hendriks, F., Tideman, M., Pelders, R., Bours, R., Liu, X.: Development tools for active safety systems: Prescan and VeHIL. In: 2010 IEEE International Conference on Vehicular Electronics and Safety (ICVES). IEEE Press, Shanghai (2010). <https://doi.org/10.1109/ICVES.2010.5550948>

6. Molenaar, R., Bilsen, A.V., Made, R.V.D., Vries, R.D.: Full spectrum camera simulation for reliable virtual development and validation of ADAS and automated driving applications. In: IEEE Intelligent Vehicles Symposium, pp. 47–52. IEEE Press, Xi'an (2015). <https://doi.org/10.1109/IVS.2015.7225661>
7. Nacu, C.R., Fodorean, D., Husar, C., Grovu, M., Irimia, C.: Towards autonomous EV by using Virtual Reality and Prescan-Simulink simulation environments. In: 2018 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), pp. 401–406. ISPECE, Xi'an (2018). <https://doi.org/10.1109/SPEEDAM.2018.8445211>
8. Zaki, S.P., William, M.M., Soliman, B.K., Alexsan, K.G., Khalil, K., El-Moursy, M.: Traffic signs detection and recognition system using deep learning [arXiv:2003.03256](https://arxiv.org/ftp/arxiv/papers/2003/2003.03256.pdf) (2020). <https://arxiv.org/ftp/arxiv/papers/2003/2003.03256.pdf>
9. Kim, Y., Tay, S., Guanetti, J., Borrelli, F., Miller, R.: Hardware-in-the-loop for connected automated vehicles testing in real traffic. In: The 14th International Symposium on Advanced Vehicle Control, AVEC 2018, Beijing (2018). [arXiv:1907.09052](https://arxiv.org/abs/1907.09052)



Keywords Extraction Algorithm of Financial Review Based on Dirichlet Multinomial Model

Jinxuan Li¹, Yawen Li², and Zhe Xue^{1(✉)}

¹ Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia, School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

xuezhe@bupt.edu.cn

² School of Economics and Management,
Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. At present, the main method of keyword extraction is usually based on topical discovery model, and the most popular model is LDA (Latent Dirichlet Allocation) topical model. However, LDA model has poor performance in dealing with short texts, which is the imperfection of the current topical discovery model. This paper propose a keywords extraction model SDMM (Statistical Dirichlet Multinomial model) to extract the keywords of short financial review. This keywords extraction model use DMM (Dirichlet Multinomial Model) for topic discovery, and calculate the similarity between the words and keywords, thus improve the performance of keyword extraction for short texts. Experimental results on several financial review datasets show that the proposed model is better than the existing models in evaluation indicator Precision, Recall and F-measure.

Keywords: Keywords extraction for short texts · DMM · Statistical method

1 Introduction

With the development of the Internet, more and more financial reviews and news content appear on the Internet, including many short messages, reviews, etc. These essays still contain a lot of valuable content. In the field of keyword extraction, the existing topical translation models use LDA topic model for topic discovery and achieving the topic consistency between words and keywords. Then we use the topic constraints to across the lexical difference between words and keywords which leads to the problem that keywords do not appear in the text to improve the keyword extraction performance of short text. When it comes to the long text topic discovery which is represented by news, LDA has a pretty good performance. The reason is that long text has more words and

text features. However, short text has few words and less features. If short text is handled directly by LDA model, the performance of short text topic develop and keyword extraction will be poor. In order to overcome the above shortcomings, this paper proposes the SDMM topic translation model using the DMM model which has good performance on the topic discovery of short text, and combines with the words statistics for the keyword extraction of short financial review text. Compared with traditional methods, this model can effectively improve the extraction performance of short text keywords.

2 Related Work

At present, keyword extraction methods are mainly divided into four categories: statistical-based method, topic discovery-based method, graph model-based method and machine learning-based method.

Statistical-based method mainly use the statistical information of words in text to extract keyword. This method does not need training data and external knowledge but uses simple statistical rules (word frequency, TF-IDF and co-occurrence information, etc.) to extract keywords of the preprocessed document [1]. Yan et al. [2] proposed a text keyword extraction method based on word frequency statistics, and verified it on the Chinese and English data sets. Su et al. [3] proposed a keyword extraction method based on word order statistical combination through word order combination, part of speech tagging, stop using word filtering. Siddiqi et al. [4] combine word frequency information with word spatial distribution and propose an automatic keyword extraction method suitable for any language.

Topic-based methods usually use the topic model LDA to extract keywords. Siu et al. [5] found topic information and keyword information by training HMM model, and got good results on test set. In literature [6, 7], LDA topic model is used to automatically extract keywords. Ding et al. Proposed the TSTM (topic specific translation model) theme translation model. The model uses LDA model to find the theme, and then calculates the alignment probability of words and keywords under the theme, which improves the performance of micro blog keyword automatic tagging [8]. The key words extraction method based on network graph should first construct the language network graph of the document, then analyze the network graph, find the words or phrases that play an important role in the whole network graph, and extract these words or phrases as the key words. Mihalcea and Tarau [9] proposed a graph model-based Textrank method, which takes the graph node as the candidate word and the edge as the co-occurrence relationship between words. According to PageRank algorithm, some of the highest-ranking keywords are selected. In reference [10, 11], Textrank algorithm is improved to improve the performance of keyword extraction. Based on machine learning method, the model parameters are first learned on the training corpus, and then the test corpus is automatically extracted with the learned parameters. Liu et al. [12] proposed using decision tree learning method to extract text keywords. Sarkark et al. [13] proposed a keyword extraction

method based on neural network for science and technology articles. In reference [14, 15], a keyword extraction method based on support vector machine is proposed. Zhang et al. [16] used conditional random fields to realize automatic tagging of keywords.

The SDMM model proposed in this paper is a topic-based keyword extraction method, which is combined with statistical machine translation to achieve keyword extraction of short text.

3 SDMM Topical Translation Model

The SDMM topic translation model of short text keyword extraction proposed in this paper mainly uses DMM as the topic discovery model. The difference between DMM and LDA is that LDA assumes that each article has a topic distribution, which is reasonable for long text with a large number of words. Therefore, LDA has achieved good performance in long text topic discovery represented by news. DMM model assumes that there is only one topic distribution in the corpus and only one topic in each article. Obviously, this is not reasonable for long text, but it is suitable for short text, because short text has few words. Therefore, each short text having one topic is more realistic than one topic distribution. This is an important theoretical basis for DMM to solve the problem of the text feature sparsity and improve the performance of short text topic discovery.

In addition, Yin et al. Proposed GSDMM model for short text clustering based on DMM model [17]. This model uses the collapsed Gibbs method to sample the topic of short text, which is more efficient than the EM method used by Nigam. This method converges fast and has done well in solving the problem of short text's sparsity. It provides the feasibility for this paper to use the DMM model of the collapsed Gibbs sampling method for subject discovery and combined with statistical machine translation for short text key word extraction.

3.1 DMM Model Background Introduction

Dirichlet multinomial mixture (DMM) model is a graph model-based method proposed by Nigam et al. Its construction diagram is shown in Fig. 1.

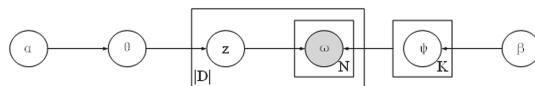


Fig. 1. DMM models construction diagram

In this diagram, k is the number of topics and D is the corpus. According to DMM model, when a document D is generated, a topic Z will be selected from the corpus topic distribution θ . Then the words in document D are generated

according to the word distribution of the given topic Z . Therefore, the likelihood function which DMM generates of its content is shown in equality 1.

$$p(d) = \sum_{k=1}^K p(d|z=k)p(z=k) \quad (1)$$

3.2 SDMM Model

The construction diagram of SDMM is shown in Fig. 2. The model includes the following two aspects:

Using DMM model. In SDMM model, when users write a short document D , they first select a topic z from the corpus topic distribution θ , and then generate words according to the word distribution θ under the topic z . The keywords of short document have the same topic z , which ensures the topic consistency of short documents words and keywords.

Combined with statistical machine translation model. Based on the topic consistency principle of short documents words and keywords, SDMM model learns the correlation probability q of words and keywords under a specific topic and achieve the purpose of generating keywords for short text without keywords pre-marked.

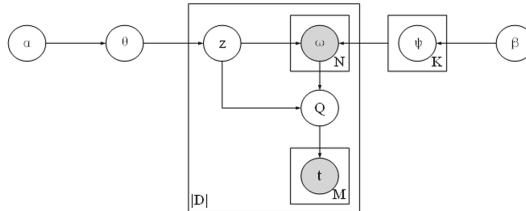


Fig. 2. SDMM models construction diagram

Using DMM which has better performance in short text topic discovery can improve the accuracy of topic discovery, and then improve the accuracy of the correlation probability of words and keywords under the constraints of the same topic, so as to achieve the purpose of short text keyword extraction.

According to Algorithm 1, all words and keywords in the short document of part d share the same topic. Q is the correlation probability of words and keywords in short document under specific topic, and is the related probability of words and keywords under the condition of topic K .

Algorithm 1. SDMM

Input: Text set, identified keywords set

Output: result of keyword extraction

Step 1. Preprocessing and segmentation of text data

Step 2. Using Gibbs sampling method to classify all words and identified keywords in text set

Step 3. After the topic classification is completed, the correlation probability matrix Q between words and keywords in corpus is calculated, then expand the keyword set and recalculate the correlation probability matrix after expanding the keyword

Step 4. For every short text in text set, calculate the topic

Step 5. According to the short text topic and correlation probability matrix Q , Calculating the probability of keyword m in short text

Step 6. Extract keywords according to probability ranking

3.3 SDMM Models Parameter Estimation

Firstly, we use the collapsed Gibbs method to sample the words and keywords in the training short document set. For short document w_d , select a topic by:

$$p \propto \frac{N_{-d,k} + \alpha}{N_{-d} + K\alpha} * \prod_{n=1}^{N_d} \frac{N_{-d,k}^{w_{dn}} + \beta}{N_{-d,k} + V\beta} * \prod_{m=0}^{M_d} \sum_{n=0}^{N_d} \frac{M_{-d,k}^{w_{dn}t_{dm}} + \beta}{M_{-d,k}^{w_{dn}} + T\beta} \quad (2)$$

Among above equality, $N_{-d,k}$ is the number of short documents whose topic are k after removing the short text w_d from corpus D , N_{-d} is the number of documents which remove the short text w_d from corpus D , that is, $N_{-d} = |D| - 1$. $N_{-d,k}^{w_{dn}}$ represents the number of w_{dn} words with the topic k after the short document w_d is removed from the corpus. $M_{-d,k}^{w_{dn}t_{dm}}$ indicates the number of times the keywords t_{dm} and w_{dn} appear together in a short document whose topic are k . Secondly, when the topic of words and keywords in short documents are stable, the correlation probability q of words and keywords under specific topic will be estimated by the following equality 3:

$$Q_{t,w}^k = \frac{N_{t,w}^k}{N_w^k}. \quad (3)$$

Among above equality, $N_{t,w}^k$ is the number of times that keyword t and word w appear together in a short document whose topic is K . When $k = 1, \dots, K$, the value of the topic with the maximum probability $p(z_d = k|d)$ is obtained as the topic value of the d -th short document. Finally, using the topic distribution and the correlation probability Q of words and keywords under a specific topic to extract keywords from the test data set through equality 4

$$p(t_{dm}|w_d, Q) \propto \sum_{n=1}^N p(t_{dm}|z_d, w_{dn}, Q) * p(w_{dn}|w_d) \quad (4)$$

4 Experiment and Result Analysis

4.1 Data Set and Experimental Environment

This paper uses the financial review data under the three topics of interest rate marketization reform, US dollar interest rate increase and Internet financial development in Sina Weibo financial topics. The specific parameters of the data set are shown in Table 1.

Table 1. Financial review dataset

Dataset name	Number of reviews	Number of keywords
Interest rate marketization reform	3000	5
US dollar interest rate increase	3000	7
Internet financial development	3000	6

The experimental environment is the tensorflow deep learning framework running in Ubuntu environment. NVIDIA cuda9.0 and cudnn7.5 deep learning library are configured to accelerate GPU calculation. Python environment is Python 3.5.

4.2 Comparison Experiment Settings

In this paper, SDMM model is compared with the following three methods:

LDA topic model: LDA is a probability generation model based on topic. When using LDA to extract keywords, firstly, the topic distribution of documents and candidate keywords is calculated; then calculate the topic similarity of the documents and candidate keywords, and choose the candidate words whose similarity are the top as the keywords.

IBM model-1: IBM model-1 is the representative of translation model. It uses the word alignment model and the co-occurrence relationship between words and keywords to learn the translation probability from words to keywords.

TSTM model: a LDA based topic translation model. This model first samples the words and keywords in the training set documents by learning LDA model; secondly, it studies the translation probability from words to keywords under specific topics; finally, it calculates the keyword score of each document by using the learned translation probability, the importance of words in the documents and the topic distribution of the test set. This method is similar to the SDMM model proposed in this paper. The difference is that the SDMM model combines the DMM model to solve the problem of short documents sparsity, and then improves the performance of topic discovery and the accuracy of keyword extraction.

4.3 Experimental Results

The initialization of SDMM model is $\alpha = 0.5$, $\beta = 0.1$, $k = 4$. The number of iterations is 2000, and the IDF value is the importance of short documents words.

The average precision, recall and F-measure value of four keyword extraction methods under three topics are shown in Table 2.

Table 2. Results of four keyword extraction methods under three topics

Topic set	Name of method	Precision	Recall	F-measure
Interest rate marketization reform	LDA	0.237	0.168	0.197
	IBM model-1	0.386	0.332	0.357
	TSTM	0.454	0.416	0.434
	FRM_DMM	0.484	0.498	0.491
US dollar interest rate increase	LDA	0.212	0.178	0.194
	IBM model-1	0.363	0.352	0.357
	TSTM	0.442	0.434	0.438
	FRM_DMM	0.463	0.442	0.452
Internet financial development	LDA	0.205	0.182	0.193
	IBM model-1	0.354	0.373	0.363
	TSTM	0.422	0.448	0.435
	FRM_DMM	0.452	0.439	0.445

Compared with the experimental algorithm, SDMM improves the evaluation indexes compared with LDA topic model and IBM model-1 algorithm, which shows that SDMM is effective in improving the performance of keyword extraction by combining topic based and statistics method. TSTM is also based on topic and statistical method, which is similar to the SDMM proposed in this paper. However, SDMM solves the topic sparsity problem of short documents by modifying the topic model, and finally improves the keyword extraction performance.

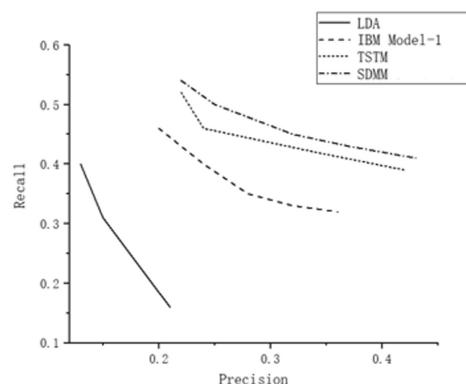


Fig. 3. Keywords extraction average PR curve

Figure 3 shows the average PR curve of keyword extraction of four algorithms. Each point on the curve represents the precision and recall values with different number of keywords extracted. The first point on the right represents the number of keywords extracted is 1, and the number of keywords extracted increases from right to left successively to 5. The closer the curve is to the upper right corner, the better the extraction performance is. Because the average number of keywords in each short document in the data set is 1.08, when the number of keywords extracted is 1, the accuracy rate is the highest and the recall rate is the lowest. When the number of keywords extracted is reduced from 5 to 1, the keywords extracted by each model will contain fewer marked keywords in the document, the recall rate is reduced and the accuracy rate is increased. The experimental results of each model are analyzed as follows.

The curve of LDA model is located at the bottom left, and the decline speed is the fastest, which shows that the extraction performance of key words is the worst, and the stability is poor when the number of key words changes. The reason is that the model only considers the topic consistency of words and keywords, and does not add the influence of short text existing words on keyword extraction. Moreover, due to the small number of words in short text and the lack of vocabulary information, LDA has defects in short text topic discovery, which affects the performance of keyword extraction.

The performance of IBM model-1 model is better than LDA model, and the performance is more stable than LDA model when the number of keywords is changed. But because it only considers the co-occurrence relationship between words and keywords, and does not consider whether the theme of words and keywords is the same, the result of extraction is not the best.

TSTM model's short text keyword extraction performance is second only to TTM - DMM model, and the curve decline trend is gentle, which shows that the model has good performance in extraction performance and stability, but the model's short text keyword extraction performance is still not the best, because although the model introduces LDA model and statistical machine translation, the advantages of the two models are integrated. Because of the defects of LDA model in short text topic discovery, the performance of model topic discovery is not good, which affects the calculation of the translation probability of words and keywords under specific topic, so the extraction performance is not the best. The curve of SDMM model is closest to the upper right corner, indicating that the extraction performance is better than the other three models. The reason is that SDMM model not only solves the problem of topic consistency and the co-occurrence relationship between words and keywords, but also uses DMM as the main model of topic discovery, which has better performance on topic discovery of short text. Combined with statistical Machine Translation, the model can be more refined the translation probability from words to keywords under a specific topic can be accurately obtained, which improves the performance of keyword extraction in short text.

5 Conclusion

This paper innovatively proposes a FRM_DMM topic translation model for short text keyword extraction. By using the DMM model which has good performance on finding short text topics with sparse features and combining with statistical machine translation, the translation probability of words and keywords under specific topics is calculated, so as to achieve the purpose of tagging keywords for short text. Experimental results show that FRM_DMM model can effectively improve the performance of short text keyword extraction. In the future, we will continue to study how to further improve the accuracy of short text keyword extraction, and combine it with practical application to extract keywords in specific fields.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant (No. 61902037, No. 61532006, No. 61772083, No. 61802028), and Science and Technology Major Project of Guangxi (GuikeAA18118054).

References

1. Jingsheng, Z., Qiaoming, Z.: Review of research in automatic keyword extraction. *J. Softw.* **28**(9), 2431–2449 (2017)
2. Yan, L., et al.: Text keyword extraction method based on word frequency statistics. *JOCA* **36**(3), 718–725 (2016)
3. Su, X., et al.: Keywords extraction based on word order statistics and combination of Chinese text theme. *Comput. Eng. Des.* (6), 1647–1651 (2015)
4. Siddiqi, S., Sharan, A.: Keyword and key phrase extraction from single Hindi document using statistical approach. In: SPIN, pp. 713–718. IEEE (2015)
5. Siu, M.H., et al.: Unsupervised training of an HMM-based self-organizing unit recognizer with applications to topic classification and key-word discovery. *Comput. Speech Lang.* **28**(1), 210–223 (2014)
6. Xiaojian, L., Xie, F.: Keyword extraction method combining topic distribution with statistical features. *Comput. Eng.* **43**(7), 217–222 (2017)
7. Cho, T., Lee, J.H.: Latent keyphrase extraction using LDA model. *J. Korean Inst. Intell. Syst.* **25**(2), 180–185 (2015)
8. Ding, Z., et al.: Automatic hashtag recommendation for microblogs using topic-specific translation model. In: COLING 2012: Posters, pp. 265–274 (2012)
9. Mihalcea, R., Tarau, P.: TextRank: bringing order into text. In: EMNLP 2004. Unt Scholarly Works, pp. 404–411 (2004)
10. Yijun, G., Tian, X.: Study on keyword extraction with LDA and TextRank combination. *New Technol. Libr. Inf. Serv.* **30**(7), 41–47 (2014)
11. Jianfei, N., Jiangzhen, L.: Using Word2vec with TextRank to extract keywords. *New Technol. Libr. Inf. Serv.* **6**, 20–27 (2016)
12. Liu, J., et al.: Keyphrase extraction based on topic feature. *Appl. Res. Comput.* **29**(11), 4224–4227 (2012)
13. Sarkar, K., et al.: A new approach to keyphrase extraction using neural networks. *IJCSI* **7**(2), 16–25 (2010)
14. Zhang, K., et al.: Keyword extraction using support vector machine. In: Lecture Notes in Computer Science, vol. 4016, pp. 85–96 (2006)

15. Chen, Y.N., et al.: Unsupervised two-stage keyword extraction from spoken documents by topic coherence and support vector machine, pp. 5041–5044 (2012)
16. Zhang, C.: Automatic keyword extraction from documents using conditional random fields. JCIS **4**, 1169–1180 (2008)
17. Nigam, K., et al.: Text classification from labeled and unlabeled documents using EM. Mach. Learn. **39**(2/3), 103–134 (2000)



Modelling of Wind Turbine with Battery Energy Storage System Based on Two-Level Switched System

Cheng Peng¹ and Yang Song^{1,2(✉)}

¹ Department of Automation, Shanghai University, Shanghai 200444, China
y.song@shu.edu.cn

² Shanghai Key Laboratory of Power Station Automation Technology,
Shanghai 200444, China

Abstract. This paper presented a new switched system model for wind turbine combined with a battery energy system (WT/BESS). There are two-level switching mechanisms in this WT/BESS model. When the maximum wind power which is captured is lower than the demand power, WT takes the maximum power point tracking (MPPT) control to meet the maximum wind power. Otherwise, WT takes variable pitch control to limit the output power to achieve the demand. The switching between the two control modes is described by a Markov process. On the other hand, in order to maintain the state of charging (SOC) of BESS, an SOC offset strategy is proposed to change the expected output power to keep the SOC at a mid-value, which is driven by the switching of the state of SOC.

Keywords: Switched system · Markov · SOC offset strategy

1 Introduction

In recent years, renewable energy such as wind energy has been used on a large scale. With the continuously increasing usage of wind energy, the control of wind turbines is receiving more and more interest around the world [1].

It is still a great challenge to establish an accurate and complete wind power generation model because of its complex dynamic characteristics. And the control strategy of the wind turbine could be diverse for different wind speed. Wind speed is divided into two regions by the required wind speed, when the wind speed rises from low to over the required wind speed, the control strategy will change from maximum power point tracking control to variable pitch control [2]. Furthermore, A Markov chain method is considered to describe the switching of the wind speed sub-regions to probably estimate of wind speed states [3].

The object of MPPT control is to meet the maximum power point by changing the load torque when the wind speed is comparatively low, and then improve the utilization rate of wind energy [4]. Many controllers are designed to improve

the control performance of MPPT control. In [5], a second-order sliding mode control (SOSMC) are proposed to ensure the better tracking. In [6], an adaptive torque controller is designed to realize tracking the maximum power point for variable-speed wind energy conversion systems.

The object of variable pitch control is to limit the output power of turbine by controlling the pitch angle when wind speed is high [7]. Several control methods have been reported such as traditional proportional-integral (PI) controller [8], gain-scheduled control method [9] have also been presented to solve the pitch angle control problem.

While the above method is not enough to maintain the output power of the wind turbine at a certain level in a rapidly changing wind speed environment. In this case, it is an effective scheme to increase the reliability and availability of the system with a battery energy storage system [10]. BESS can not only compensate power to grid when the wind is too weak, but also absorb the surplus output power of turbine for the case of high wind speed.

However, Combining the wind turbine with a battery energy storage system will also make the model and control of WT/BESS more complex. Considering the above, we have proposed a multi point piecewise linearization method combined with switched system, modelling the WT/BESS as a two-level switched system. One switching in the WT/BESS relies on the state of battery energy, while another switching is dependent on the switching of wind speed Markov model.

The remainder of this paper is organized as follows: Sect. 2, the modelling process of WT/BESS is explained. Section 3 presents and discusses the hierarchical modeling and simulation of WT/BESS with stateflow in MATLAB/Simulink. Finally, Sect. 4 draws some conclusions.

2 Modelling of Wind Turbine

2.1 Modelling of Wind Speed

It is clear that the characteristics of wind speed affect the dynamics performance of WT/BESS dominantly. In this paper, we use a transition matrix to describe wind speed $v(t)$ as following,

$$v(t) = v_{\sigma(t)} + v_{\omega}(t) \quad (1)$$

where $v_{\omega}(t)$ is the disturbance of wind speed, $v_{\omega}(t) \in (-d/2, d/2)$, and average wind speed $v_{\sigma(t)}$ is modeled by a Markov process $\sigma(t) \in \{1, \dots, N\}$. The transition probability is given by $p_{ij} = P(\sigma(k+1) = j | \sigma(k) = i), \forall k \in N$.

2.2 Modelling of WT

As shown in Fig. 1, a WT/BESS system consist of two mainly parts: a wind turbine system combines a battery energy storage system.

The expected output power of the wind turbine P_e^* which is related to the state of SOC and the demand power of power grid P_d , and in this paper, we

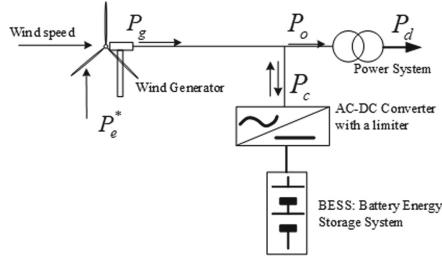


Fig. 1. WT/BESS system

assume P_d is a constant. P_o is the output power of WT/BESS, and it is composed by the instantaneous power generated by the wind turbine P_g and the compensation power P_c of battery system, where P_c can be positive or negative, representing the charge or discharge SOC of BESS respectively.

The dynamics of the wind turbine can be expressed as

$$\dot{\omega} = \frac{1}{J} (\tau_{aero} - \tau_{\sigma(t)}) \quad (2)$$

where ω is the rotor angular velocity, J and $\tau_{\sigma(t)}$ are the rotational inertia and the load torque of the turbine respectively, τ_{aero} the aerodynamic torque is given by

$$\tau_{aero} = \frac{1}{2\omega} \rho \pi R^2 v^3 C_p(\lambda, \beta) \quad (3)$$

where ρ is the air density, R is the rotor radius, v is the wind speed, β is the pitch angle of the turbine blade, power coefficient $C_p(\lambda, \beta)$, which is given by the following reflects the conversion efficiency of wind energy of the WT [11].

$$C_p(\lambda, \beta) = 0.4654 \left(\frac{116}{\lambda'} - 0.4\beta - 5 \right) e^{-\frac{20.24}{\lambda'}} \quad (4)$$

where $\frac{1}{\lambda'} = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{\beta^3 + 1}$, tip-speed ratio $\lambda = \omega R/v$.

Figure 2 shows the instantaneous output power curve of turbine, where $P_g = \tau_{\sigma(t)}\omega$. The operation of the turbine is divided into three regions successively with the increase of wind speed. In Region 1, wind turbine does not start since the wind speed is too slow (smaller than the cut-in wind speed).

In Region 2, also called the under-power region, the wind turbine can output power but cannot reach the demand. In this region, the control object is to maximize aerodynamic efficiency, i.e. the power coefficient C_p achieve the maximum power coefficient C_{pmax} at $\beta = 0^\circ$, the control strategy is adjusting the load torque $\tau_{\sigma(t)}$ in real time and keeping the pitch angle facing the winds (zero degree). Generally, we use PI controller as below to control the pitch angle quickly and accurately. Select k_p, k_i as the PI controller parameters,

$$\dot{\beta} = k_p \dot{\omega} + k_i (\omega_e - \omega) \quad (5)$$

where ω_e is the rotor angular velocity of the operation point (ω_e, β_e, v_e) .

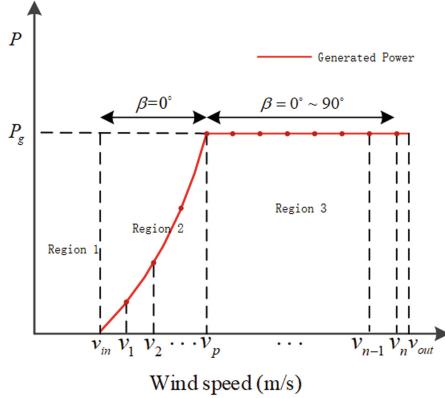


Fig. 2. The power curve of turbine

On the other hand, the load torque is set to $\tau_{\sigma(t)} = K_{\tau}\omega^2$, where coefficient $K_{\tau} = \frac{1}{2\lambda_d^3}\rho\pi R^5 C_{p_{max}}$, and λ_d is the tip-speed ratio at $C_{p_{max}}$. Apply it into Eq. (2), we can get

$$\dot{\omega} = \frac{1}{2J}\rho\pi R^5\omega^2\left(\frac{C_p}{\lambda^3} - \frac{C_{p_{max}}}{\lambda_d^3}\right) \quad (6)$$

Equation (6) is a highly nonlinear function, usually the control method is based on linearizing the nonlinear turbine dynamics at a special operation point. However, single operation point can not accurately show the dynamic characteristic of WT, so a multipoint piecewise linearization method can be used to deal with the problem. As shown in region 2 in Fig. 2, the operation point of each wind speed subinterval is its mid-value, which can be switched at the same time when the wind speed changes. And the Taylor expansion can be used to linearize the Eq. (6) and bring it into Eq. (5) to get

$$\begin{aligned} \dot{\omega} &= f_{\beta}^1(\beta_{i,e})(\beta - \beta_{i,e}) + f_{\omega}^1(\omega_{i,e})(\omega - \omega_{i,e}) + f_v^1(v_{i,e})(v - v_{i,e}) \\ \dot{\beta} &= k_P(f_{\beta}^1(\beta_{i,e})(\beta - \beta_{i,e}) + f_v^1(v_{i,e})(v - v_{i,e})) \\ &\quad + (k_p f_{\omega}^1(\omega_{i,e}) + k_I)(\omega - \omega_{i,e}) \end{aligned} \quad (7)$$

Where

$$\begin{aligned} f_{\beta}^1(\beta_{i,e}) &= \frac{\partial f^1(\beta, \omega, v)}{\partial \beta} \Big| (\beta_{i,e}, \omega_{i,e}, v_{i,e}), \quad f_{\omega}^1(\omega_{i,e}) = \frac{\partial f^1(\beta, \omega, v)}{\partial \omega} \Big| (\beta_{i,e}, \omega_{i,e}, v_{i,e}), \\ f_v^1(v_{i,e}) &= \frac{\partial f^1(\beta, \omega, v)}{\partial v} \Big| (\beta_{i,e}, \omega_{i,e}, v_{i,e}), \quad f^1 = \frac{1}{2J}\rho\pi R^5\omega^2\left(\frac{C_p}{\lambda^3} - \frac{C_{p_{max}}}{\lambda_d^3}\right), \end{aligned}$$

$i \in \{1, 2, \dots, p\}$, and the wind speed $(v - v_{i,e})$ is the disturbance.

In Region 3, the so called over-power region, the wind energy will exceed the demand power, the control object is to limit the output power by adjusting blade pitch angle β to change the power coefficient C_p , the load torque $\tau_{\sigma(t)}$

would be set as the ratio of the excepted output P_d and $\omega_{i,e}$. Similarly, by taking multiple operation points, Linearizing Eq. (6) in this region to get

$$\begin{aligned}\dot{\omega} &= f_\beta^2(\beta_{i,e})(\beta - \beta_{i,e}) + f_\omega^2(\omega_{i,e})(\omega - \omega_{i,e}) + f_v^2(v_{i,e})(v - v_{i,e}) \\ \dot{\beta} &= k_P(f_\beta^2(\beta_{i,e})(\beta - \beta_{i,e}) + f_v^2(v_{i,e})(v - v_{i,e})) \\ &\quad + (k_p f_\omega^2(\omega_{i,e}) + k_I)(\omega - \omega_{i,e})\end{aligned}\quad (8)$$

Where

$$\begin{aligned}f_\beta^2(\beta_{i,e}) &= \frac{\partial f^2(\beta, \omega, v)}{\partial \beta} |_{(\beta_{i,e}, \omega_{i,e}, v_{i,e})}, \quad f_\omega^2(\omega_{i,e}) = \frac{\partial f^2(\beta, \omega, v)}{\partial \omega} |_{(\beta_{i,e}, \omega_{i,e}, v_{i,e})}, \\ f_v^2(v_{i,e}) &= \frac{\partial f^2(\beta, \omega, v)}{\partial v} |_{(\beta_{i,e}, \omega_{i,e}, v_{i,e})}, \quad f^2 = \frac{1}{J} \left(\frac{1}{2} \rho \pi R^5 \omega^2 \frac{C_p}{\lambda^3} - \frac{P_d}{\omega_{i,e}} \right), \\ i &\in \{p+1, p+2, \dots, n\}.\end{aligned}$$

2.3 Modelling of BESS

To alleviate the fluctuation of the output power of WT, the compensation power P_c of BESS is generally set to be $P_c = P_g - P_d$, i.e. the difference between the instantaneous power P_g and the demand power P_d . The SOC of battery as

$$SOC = \frac{P_g - P_d}{n V_b C_b K_b} \quad (9)$$

n is the number of batteries, C_b is the nominal capacity and V_b is the terminal voltage of battery.

In addition, in order to prolong the battery life, we proposed an SOC offset strategy as following. When the SOC is below a certain limit, we charge the battery by increasing the expected output power of the wind turbine, i.e. increasing the operation point. On the other hand, if the SOC is too high, the expected output power will be rised for BESS charging. The battery capacity is divided into several sections, from low to high, expressed by $S_1, S_2 \dots S_n$ respectively. In order to maintain the battery storage energy in the middle level, each section corresponds to the expected output power of a wind turbine, expressed by P_1, P_2, \dots, P_n respectively.

Suppose that when the expected output power changes from P_1 to P_2 , part of the operation point will shift. i.e. at the same wind speed v_s , the expected pitch angle β_s at the operation point will shift from β^1 to β^2 , and the expected rotor speed ω_s will also change from ω^1 to ω^2 . Hence, the model of WT/BESS is linearized at (β^1, ω^1) and (β^2, ω^2) respectively by distinguishing the different expected output power.

2.4 Modelling of WT/BESS

Figure 3 shows the switching process of the WT/BESS system, the operation point of the turbine is related to two switching conditions: SOC based switching

$\gamma(\xi(t))$ and wind speed Markov switching $\sigma(v(t))$, is the junction point between region 2 and region 3. In region 2 ($v < v_p$), the operation point is determined only by $\sigma(v(t))$, while in region 3 ($v_p < v$), the operation point would be determined by $\sigma(v(t))$ and $\gamma(\xi(t))$ together.

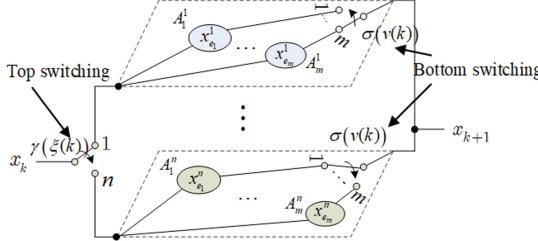


Fig. 3. Effect of battery storage energy on operation point

After linearization on the different equilibriums and notice (7)–(8), a two-level switched system with multiple equilibriums is presented as below,

$$\dot{x} = \left(A_i^j + BK_i^j \right) (x - x_i^j) + D_i^j (v_t - v_i^j) \quad (10)$$

Where

$$x = \begin{bmatrix} \beta \\ \omega \end{bmatrix}, x_i^j = \begin{bmatrix} \beta_i^j \\ \omega_i^j \end{bmatrix}, K_i^j = \left[k_{P,i}^j f \Big|_{\beta_i^j} \quad k_{P,i}^j f \Big|_{\omega_i^j} + k_{I,i}^j \right],$$

$$A_i^j = \begin{bmatrix} 0 & 0 \\ f \Big|_{\beta_i^j} & f \Big|_{\omega_i^j} \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, D_i^j = \begin{bmatrix} k_{P,i}^j f \Big|_{v_i^j} \\ f \Big|_{v_i^j} \end{bmatrix},$$

$$f \Big|_{\beta_i^j} = \begin{cases} \frac{\partial f_L(\omega, \beta, v)}{\partial \beta} \Bigg|_{\theta_i^j}, & i \in L^j \\ \frac{\partial f_H(\omega, \beta, v)}{\partial \beta} \Bigg|_{\theta_i^j}, & i \in H^j \end{cases}, f \Big|_{\omega_i^j} = \begin{cases} \frac{\partial f_L(\omega, \beta, v)}{\partial \omega} \Bigg|_{\theta_i^j}, & i \in L^j \\ \frac{\partial f_H(\omega, \beta, v)}{\partial \omega} \Bigg|_{\theta_i^j}, & i \in H^j \end{cases},$$

$$f \Big|_{v_i^j} = \begin{cases} \frac{\partial f_L(\omega, \beta, v)}{\partial v} \Bigg|_{\theta_i^j}, & i \in L^j \\ \frac{\partial f_H(\omega, \beta, v)}{\partial v} \Bigg|_{\theta_i^j}, & i \in H^j \end{cases},$$

$$\theta_i^j = \theta_{\sigma(v(t))}^{\gamma(\xi(t))} := (\omega_{\sigma(v(t))}^{\gamma(\xi(t))}, \beta_{\sigma(v(t))}^{\gamma(\xi(t))}, v_{\sigma(v(t))}^{\gamma(\xi(t))}),$$

$$i = \sigma(v(t)) = \{1, 2, \dots, N\} \text{ and } j = \gamma(\xi(t)) \in \{1, 2, \dots, M\},$$

$$f_L(\omega, \beta, v) = \frac{1}{2J} \rho \pi R^5 \omega^2 (C_p \lambda^{-3} - C_{p_{\max}} \lambda_d^{-3}),$$

$$f_H(\omega, \beta, v) = \frac{1}{J} \left(\frac{1}{2} \rho \pi R^5 \omega^2 \frac{C_p}{\lambda^3} - \frac{P_d}{\omega_i^j} \right)$$

L and H represent the wind speed set that makes the wind turbine work in the under-power state and the maximum-power state, respectively.

3 Simulation Result

Simulations are carried out using MATLAB/stateflow to show the composite switching characteristic of system (10). Considering a practical WT/BESS given in [12], where the turbine radius $R = 9$ m, Rotor inertia $J = 26000 \text{ kg} \cdot \text{m}^2$, maximum power coefficient $C_{pmax} = 0.4101$, air density $\rho = 1.29 \text{ kg/m}^3$, number of batteries $n = 30$, terminal voltage of a single battery $V_b = 12 \text{ V}$, nominal energy capacity of a single battery $C_b = 200 \text{ Ah}$, battery charging-discharging power threshold $P_{limit} = 2.5 \text{ kw}$.

In this paper, we divide the wind speed scope into 3 intervals $\sigma(v(k)) = 1, 2$ and 3 , corresponding to the average wind speed $v_{\sigma(t)}$ of 3 m/s , 7 m/s 11 m/s respectively. The transition probability matrix is shown as following

$$\begin{bmatrix} 0.759 & 0.241 & 0 \\ 0.066 & 0.786 & 0.148 \\ 0 & 0.081 & 0.919 \end{bmatrix} \quad (11)$$

Clearly, the width constant $d = 4 \text{ m/s}$, so the random signal $v_{\omega}(t) \in (-2, 2)$. The complete wind speed simulation results are shown in Fig. 4.

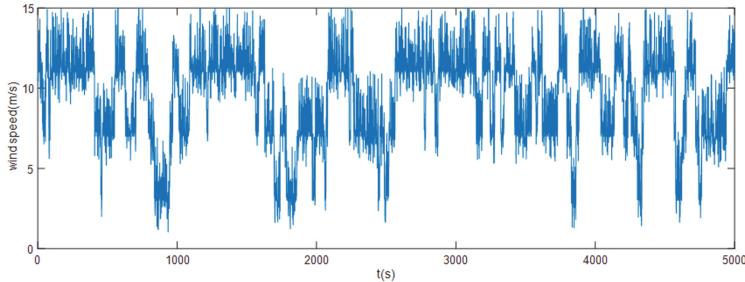


Fig. 4. Input wind speed

According to the battery power management strategy, we divide the SOC of BESS into three regions, named S1, S2 and S3 respectively, correspond to three expected output power P_1, P_2, P_3 where $P_1 = 27.5 \text{ kw}$, $P_2 = 25 \text{ kw}$, $P_3 = 22.5 \text{ kw}$. And when $SOC \leq 0.4$, the expected output power P_g^* changes to P_1 ; if SOC between 0.5 and 0.6, $P_g^* = P_2$; and if $0.7 \leq SOC$, $P_g^* = P_3$.

Thus the WT/BESS model can be built, Fig. 5 is the comparison of the total output power of WT/BESS and WT system, Fig. 6 is the curve of SOC, from which we can know that:

- 1) Although PI control can basically meet the control object under the switching of wind speed state, the output power still fluctuates according to the fluctuation of wind speed. The SOC of BESS can be maintained at an appropriate level with the switching of the expected output power
- 2) Compared with the single WT without BESS system, WT/BESS system has more stable and smooth output, which is synchronous with the composite switching.

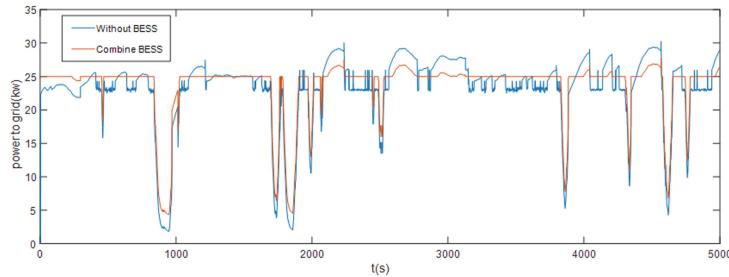


Fig. 5. The total output power supplied to grid

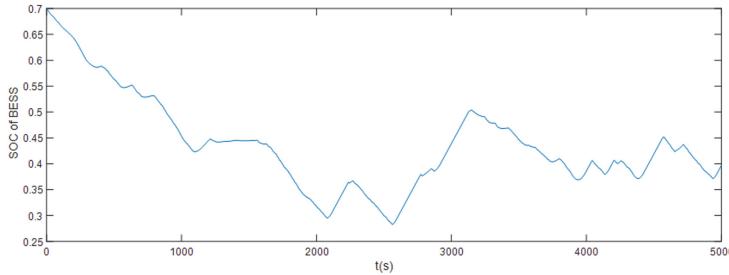


Fig. 6. The SOC of the BESS

4 Conclusion

The paper presents the switching between the operation points of the WT/BESS during different battery states and different wind speed states. A new switched system with two-level switched mechanisms is built to describe the WT/BESS. Combined with PI control, this new system can well describe the working characteristics of WT/BESS. In the end, we show the definite smoothing effect of BESS and the important effect of the SOC offset strategy to the stability of SOC, and the results are verified by simulation.

Acknowledgments. This work was supported by the National Natural Science Funds of China (61573237), 111 Project (D18003), Program of China Scholarship Council (20190-6895021).

References

1. Yin, M., Xu, Y., Shen, C., et al.: Turbine stability-constrained available wind power of variable speed wind turbines for active power control. *IEEE Trans. Power Syst.* **32**, 2487–2488 (2016). <https://doi.org/10.1109/TPWRS.2016.2605012>
2. Pao, L.Y., Johnson, K.E.: Control of wind turbines. *IEEE Control Syst. Mag.* **31**, 44–62 (2011). <https://doi.org/10.1109/MCS.2010.939962>
3. Lin, Z., Liu, J., Wu, Q., et al.: Mixed H_2/H_∞ pitch control of wind turbine with a Markovian jump model. *Int. J. Control.* **117**, 156–169 (2018). <https://doi.org/10.1080/00207179.2016.1272716>
4. Kharchouf, I., Essadki, A., Fdaili, M., et al.: Comparative study of MPPT and pitch angle using PI and fuzzy logic controllers. In: 6th International Renewable and Sustainable Energy Conference, pp. 868–873. IEEE Press, Rabat (2018). <https://doi.org/10.1109/IRSEC.2018.8702882>
5. Abolvafaei, M., Ganjefar, S.: Maximum power extraction from a wind turbine using second-order fast terminal sliding mode control. *Renew. Energy* **139**, 1437–1446 (2019). <https://doi.org/10.1016/j.renene.2019.03.044>
6. Meng, W., Yang, Q., Sun, Y.: Guaranteed performance control of DFIG variable-speed wind turbines. *IEEE Trans. Control Syst. Technol.* **24**, 1–9 (2016). <https://doi.org/10.1109/TCST.2016.2524531>
7. Roy, S.: Power output by active pitch-regulated wind turbine in presence of short duration wind variations. *IEEE Trans. Energy Convers.* **28**, 1018–1025 (2013). <https://doi.org/10.1109/TEC.2013.2282992>
8. Rahimi, M., Asadi, M.: Control and dynamic response analysis of full converter wind turbines with squirrel cage induction generators considering pitch control and drive train dynamics. *Int. J. Electr. Power Energy Syst.* **108**, 280–292 (2019). <https://doi.org/10.1016/j.ijepes.2019.01.018>
9. Howlader, A.M., Matayoshi, H., Senju, T.: A robust H_∞ controller based gain-scheduled approach for the power smoothing of wind turbine generator with a battery energy storage system. *Electr. Mach. Power Syst.* **43**, 2156–2167 (2015). <https://doi.org/10.1080/15325008.2015.1076089>
10. Zhao, T., Ding, Z.: Cooperative optimal control of battery energy storage system under wind uncertainties in a microgrid. *IEEE Trans. Power Syst.* **33**, 2292–2300 (2017). <https://doi.org/10.1109/TPWRS.2017.2741672>
11. Yilmaz, A.S., Ozer, Z.: Pitch angle control in wind turbines above the rated wind speed by multi-layer perceptron and radial basis function neural networks. *Expert Syst. Appl.* **36**, 9767–9775 (2009). <https://doi.org/10.1016/j.eswa.2009.02.014>
12. Palejiya, D., Chen, D., Mecklenborg, C.: Stability of wind turbine switching control in an integrated wind turbine and rechargeable battery system: a common quadratic Lyapunov function approach. *J. Dyn. Syst. Meas. Control*, **135**, 44–45 (2013). <https://doi.org/10.1109/ACC.2011.5991267>



Integrated Marine Environment Observation System Based on Iridium Satellite Communication

Fei Yu and Jia Sun^(✉)

School of Automation and Electrical Engineering,
Qingdao University of Science and Technology, Qingdao 266100, China
827250237@qq.com

Abstract. In this paper, iridium satellite communication based integrated Marine environment observation system is studied. There are some problems in the deep ocean, such as bad environment, difficult release of anchored buoy, single measurement of Marine environmental data and low accuracy of Marine environmental data. The system is designed to carry a variety of sensors that transmit real-time observations to a land-based monitoring center via iridium satellite communications. In this way we can get integrated Marine environmental data. The data of systematic observation is of great significance for improving the accuracy of Marine meteorological forecast, Marine ecological environment protection, further research on global Marine climate change and network observation.

Keywords: Buoy control system · Marine environment observation · Iridium satellite communication · Shore-based monitoring

1 Introduction

At present, the global climate is gradually warming. As a result, the Marine environment and Marine meteorology have become more and more complex. Tsunami, typhoon and other natural disasters on the sea have brought huge impact on the production and life of human beings [1]. Accurate observation of Marine meteorology is of far-reaching significance for disaster warning, navigation guidance of Marine ships and environmental protection in offshore areas. At present, most Marine environment observation systems used in the world adopt fixed-point buoy measurement [2,3]. The mooring buoy is fixed in the observation position and measures the Marine environmental data within a fixed range. This kind of anchor buoy has high manufacturing cost, and it is difficult to lay down in the deep sea area and inconvenient to maintain in the later period.

To solve these problems, this paper studies iridium satellite communication based Marine environment integrated observation system. In order to solve the problem that the mooring buoy is difficult to release due to the severe Marine

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 126–134, 2021.

https://doi.org/10.1007/978-981-15-8458-9_14

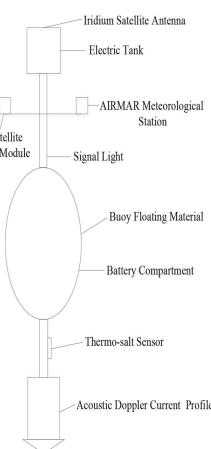
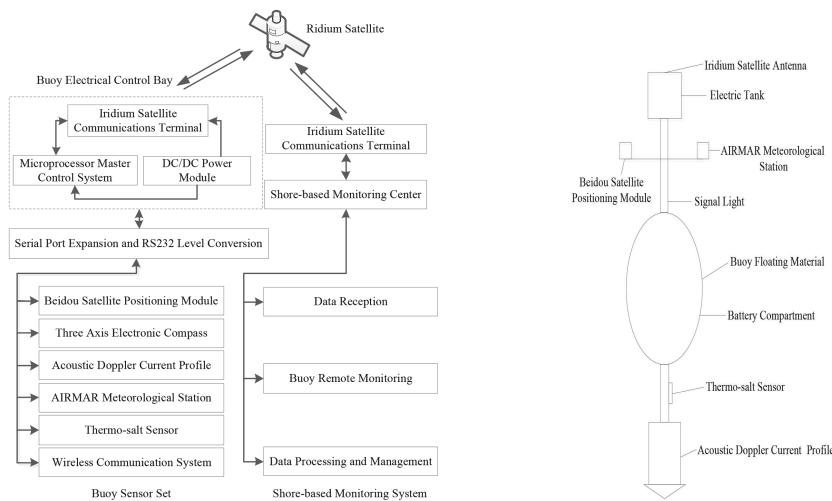
environment, this system adopts the tractive mobile observation buoy. The master control system is designed to carry a variety of sensors, and transmit the observation data to the shore-based monitoring center in real time through iridium satellite communication to obtain integrated Marine environment data [4–6].

The buoy system is towed by a mobile offshore platform and can be used for long-term offshore observation at specific target points. The design ensures the accurate observation of specific areas, and at the same time, the communication mode of global coverage improves the observation range of Marine environmental data [7–9].

2 Overall Design of System

2.1 Overall Structure Design of the Observation System

The overall structural design of the observation system includes maritime data acquisition, satellite wireless data transmission, shore-based monitoring system data reception and processing. The overall structure of the system is shown in Fig. 1.



This system takes Marine environmental data as the research object and relies on iridium satellite communication network. From data collection, compression and encryption, data remote transmission, to data reception and storage, the whole link system is designed. It provides a new solution for the collection and transmission of Marine environmental data.

The observation buoy is towed by a mobile offshore platform and can be observed for a long time at a set target point. The buoys are equipped with a variety of sensors to observe Marine environmental data, The internal master control system integrates the data and transmits the data to the iridium satellite network through satellite communications. The iridium satellite network receives data and sends it to the land-based monitoring system. The data received include air temperature, air pressure, wind speed, wind direction, water meter temperature, salinity, wave height, wave direction, flow velocity, flow direction and buoy position.

2.2 Observation System Buoy Structure Design

The mobile observation buoy mainly includes buoy floating material, Marine environment observation sensor group, electrical module, battery module, satellite communication antenna, AIS system, signal lamp, etc. The buoy structure is shown in Fig. 2.

The overall structure of the system is designed closely with high observation accuracy. It makes up for the limited measurement range of mooring buoy and the difficulty of offshore observation, and provides a technical scheme for further global netted Marine environment observation.

3 Main Function System Design

This paper mainly focuses on the Marine environment data acquisition system with high accuracy and reliability. The research content includes the design of the microprocessor master control system and the integrated data acquisition system of the Marine environment observation sensor, the satellite communication data transmission system and the shore-based monitoring system.

3.1 Microprocessor Master Control System Design and Integration of Data Acquisition System

The main control system adopts STM32F103VCT6 microprocessor of ST semiconductor company as the main control unit. The microprocessor master control system is composed of three systems, which are data acquisition system, communication system and power supply system. The minimal system composition and functions of the controller are shown in Fig. 3.

As the core of the whole system, microprocessor mainly has the following three tasks. The first task is to undertake the entire system power management. The second task is to communicate with iridium satellite module, receive shore-station instructions and send data to the shore-station. The third task is to communicate with the data acquisition module, send acquisition instructions and receive data. The communication system includes wireless communication and satellite communication. Wireless communication can be used for program

debugging, status testing and small range of maritime stereo monitoring network. Iridium satellite communications provides a remote interactive platform for ocean environment observers and shore-based monitoring, as well as data transmission and remote control for shore-based centers.

The master control system and the communication system are realized by serial communication using TTL level. The communication between the master control system and the data acquisition system is realized by RS232. The power supply system utilize 24 V secondary lithium battery pack, with the dc-dc step-down module providing the required voltage for the microprocessor and sensor.

The data acquisition system has two main tasks. The first task is to communicate with the microprocessor master control system, receive instructions and send the collected data. The second task is to collect, process and store the data collected by various sensors according to the custom protocol. The main workflow of the data acquisition system is shown in Fig. 4.

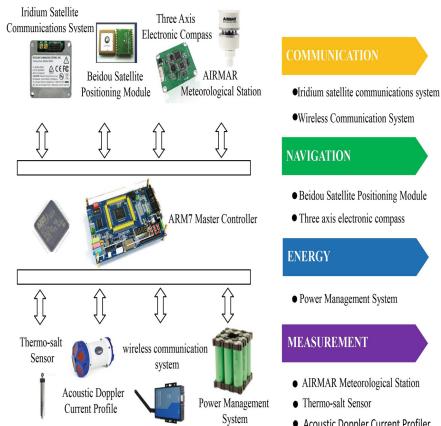


Fig. 3. Controller minimum system composition and function

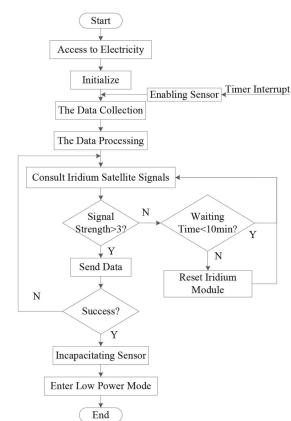


Fig. 4. Master system overall flow chart

In order to make TTL level of master control system compatible with RS232 level of data acquisition system, the system uses SP3232 chip for level conversion. The microprocessor has a limited number of serial ports. To carry more sensors, we used the WK2124 chip to extend the serial port of the microprocessor. The level conversion circuit and the serial port expansion circuit are shown in Fig. 5. Under the joint action of these systems, the ocean environment observer can operate safely and stably.

Marine environmental observers need to ensure the long-term and reliable operation of the system in the course of deep ocean observation. Integrated with a variety of sensors, the small measurement system is highly efficient and energy efficient to ensure long term Marine operation. The data acquisition and pre-processing of the master control system play an important role in improving

the measurement accuracy and data transmission efficiency of the sensor. In addition, the electronic control system can be equipped with different hydrometeorological sensors according to different requirements, such as wave sensor, nuclear radiation sensor, seawater chlorophyll sensor, etc.

3.2 Iridium Satellite Communication System Design

When the buoys are at sea, data transmission depends on satellite systems. This system uses iridium satellite's SBD9602 module as the communication module for wireless data transmission. Iridium satellite communication system can achieve global coverage of communication signals, and it has the characteristics of communication security, reliability and strong real-time performance to meet the design needs of the system.

The module interface to user connector includes DC5V power input, UART serial data interface, power supply voltage indicator. Iridium satellite module through UART interface, combined with the corresponding AT command protocol to complete the SBD business. The data writing and sending process of iridium satellite communication module is shown in Fig. 6.

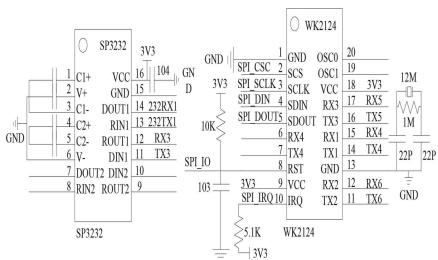


Fig. 5. Level conversion circuit and serial expansion circuit

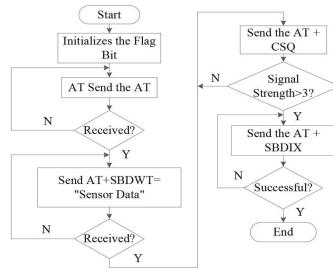


Fig. 6. Iridium satellite communication module workflow chart

3.3 Design of Shore-Based Monitoring System

Shore-based monitoring system mainly includes data receiving base station and monitoring host computer. Among them, the data receiving base station mainly receives satellite return data and uploads it to the Windows service server, performs service monitoring and achieves data acquisition. Monitoring the upper computer mainly refers to the server side that realizes human-computer interaction and data visualization. Specific functions include: terminal configuration, login user management, location tracking, track display, data audit, data storage, data export, data icon view, sending commands, monitoring and warning, etc.

The upper computer interface is shown in Fig. 7. The system workflow is shown in Fig. 8.

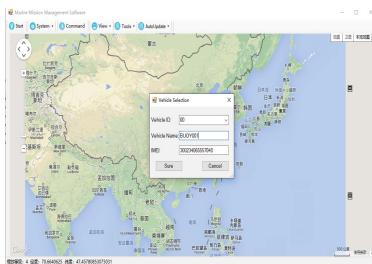


Fig. 7. Shore-based monitoring host interface

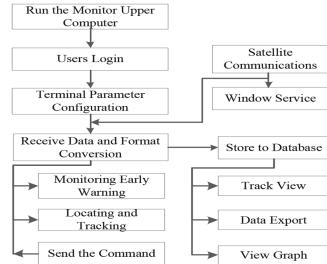


Fig. 8. Flow chart of shore-based monitoring upper computer

4 System Application Testing and Analysis

4.1 Sea Experiment

In March 2019, the system carried out a seven-day maritime monitoring performance test in Qingdao, China. This test mainly tests the design rationality of the monitoring system, the long-term communication ability, the data acquisition stability and the track tracking performance. The ocean test diagram is shown in Fig. 9. The sea track diagram of the system during the sea test is shown in Fig. 10.



Fig. 9. Actual operation test diagram of the buoy system at sea

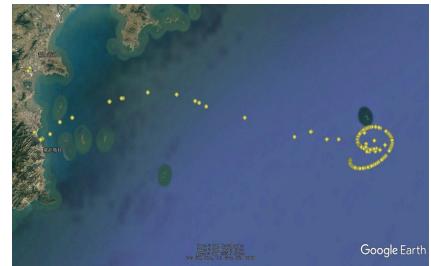


Fig. 10. Buoy sea trial run track chart

In the sea trial, under the influence of the complicated environment such as typhoon and sea wave, the data acquisition system still accurately recorded the real-time changing data of the ocean environment, and carried out a seven-day sea trial at sea. During the sea trial, the device sent back a set of data every 10 min. In the 7-day theory, the data returned was 1008 groups. In the actual test, due to the unstable signals in some offshore areas, 995 groups of data were actually received, and the data reception rate was 98.7%, which met the design requirements. The equipment is running stably in the sea trial, and the returned data will be further analyzed in the next step.

4.2 Return Data Analysis

In the shore-based monitoring center designed by the system, the preliminary analysis of the returned data can be carried out. After the login interface configuration is completed, enter the visual monitoring interface, and the data returned from the export section is shown in Fig. 11.

ID	COMMUNICATION ID	DATE	TIME	VOLTAGE	LATITUDE	LONGITUDE	PITCH	ROLL	ORIENTATION	TEMP	AIR PRESSURE	WING SPEED	WIND DIRECTION	WATER TEMP	SALINITY
734	300234065557040	2019/9/23	10:59:40	33.6585	36.215463	121.4095	-3.50	-13.72	257.29	6.2692	101.0185	7.95017	201.8893	5.02695	31.71
735	300234065557040	2019/9/23	11:09:47	33.658	36.21347	121.3901	0.49	-22.01	247.02	6.3625	101.0259	7.60398	198.4246	5.13835	31.71
736	300234065557040	2019/9/23	11:19:49	33.658	36.21347	121.3901	-30	1745	236.47	6.59858	101.05236	8.827219	204.1551	5.02356	31.72
737	300234065557040	2019/9/23	11:32:04	33.658	36.21391	121.4038	2.44	-5.68	243.75	6.5892	101.17584	8.202729	205.6514	5.12356	31.73
738	300234065557040	2019/9/23	11:40:08	33.658	36.21391	121.4038	3.92	6.96	291.74	6.98563	101.1259	8.100893	205.4336	5.03398	31.73
739	300234065557040	2019/9/23	11:48:04	33.658	36.21508	121.4103	-19.82	-8.62	207.03	6.95026	101.2039	8.898551	205.8647	5.04569	31.73
740	300234065557040	2019/9/23	12:07:55	33.658	36.21602	121.4139	1.5	27.11	195.94	7.01256	101.3595	8.230905	201.5916	5.16965	31.72
741	300234065557040	2019/9/23	12:27:55	33.657	36.21602	121.4139	-1.5	1.43	94.15	7.01256	101.3595	8.01092	202.0501	5.13695	31.72
742	300234065557040	2019/9/23	12:27:55	33.657	36.21602	121.4139	-18.47	-14.43	171.29	7.12695	101.2039	8.397281	197.756	5.16985	31.74
743	300234065557040	2019/9/23	12:38:13	33.657	36.21602	121.4139	19.17	-5.21	197.37	7.12899	101.2351	7.643832	200.2599	5.07205	31.72
744	300234065557040	2019/9/23	12:49:10	33.657	36.21602	121.4139	9.32	-6.55	99.37	7.25698	101.4569	8.775824	201.5372	5.07235	31.74
745	300234065557040	2019/9/23	12:57:48	33.657	36.21602	121.4139	11.17	-3.61	109.55	7.12547	101.1441	8.774817	202.2512	5.07569	31.73
746	300234065557040	2019/9/23	13:07:48	33.657	36.22377	121.4248	-10.52	-17.55	111.97	7.13698	101.14785	9.255445	205.5664	5.17458	31.74
747	300234065557040	2019/9/23	13:17:48	33.657	36.22377	121.4248	37.53	0.68	81.43	7.156935	101.15236	7.023695	214.841	5.18569	31.74
748	300234065557040	2019/9/23	13:27:47	33.657	36.22377	121.4248	-1.86	-1.83	94.15	7.15698	101.15236	8.01092	202.3377	5.13695	31.76
749	300234065557040	2019/9/23	13:37:47	33.657	36.22377	121.4248	14.72	3.77	90.01	7.35698	101.15598	8.230347	208.3277	5.18147	31.75
750	300234065557040	2019/9/23	13:48:07	33.658	36.22377	121.4249	3.38	-1.43	71.46	7.26589	101.66339	7.829078	202.8229	5.12569	31.77
751	300234065557040	2019/9/23	13:58:23	33.658	36.22377	121.4249	15.37	-7.96	71.72	7.26359	101.02036	8.194903	201.7916	5.18569	31.74
752	300234065557040	2019/9/23	14:08:18	33.658	36.22377	121.4249	-10.97	0.4	68.81	7.45689	101.15896	8.363647	204.4083	5.22365	31.72
753	300234065557040	2019/9/23	14:17:58	33.658	36.22377	121.4249	6.5	-2.54	62.5	7.56369	101.14782	7.929078	201.7916	5.23695	31.73
754	300234065557040	2019/9/23	14:27:48	33.658	36.22377	121.4249	-1.86	-26.11	56.04	7.24589	101.10302	8.184934	201.8654	5.26986	31.72
755	300234065557040	2019/9/23	14:37:47	33.658	36.22377	121.4249	-1.86	-2.89	55.99	7.24589	101.10302	10.01488	210.0505	5.30695	31.75
756	300234065557040	2019/9/23	14:48:01	33.658	36.22377	121.4249	-0.72	-7.05	35.49	7.28650	101.44115	8.085106	202.2339	5.25698	31.74
757	300234065557040	2019/9/23	14:47:48	33.658	36.22369	121.4018	11.17	12.04	42.58	7.35685	101.74736	8.636339	204.5283	5.32659	31.75
758	300234065557040	2019/9/23	14:57:40	33.658	36.23069	121.4018	5.93	19.23	21.92	7.36596	101.20233	8.388352	204.0395	5.36258	31.76
759	300234065557040	2019/9/23	15:07:48	33.658	36.23069	121.4018	2.18	-3.29	354.23	7.48503	101.21246	8.156335	204.3652	5.36595	31.77
760	300234065557040	2019/9/23	15:17:46	33.658	36.23069	121.4018	3.41	-8.11	328.71	7.55236	101.2569	9.350391	205.2265	5.24856	31.75
761	300234065557040	2019/9/23	15:27:42	33.658	36.23069	121.4018	11.1	-26.35	313.62	7.6965	101.2669	7.892726	206.9532	5.36985	31.74
762	300234065557040	2019/9/23	15:39:04	33.658	36.22553	121.3839	17.52	7.75	329.43	7.62568	101.15893	8.328182	204.3625	5.24968	31.76
763	300234065557040	2019/9/23	15:47:51	33.658	36.22553	121.3839	1.83	2.83	267.75	7.5232	101.1456	8.346445	206.2145	5.45265	31.72

Fig. 11. Shore-based monitoring upper computer database to export data

In order to further analyze the returned data of the system, we used MATLAB software to further process the data derived from the upper computer monitoring system. The attitude diagram of the buoy system during the sea trial, the seawater velocity diagram of the acoustic doppler current profiler, the meteorological information diagram of AIRMAR, and the data diagram of the temperature and salt sensor are obtained, such as the chart.

Figure 12 draws the attitude information diagram of the buoy system during sea test, including the buoy's pitch Angle, roll Angle and azimuth Angle. Buoys wobble under the influence of waves during sea trials. The acoustic doppler current profiler installed on the buoy needs to be in a relatively stable attitude to obtain more accurate velocity measurement information. In Fig. 12, we can intuitively see that the roll Angle and pitch Angle of the system are basically at $\pm 30^\circ$ during the sea test, and the buoy posture is generally stable during the sea test.

The acoustic doppler current profiler can collect information about the velocity of the sea water at a depth of 140 m. For this sea trial, we collected information about the velocity of seawater at a depth of 30 m. It was divided into 30 layers, each of which was measured at a depth of 1 m. The velocity of seawater in the first layer is shown in Fig. 13, which are vertical velocity, due north velocity and due east velocity respectively.

The AIRMAR meteorological station equipped with the system can measure air temperature, air pressure, wind speed, wind direction and other data at the sea-air interface. The meteorological data measured during the sea trial are

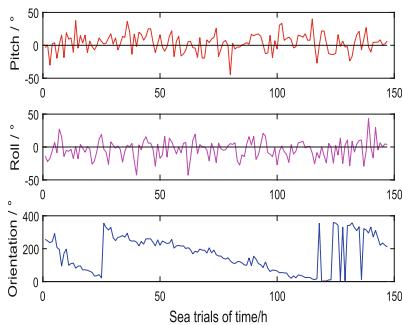


Fig. 12. Buoy system attitude diagram

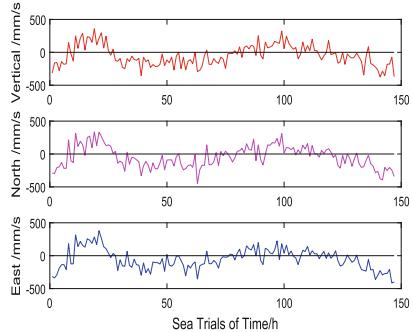


Fig. 13. Acoustic doppler current profiler data

shown in Fig. 14. As can be seen from the figure, the sea wind speed is relatively stable and the meteorological environment is better during the sea trial.

The temperature and salinity of sea water in the qianliyan sea area measured by the temperature and salinity sensor equipped with the system are shown in Fig. 15. The figure visually shows the temperature and salinity change curve of the sea water during the sea trial.

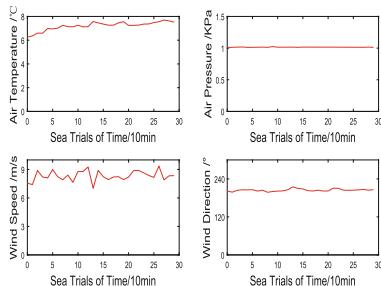


Fig. 14. AIRMAR weather station information

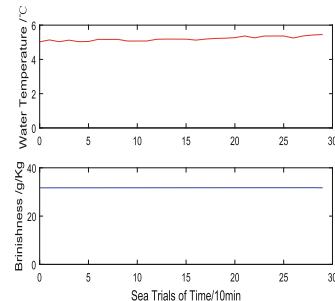


Fig. 15. Thermohaline sensor data

5 Conclusion

The Marine environment data collection and return system is studied in this paper. The main objective of this paper is to accurately monitor various Marine environmental parameters such as current velocity, flow direction, air temperature, air pressure, wind speed, wind direction, seawater salinity, temperature and so on. The data are transmitted in real time to ground monitoring stations for processing by ground meteorological technicians, so as to achieve the purpose of Marine meteorological prediction and disaster warning. In this paper, by

using mobile observation buoy, a variety of sensors are designed Through iridium satellite communication, the observation data are sent to the shore-based monitoring center in real time to obtain integrated Marine environment data and meet the requirements of accurate observation in specific areas. Finally, the sea experiment verifies the feasibility, effectiveness and practicability of the monitoring system. The data of systematic observation is of great significance for improving the accuracy of Marine meteorological forecast, Marine ecological environment protection, further research on global Marine climate change and network observation.

Acknowledgement. This work was supported by Pilot National Laboratory for Marine Science and Technology (Qingdao) (2017WHZZB0101) and National Natural Science Foundation of China (61803219).

References

1. Stanitski, D., Johnson, M., Thurston, S.: Building a global ocean observing system for climate. In: Proceedings of OCEANS 2005 MTS/IEEE, Washington, DC, pp. 1894–1898 (2005). <https://doi.org/10.1109/oceans.2005.1640035>
2. Ichikawa, M., Hirate, K.: Autonomous fixed point mooring oceanographic observation buoy system (OOS). In: 2017 IEEE Underwater Technology, Busan, pp. 1–3 (2017). <https://doi.org/10.1109/ut.2017.7890322>
3. Li, X., Zou, D., Feng, W., Xie, W., Shi, L.: Study of quality control methods for moored buoys observation data. In: 2019 International Conference on Meteorology Observations, Chengdu, pp. 1–4 (2019). <https://doi.org/10.1109/ICMO49322.2019.9026114>
4. Liu, W., Li, G., Gao, L., Li, L., Li, Z.: Monitoring and communication system design for a deep-sea unmanned submersible. In: 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans, Kobe, pp. 1–5 (2018). <https://doi.org/10.1109/oceanskobe.2018.8559173>
5. Jiang, Y., Li, J., Guo, Z.: Design and implementation of a prototype system of ocean sensor web. In: IET International Conference on Wireless Sensor Network 2010, Beijing, pp. 21–26 (2010). <https://doi.org/10.1049/cp.2010.1021>
6. Heitesenrether, R., Fiorentino, L., Breuer, E., Krug, W., Teng, C.-C.: Subsurface ADCP buoy for near surface current observations in 200–300 m coastal waters. In: OCEANS 2018 MTS/IEEE Charleston, Charleston, pp. 1–8 (2018). <https://doi.org/10.1109/oceans.2018.8604668>
7. Laun, L.A., Pittman, E.E.: Development of a small, low-cost, networked buoy for persistent ocean monitoring and data acquisition. In: OCEANS 2018 MTS/IEEE, Charleston, pp. 1–6 (2018). <https://doi.org/10.1109/oceans.2018.8604510>
8. Cui, X.: Attitude integrate monitor system design for direction finding of vector hydrophone on buoy platform. In: 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference, Xi'an, pp. 2177–2180 (2018). <https://doi.org/10.1109/IMCEC.2018.8469324>
9. Earle, M.D., Brown, M.R., Shih, H.H., Sprenke, J.J., Collier, W., Crump, D.R.: GPS-tracked buoy for hydrographic survey applications. In: Proceedings of OCEANS 2005 MTS/IEEE, Washington, DC, pp. 1263–1267 (2005). <https://doi.org/10.1109/oceans.2005.1639928>



Intelligent Recognition of Bladder Cancer Based on Convolutional Neural Network

Xinguang Jin¹(✉) and Qi Zhang^{1,2}

¹ School of Information Technology and Management,
University of International Business and Economics, Beijing 100029, China
409683945@qq.com

² Key Laboratory of Machine Perception (Ministry of Education),
Peking University, Beijing 100871, China

Abstract. Bladder cancer is the most common malignant tumor in urology. However, due to the diverse morphology of bladder tumor, including villous, follicular, flat erythema, etc., it is sometimes difficult for doctors to distinguish by eye from a variety of benign inflammatory diseases. This may lead to misdiagnosis or missed diagnosis. In this paper, we present a convolutional neural network based deep learning model for tumor recognition that efficiently realizes the intelligent recognition of cystoscope images and achieves a good recognition accuracy in experiments.

Keywords: Bladder cancer · Cystoscopy · CNN · BP neural network

1 Introduction

In China, bladder cancer is the most common malignant tumor in urinary system. The number of patients diagnosed with bladder cancer is more than 60,000 every year, and its incidence has surpassed that of kidney cancer and prostate. Although making timely treatment upon the bladder cancer can get almost 90% cure rate for patients, there are still about 5% of the patients encountering the misdiagnosis or missed diagnosis issues. At present, cystoscopy combined with histopathology is still the gold standard for the diagnosis of bladder cancer, but some of the bladder tumors have similar morphology with inflammatory lesions. This may lead to misdiagnose of identifying tumor lesions as inflammatory ones, which thus delays the treatment of patients. In view of the practical problems encountered in the diagnostic process of bladder cancer, it is desirable to implement intelligent recognition and diagnosis of the tumors in cystoscope images by artificial intelligence technology, so as to assist the cystoscopy operators in the diagnosis of benign and malignant tumors.

Convolutional neural network (CNN) has gained a great popularity in image recognition field. Since AlexNet in 2012 [1], the complex CNN models supported by GPU computing clusters have been the winning algorithm of ILSVRC for many times and applied in various recognition tasks. O. Eminaga et al. (2018)

developed deep convolutional neural network models and evaluated those using F1 scores, they finally found that the Xception-based model got the highest score [2]. A. Ikeda et al. (2020) aimed to support the cystoscopic diagnosis of bladder cancer using artificial intelligence based on CNN and realized the possibility to classify the image, including tumor lesions and normality [3]. Their research shows the potential of deep learning for the diagnostic classification of cystoscope images. In this paper, we establish a deep learning recognition model for bladder tumors based on CNN, and fulfill the intelligent recognition by using the Keras deep learning framework. The recognition accuracy of the proposed tumor recognition model is 91.67%, which indicates that our model can effectively distinguish the bladder malignant tumors from inflammatory lesions.

2 Materials and Network Architecture

2.1 Dataset Pretreatment

The cystoscope image data in this research comes from the department of urology, Peking University third hospital. The following figures are from the original cystoscopy data set. It shows that the bladder tumor has a variety of morphologies, such as villous, follicular, and flat erythema, often similar to a variety of inflammatory lesions. Malignant lesions and benign ones are difficult to be distinguished by the naked eyes, which leads to misdiagnosis or missed diagnosis. Therefore, we expect to establish a machine learning model to learn the intrinsic features in images of the malignant tumors and benign ones (Fig. 1).



Fig. 1. Dataset sample

The first step is to mark the original image data set. In this paper, we study the bladder cancer and non-cancer of binary classification problem, therefore, the cystoscopy images are labeled as malignant tumors and benign lesions with 0–1 labels, where the cancer images set as 1 and the non-cancer as 0. The second step of data preprocessing is pixel standardization. Due to the pixel value of the original image ranges from 0 to 255, and the image belongs to digital image of RGB three-channel mode, which is a high-dimensional tensor in space. If such tensor is directly input into a learning model, it will bring a huge computation overhead. Therefore, it is necessary to standardize the pixel value of the original image. Here, the pixel values are mapped to the interval of 0–1 by the normalization formula as follows.

$$x' = \frac{x - X_{min}}{X_{max} - X_{min}},$$

where x is the pixel value of image, X_{min} is the smallest pixel value 0 and X_{max} is the maximum 255.

2.2 Construction of Deep Convolutional Neural Network

CNN is a multilayer neural network structure, each layer is composed of several two-dimensional planes as shown in the following figure. In a CNN recognition model, the first layer are the input images, and the second layer is a number of feature maps obtained by convolution operation. The following is the down sampling layer, which is to reduce the resolution of the feature map, so as to retain valid feature information while reducing the amount of processing parameters. Until entering the dense layer. The feature map tensors obtained after several convolution and pooling are flattened into a vector as the input of BP neural network. BP neural network as the classifier, trains all the features of input to get the final classification output (Fig. 2).

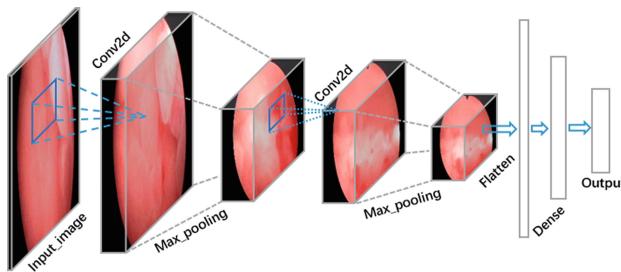


Fig. 2. CNN network structure

3 Methods and Implementations

3.1 Feature Extraction Based on Convolution Operation

The convolutional layer is a unique structure in CNN. Convolution operation enhances the features of input original signal while reducing noise. The convolution kernel is applied to such digital image, which is actually a local weight sharing structure and shown as a weight matrix in space. Different features of the input image can be extracted by setting different kernels. Each kernel convolved with the input image can obtain a feature map, which with specific features of the input. In a traditional fully-connected neural network, all the neurons in the input layer are connected to each neuron in adjacent hidden layer. While in CNN, the first hidden layer is only connected to neurons in a local area of the input. Convolution operation involves matrix dot product and sliding window. The dot product is shown as the following figure, each element in the $3 * 3$ region of the $6 * 6$ input pixel matrix is multiplied by its corresponding weight matrix and finally add them up. The result of the dot product constitutes the elements of the output matrix. Sliding window is carried out from left to right and top to down upon the input image to generate feature map after a traversal (Fig. 3).

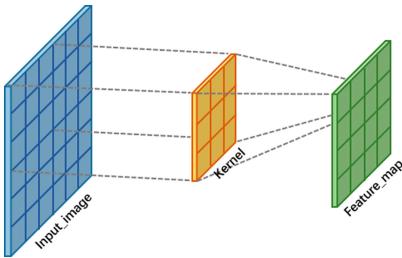


Fig. 3. Convolution operation

Because of this local connection and weight sharing mechanism, the convolution operation can effectively reduce the amount of parameter, while it is hard to do in a fully-connected network. Features in different directions can be extracted by setting kinds of kernels, such as vertical or horizontal and so on. While if using a kernel of Laplace operator, the feature map shows the irregular feature information. For areas of the image without edges, such as the black shaded areas in the cystoscope images, most pixel values are almost the same, so the kernel outputs the same values. While for an image with edges, since the pixel values on both sides of the edge are different, the result of convolution is also different, thus the edge features can be detected. For RGB three-channel digital image, convolution operation is performed on a three parallel channels in space, and then overlay the values of three matrixs at the same position to synthesize the final feature map.

3.2 Down Sampling and Pooling Operation

Down sampling is actually a process of dimensionality reduction while preserving the original image feature. The summary statistical features obtained by pooling not only have lower dimension, but also improve the training effect of model, which can effectively reduce overfitting. In many CNN models, pooling is a common method of down sampling. Similar to the process of convolution's sliding window, pooling slides on the input and do an aggregation on pixel matrix in a specific area. Max pooling and mean pooling are two methods for this aggregation. The former takes the maximum of all pixel values in pooling area, the features obtained by this method are more sensitive to texture. The latter is to do an average and get features are more sensitive to background information of image. Because of the abundant texture feature of bladder tumor tissue in cystoscope image, this experiment adopts the max pooling. The figure below shows the max pooling operation, it can be seen that after a max pooling, the size of the image changed to original 1/4. Of course, pooling is not the only down sampling method, the convolution with stride larger than 1 can make it as well (Fig. 4).

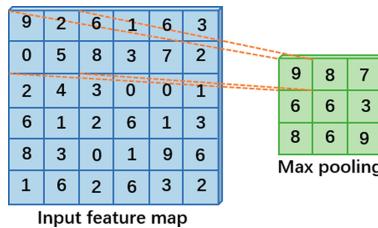


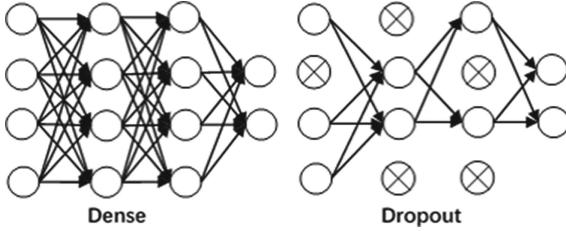
Fig. 4. Max pooling operation

3.3 BP Neural Network Classifier and Dropout

After several convolution and pooling operations, the final step is to build a classifier. In this experiment, we use the BP neural network as the classifier model. Before entering the dense layer, it is necessary to flatten all the feature map tensors into a vector. This vector is used as input node of the neural network. Several hidden layers are setting after the input layer, and using ReLU activation function for each neuron in hidden layer. The output layer is two neuron nodes using sigmoid as the activation, which can solve the binary classification problem effectively. In order to improve the training speed and accelerate the convergence of model, this experiment carrys the enhanced sigmoid function, namely the hard sigmoid and its formula as follows.

$$\text{Hard_Sigmoid}(x) = \begin{cases} 0 & x < -2.5 \\ 0.2x + 0.5 & -2.5 \leq x \leq 2.5 \\ 1 & x > 2.5 \end{cases}$$

In CNN model, the down sampling layer realizes the proportional dimension reduction of feature map and reduces overfitting to some extent. Another technique that prevents model from overfitting in dense layer is dropout. Dropout enables that the neuron nodes in hidden layers are randomly abandoned with a certain probability, and sets the output of these neurons to 0. Thus certain proportion of neurons cannot participate in feedforward and back-propagation training process. Due to this mechanism, a batch of data is imported into the model each time can actually train different networks. Which means that each sample input is equivalent to trying a new network structure, while all these different networks are sharing the same weight. Dropout reduces the complex mutual adaptation between neurons and thus reduces overfitting by generating different network structures for combination. After each dropout, it is equivalent to find a network structure with a certain degree of sparsity from the original dense network, as shown in the following figure. It should be noted that the total number of model training parameters has not changed, yet can effectively reduce the time-consuming problem of deep learning model training (Fig. 5).

**Fig. 5.** Dense and dropout

3.4 Optimizer and Loss Construction

After the construction of this CNN model, we start to do training. Training model involves two key issues. One is the selection of optimizer, another is the construction of loss function. The choice of the optimizer determines the speed and effect of model training. We select Adam optimizer (i.e., an algorithm for first-order gradient-based optimization of stochastic objective functions) based on adaptive estimates of lower-order moments. The method is straightforward to implement and computationally efficient, which requires little memory consumption. Is invariant to diagonal rescaling of the gradients, and is well suited for problems that are large in terms of data or parameters. The final prediction result of the model is binary, i.e., cancer and non-cancer. Therefore, the loss function is constructed based on the binary cross entropy. The binary cross entropy can be regarded as the special case of cross entropy, which defined as

$$\text{Cross_Entropy}(p, q) = - \sum_{i=1}^n p(x_i) \log q(x_i).$$

There are only two categories here including x_1 and x_2 , then

$$\text{Cross_Entropy}(p, q) = -(p(x_1) \log q(x_1) + p(x_2) \log q(x_2)),$$

because there are only two choices, then $p(x_1) + p(x_2) = 1$ and $q(x_1) + q(x_2) = 1$. Suppose that the probability of x_1 in the training sample is p , then the probability of x_2 is $1 - p$; the predicted probability of x_1 is q , then the predicted probability of x_2 is $1 - q$. So the above equation can be rewritten as follows.

$$\text{Cross_Entropy}(p, q) = -(p \log q + (1 - p) \log(1 - q)).$$

The loss function L formulated below is the target function that needs to be optimized in the model training process,

$$L = -\frac{1}{n} \sum_{i=1}^n (y_i^* \log H_i + (1 - y_i^*) \log(1 - H_i)),$$

where the probability predicted for each category is H and $1 - H$, y^* is the label for each sample.

3.5 Experiment and Evaluation

This experiment is implemented on the Keras deep learning framework. After a certain number of iterative training, the constructed deep learning model finally got the convergence of loss function. In the process of model iterative training, the accuracy and convergence of loss on the training set and validation set are shown in the following figure (Fig. 6). The above two curves are the training results of 20 iterations and the below curves show the variation tendency of 25 iterations. It can be seen that the model shows some fluctuations in the validation set, and the performance of the model tends to be better with the moderate increase of iterations. The accuracy of the recognition model in the final test set reached 91.67%, which indicates that a good prediction performance can be achieved.

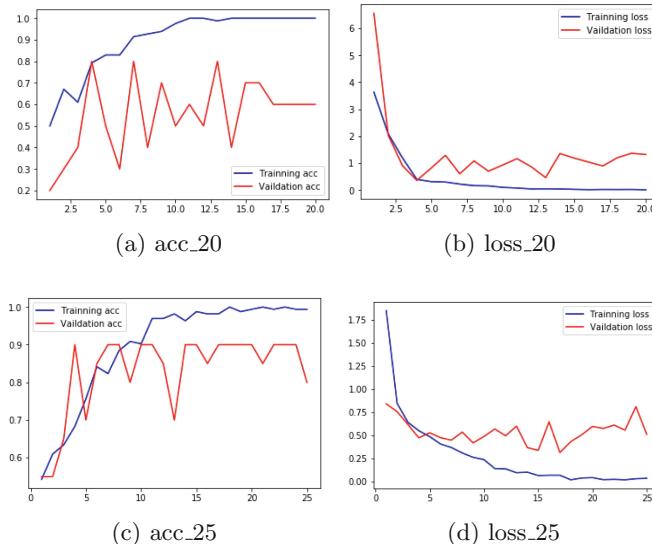


Fig. 6. Training accuracy and loss

4 Conclusion

The conclusion of this study is expected to provide a powerful auxiliary reference for clinicians in the diagnosis of benign and malignant tumor lesions based on cystoscope images. Cystoscope image can be pre-diagnosed and recognized by the proposed deep learning based recognition model, which provides model support for the recognition of tumors from cystoscope. This is of much help for diagnosing the urological malignancy, with experimental results verifying the feasibility and effectiveness of the machine learning based intelligent diagnose. In order to further improve the prediction accuracy and generalization ability of

the model, the subsequent directions are as following. First, noise processing and image detection segmentation could be carried out during image preprocessing. Second, the structure of the network model could be adjusted appropriately, and the hyperparameters of the network model could be further optimized.

Acknowledgement. This paper is funded by Beijing Natural Science Foundation (19L2037) and the Excellent Young Scholars Funding Project in UIBE (No. 19YQ10).

References

1. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: NIPS, vol. 25. Curran Associates Inc. (2012)
2. Eminaga, O., Semjonow, A., Breil, B.: Diagnostic classification of cystoscopic images using deep convolutional neural networks. *Eur. Urol. Suppl.* **17**(2), e1232 (2018)
3. Ikeda, A., Nosato, H., Kochi, Y., Kojima, T., Nishiyama, H.: Support system of cystoscopic diagnosis for bladder cancer based on artificial intelligence. *J. Endourol.* **34**, 352–358 (2019)
4. Moeskops, P., Viergever, M.A., Mendrik, A.M., Vries, L.S.D., Benders, M.J.N.L., Işgum, I.: Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Trans. Med. Imaging* **35**(5), 1252–1261 (2017)
5. Shin, H.C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., et al.: Deep convolutional neural networks for computer-aided detection: CNN, architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **35**(5), 1285–1298 (2016)
6. Zhou, D.X.: Theory of deep convolutional neural networks: downsampling. *Neural Netw.* **124**, 319–327 (2020)
7. Garbin, C., Zhu, X., Marques, O.: Dropout vs. batch normalization: an empirical study of their impact to deep learning. *Multimed. Tools Appl.* **79**(19), 12777–12815 (2020)
8. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. *Computer Science* (2014)
9. Zhang, Z., Sabuncu, M.R.: Generalized cross entropy loss for training deep neural networks with noisy labels (2018)



Distributed Consensus Backstepping Control in Networked Flexible Joint Manipulator Systems

Chengye Ying¹, Lixia Liu¹, Zhonghua Miao^{2(✉)}, and Jin Zhou¹

¹ Shanghai Institute of Applied Mathematics and Mechanics, and Shanghai Key Laboratory of Mechanics in Energy Engineering, School of Mechanics and Engineering Science, Shanghai University, Shanghai, People's Republic of China

² School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, People's Republic of China

zhhmiao@shu.edu.cn

Abstract. In this brief, the problem of distributed consensus backstepping control in networked flexible-joint robotic systems is investigated. A unified distributed consensus tracking strategy is proposed for the cases of the directed network topology with a spanning tree by systematically combining backstepping approach and sliding mode technique. It is shown that the developed cooperative algorithms can always guarantee the desired tracking consensus of networked nonidentical flexible-joint manipulator systems, and it also has the robustness to some unmatched uncertainties. Finally, the simulation illustrative examples are provided to illustrate the proposed control algorithms.

Keywords: Distributed cooperative consensus control · Networked multi-agent systems · Flexible joint manipulator · Backstepping control

1 Introduction

There has been an increasing interest together with a significant advance in distributed cooperative control in large-scale networks of multi-agent systems in recent years [1, 2]. In particular, special attention has been focused on networked multiple robotic systems formulated by Euler-Lagrange (EL) dynamics due to its potential practical applications especially in complex unstructured task environments, such as underwater and space exploration, hazardous environments, and service robotics [3]. As a result, this has led to a large variety of synchronization and consensus protocol (or algorithm) for fully actuated networked EL systems from various perspectives.

As is generally known, robotic manipulators with joint flexibilities have some potential advantages over their single rigid manipulators including more flexibility, high security and low energy consumption, such as surgical and space robots [4]. Accordingly, the modeling and control of flexible-joint robots has become an interesting challenge topic over the past decades along with the development of several fundamental control schemes such as the feedback linearization method, the singular perturbation approach, the integral manifold control, and the adaptive sliding mode technique, etc. [5–7]. As

the aforementioned schemes, the conventional backstepping control is viewed as a systematic methodology following a step by step based on the construction of feedback control law and Lyapunov functions, in some case, so it is effective and relatively easily implemented for solving uncertain nonlinear systems which do not satisfy the matching conditions. For example, Jong et al. put forward an adaptive backstepping design approach to the control of flexible joint robot manipulators [8].

It is widely accepted that, the distributed consensus control of networked flexible-joint robotic systems modelled by under-actuated EL dynamics is more challenging than that for their single-robot counterpart because of the inherent highly nonlinearity and stronger coupling involving multiple dynamics [3]. More importantly, the design of cooperative consensus control scheme becomes more difficult when the communication interactions among robots are taken into account simultaneously especially for regulating a large group of robots [9, 10]. As a consequence, up to now, very little research work was reported on the topic, and thus these observations motivate to conduct our present research in this brief.

The main objective of this brief is to further extend the trajectory tracking control of the single flexible-joint robot to the case of distributed consensus control of networked flexible-joint robotic systems via the backstepping schemes. The main contribution of the paper is two-fold compared with the latest reported references, namely: (i) A fully distributed consensus control algorithm is developed for directed network topology graph with a spanning tree by a systematic combination of both backstepping approach and sliding mode technique, and so it would be easily realized for controlling a large group of flexible-joint robots. (ii) The proposed backstepping consensus algorithm is mainly based on routine Lyapunov-like analysis and careful gain selections, and thus it can fully guarantee globally asymptotic stability of the closed-loop controlled system, which also yields a unified additional robust consensus controller for uncertain networked flexible-joint robotic systems.

The rest of the paper is organized as follows: Some preliminaries and problem formulation are presented in Sect. 2. Section 3 gives the main results of distributed consensus control of networked flexible-joint robotic systems. Section 4 provides the simulation illustrative examples, and the concluding remarks are made in Sect. 5.

2 Preliminaries

2.1 Graph Theory

Consider a networked agent system consisting of n nodes, which denoted by a weighted directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ including node set $\mathcal{V} = \{1, \dots, n\}$, the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ and adjacency matrix $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{n \times n}$. For any $i \neq j$, $i, j \in \mathcal{V}$, the adjacency element $a_{ij} \neq 0$, $(i, j) \in \mathcal{E}$. When $i = j$, $a_{ij} = 0$. A path from node i_1 to node i_p is a sequence of edges $(i_1, i_2), (i_2, i_3), \dots, (i_{p-1}, i_p)$. \mathcal{G} is called a connected graph if any $i, j \in \mathcal{V}$. A directed graph has a spanning tree with at least one node i , which indicates that there is a path from i to j for any other node j , $j \neq i$. As usual, the Laplacian matrix $\mathcal{L} \triangleq [l_{ij}] \in \mathbb{R}^{n \times n}$ with respect to \mathcal{A} is defined as $l_{ii} = \sum_{j=1, i \neq j}^n a_{ij}$ and $l_{ij} = -a_{ij}$, $i \neq j$.

2.2 Networked Flexible-Joint Robotic Systems

Consider a networked agent system of n flexible joint manipulators, where the dynamics model of the i -th flexible joint manipulator can be represented by the classic Euler-Lagrangian (EL) dynamics [4]

$$D_i(q_{1i})\ddot{q}_{1i} + C_i(q_{1i}, \dot{q}_{1i})\dot{q}_{1i} + G_i(q_{1i}) + K_i(q_{1i} - q_{2i}) = 0, \quad (1)$$

$$J_i\ddot{q}_{2i} + K_i(q_{2i} - q_{1i}) = u_i, \quad i = 1, 2, \dots, n, \quad (2)$$

where $q_{1i} \in \mathbb{R}^p$ is the link angle of flexible robotic arm i , $q_{2i} \in \mathbb{R}^p$ represents the motor angles. $D_i(q_{1i}) \in \mathbb{R}^{p \times p}$ is the link inertia matrix, $C_i(q_{1i}, \dot{q}_{1i}) \in \mathbb{R}^{p \times p}$ is the Coriolis and centrifugal matrix, $G_i(q_{1i}) \in \mathbb{R}^p$ represents the gravitational terms, coefficient K_i is the elastic stiffness of the flexible robotic arm, and J_i represents the inertia matrix of the motor, $u_i \in \mathbb{R}^p$ is the input torque.

In general, it is assumed that the dynamics model (1) of flexible joint manipulator have the following fundamental properties:

P1: There exist positive constants β_{Di} , β_{ni} , β_{Ci} , β_{gi} such that

$$\|D_i(q_{1i})\| \leq \beta_{Di}, \|D_i(q_{1i})^{-1}\| \leq \beta_{ni}, \|C_i(q_{1i}, \dot{q}_{1i})\| \leq \beta_{Ci}, \text{ and } \|G_i(q_{1i})\| \leq \beta_{gi}.$$

P2: The matrix $\dot{D}_i(q_{1i}) - 2C_i(q_{1i}, \dot{q}_{1i})$ is skew symmetric.

It is obvious that the flexible joint manipulator system is composed of the cascade of two parts with robot link dynamics (1) that are actuated by the rotation angle q_{2i} and motor dynamics (2) that are driven by the motor input torque u_i . According to the cascade control method, a virtual control input is first developed for the link dynamics, then the motor dynamic torque control inputs in realizing the consensus tracking performance are presented subsequently.

For the sequel use, we first define two auxiliary variables $q_{2i} = x_{1i}$, $\dot{q}_{2i} = x_{2i}$, then the dynamic model (1) and (2) can be divide into following three parts:

$$\begin{cases} D_i(q_{1i})\ddot{q}_{1i} + C_i(q_{1i}, \dot{q}_{1i})\dot{q}_{1i} + G_i(q_{1i}) + K_i q_{1i} = K_i x_{1i}, \\ \dot{x}_{1i} = x_{2i}, \end{cases} \quad (3)$$

$$\dot{x}_{2i} = J_i^{-1}u_i - J_i^{-1}K_i(x_{1i} - q_{1i}). \quad (4)$$

$$(5)$$

3 Distributed Consensus Backstepping Control

3.1 Control Objective

The issue of distributed consensus control in networked flexible-joint robotic systems is addressed based on backstepping approach. The control goal is to design a suitable distributed algorithm with local communication interactions for networked flexible-joint robotic manipulators such that each of robot can track the desired consensus reference trajectory.

Now it is ready to define the concept of consensus tracking of networked flexible-joint robotic systems below.

Definition 1. The control protocols u_i , ($i = 1, 2, \dots, n$) for each manipulator i are said to solve the consensus tracking problem, if each manipulator i can get cooperative tracking the desired trajectory q_{1d} , that is, $\lim_{t \rightarrow \infty} \|q_{i1} - q_{1d}\| = 0$, $\lim_{t \rightarrow \infty} \|\dot{q}_{i1} - \dot{q}_{1d}\| = 0$.

3.2 Controller Design

It should be noted that in traditional backstepping technique, the linear part in the control model should be required to be of minimum phase with relative degree one. However, the linear part (4) and (5) of the dynamics model (2) with respect to the motor unit in this paper is actually of relative degree two. Therefore, it is highly expected to developed conventional backstepping for designing the distributed consensus tracking control strategy for networked flexible-joint robotic systems (1)–(2).

In order to perform the distributed consensus tracking control scheme for networked flexible-joint robotic manipulator systems (1)–(2), the consensus tracking torque for dynamics (3)–(5) will be derived by the procedure of backstepping technique step by step.

As mentioned in the previous section, this paper has the following assumption for directed interconnection network topology.

Assumption 1. The digraph \mathcal{G} contains a directed spanning tree.

As usual, it is always assumed that a feasible desired consensus reference trajectory q_{1d} satisfying following assumption [8]:

Assumption 2. q_{1d} is differentiable up to fourth order, i.e., q_{1d} and $q_{1d}^{(m)}$, ($m = 1, 2, 3, 4$) are bounded.

For the dynamics model (3), assume that the system parameters are exactly known, and let x_{1i} be regarded as its control inputs. In what follows, sliding mode control approach is applied to design a virtual control input u_{1i} . We first define the consensus tracking position error $q_{ie} = q_{1i} - q_{1d}$, and its derivative along time is $\dot{q}_{ie} = \dot{q}_{1i} - \dot{q}_{1d}$.

Then, a sliding mode auxiliary variable $\hat{q}_{i1} \in \mathbb{R}^p$ for the i -th manipulator system can be designed as

$$\hat{q}_{i1} = \sum_{j=1}^n a_{ij} (q_{1i} - q_{1j}) + c_{id} q_{ie}. \quad (6)$$

Therefore, a novel sliding mode vector $s_i \in \mathbb{R}^p$ is introduced below

$$s_i = \dot{q}_{ie} + \Lambda_i \hat{q}_{i1}, \quad (7)$$

where Λ_i is a diagonal matrix with positive elements.

By the above preparations, the virtual controller u_{1i} can be constructed as:

$$u_{1i}(q_{1i}, \dot{q}_{1i}, q_{1d}, \dot{q}_{1d}, \ddot{q}_{1d}) = q_{1i} + K_i^{-1} [D_i(q_{1i})a_{1i} + C_i(q_{1i}, \dot{q}_{1i})v_{1i} + G_i(q_{1i}) - K_{si}s_i], \quad (8)$$

where $v_{1i} = \dot{q}_{1d} - \Lambda_i \hat{q}_{i1}$, $a_{1i} = \dot{v}_{1i}$, and K_{si} is a diagonal array with positive elements.

Next, we further modify the error between the motor rotation angle x_{1i} and the virtual control input u_{1i} by defining $z_{i1} = x_{1i} - u_{1i}$, and substituting it into Eq. (3) gives

$$D_i(q_{1i})\dot{s}_i + C_i(q_{1i}, \dot{q}_{1i})s_i + K_{si}s_i = K_iz_{i1}. \quad (9)$$

For the closed loop system (9), consider Lyapunov function candidate

$$V_{i1} = \frac{1}{2}s_i^T D_i(q_{1i})s_i, \quad (10)$$

and then its derivative is computed as

$$\begin{aligned} \dot{V}_{i1} &= \frac{1}{2}s_i^T \dot{D}_i(q_{1i})s_i + s_i^T D_i(q_{1i})\dot{s}_i \\ &= \frac{1}{2}s_i^T \dot{D}_i(q_{1i})s_i + s_i^T [-C_i(q_{1i}, \dot{q}_{1i})s_i - K_{si}s_{1i} + K_iz_{i1}] \\ &= \frac{1}{2}s_i^T (\dot{D}_i(q_{1i}) - 2C_i(q_{1i}, \dot{q}_{1i}))s_i - s_i^T K_{si}s_{1i} + s_i^T K_iz_{i1} \\ &= -s_i^T K_{si}s_{1i} + s_i^T K_iz_{i1}. \end{aligned} \quad (11)$$

Obviously, according to Lyapunov stable theory, if $z_{i1} = 0$, $\dot{V}_{i1} < 0$, we have $s_i \rightarrow 0$ as $t \rightarrow \infty$. In that way, a new virtual controller needs to be designed to stabilize $z_{i1} \rightarrow 0$ as $t \rightarrow \infty$.

In this way, second Lyapunov function are selected as

$$V_{i2} = V_{i1} + \frac{1}{2}z_{i1}^T K_i z_{i1}. \quad (12)$$

By using $\dot{z}_{i1} = \dot{x}_{1i} - \dot{u}_{1i} = x_{2i} - \dot{u}_{1i}$, we have

$$\dot{V}_{i2} = \dot{V}_{i1} + z_{i1}^T K_i \dot{z}_{i1} = -s_i^T K_{si}s_i + s_i^T K_i z_{i1} + z_{i1}^T K_i (x_{2i} - \dot{u}_{1i}). \quad (13)$$

The second virtual control input thereby designed as $x_{2i} = -s_i - c_{i1}z_{i1} + \dot{u}_{1i} + z_{2i}$, which further lead to that $z_{2i} = s_i + c_{i1}z_{i1} + \dot{u}_{1i} + x_{2i}$, where c_{i1} is positive diagonal matrix. Then one can get

$$\dot{V}_{i2} = -s_i^T K_{si}s_i - z_{i1}^T c_{i1} K_i z_{i1} + z_{i1}^T K_i z_{i2}. \quad (14)$$

Following the above same analysis procedure, we have $s_{1i} \rightarrow 0$, $z_{i1} \rightarrow 0$ as $t \rightarrow \infty$ if $z_{i2} \rightarrow 0$, $\dot{V}_{i2} < 0$. Another Lyapunov function will be constructed to guarantee $z_{i2} \rightarrow 0$ as $t \rightarrow \infty$.

By the use of repeating previous procedures, choosing a new Lyapunov function candidate

$$V_i = V_{i2} + \frac{1}{2}z_{i2}^T K_i z_{i2}, \quad (15)$$

and computing its derivative \dot{V}_i , we have

$$\dot{V}_i = \dot{V}_{i2} + z_{i2}^T K_i \dot{z}_{i2}. \quad (16)$$

Differentiating z_{i2} yeilds:

$$\dot{z}_{i2} = \dot{s}_i + c_{i1}\dot{z}_{i1} + \dot{x}_{2i} - \ddot{u}_{1i} = \dot{s}_i + c_{i1}\dot{z}_{i1} - \ddot{u}_{1i} + J_i^{-1}u_i - J_i^{-1}K_i(x_{1i} - q_{1i}). \quad (17)$$

By substituting \dot{V}_{i2} , \dot{z}_{i2} into \dot{V}_i , one can obtain:

$$\begin{aligned} \dot{V}_i &= -s_i^T K_{si} s_{1i} + s_i^T K_i z_{i1} + z_{i1}^T K_i (x_{2i} - \dot{u}_{1i}) \\ &\quad + z_{i2}^T K_i (\dot{s}_i + c_{i1}\dot{z}_{i1} - \ddot{u}_{1i} + J_i^{-1}u_i - J_i^{-1}K_i(x_{1i} - q_{1i})). \end{aligned} \quad (18)$$

Thus, the consensus tracking control protocol for the networked dynamic system (5) can be derived as

$$u_i = -J_i[\dot{s}_i + c_{i1}\dot{z}_{i1} - \ddot{u}_{1i} + c_{i2}z_{i2} + z_{i1}] + K_i(x_{1i} - q_{1i}), \quad (19)$$

where c_{i2} is a diagonal matrix with positive elements, then it gives

$$\dot{V}_i = -s_i^T K_{si} s_i - z_{i1}^T c_{i2} K_i z_{i1} - z_{i2}^T c_{i2} K_i z_{i2} \leq 0. \quad (20)$$

According to the Lyapunov stability, the consensus tracking control protocol (19) guarantees that the equilibrium point $s_i = 0$, $z_{i1} = 0$ and $z_{i2} = 0$ is asymptotically stable.

Finally, for our convenience, Eq. (7) can be rewritten in a vector form

$$\dot{\mathbf{q}}_e = -\Lambda_i \hat{\mathbf{q}} + \mathbf{s} = -\Lambda [(\mathbf{L} + \mathbf{A}) \otimes \mathbf{I}_p] \mathbf{q}_e + \mathbf{s}. \quad (21)$$

where $\mathbf{q} = [q_{1e}, q_{2e}, \dots, q_{ne}]^T$ and $\mathbf{s} = [s_1, s_2, \dots, s_n]^T$ and $\mathbf{A} = \text{diag}[c_{1d}, c_{2d}, \dots, c_{nd}]$. By Assumption 1, then the matrix $-(\mathbf{L} + \mathbf{A})$ is Hurwitz stable. Therefore, it can be proved that concluded that q_{1d} , that is, $\lim_{t \rightarrow \infty} \|q_{i1} - q_{1d}\| = 0$, $\lim_{t \rightarrow \infty} \|\dot{q}_{i1} - \dot{q}_{1d}\| = 0$, ($i = 1, 2, \dots, n$).

On the basis of the above discussions, the main results of distributed consensus backstepping control in networked flexible-joint robotic systems are now readily given in the following.

Theorem 1. Under Assumptions 1–2, by using control torque (19), the networked flexible joint manipulator systems (3)–(5) can realize the consensus tracking performance, that is, q_{1d} , that is, $\lim_{t \rightarrow \infty} \|q_{i1} - q_{1d}\| = 0$, $\lim_{t \rightarrow \infty} \|\dot{q}_{i1} - \dot{q}_{1d}\| = 0$, ($i = 1, 2, \dots, n$).

Remark 1. It is interesting to point out that different from some previous consensus scheme for the networks of flexible-joint manipulators, the developed distributed cooperative consensus scheme in this paper can fully guarantee to track a desired common reference trajectory with a time-varying velocity in the sense of Definition 1. It is obvious that the proposed consensus controller in Refs [6] and [7] are no longer valid for realizing consensus case of a time-varying velocity reference trajectory. Apparently, Theorem 1 is actually an important improvement of the backstepping tracking control technique for a single flexible-joint manipulator with the non-adaptive case in the literature [8].

4 Simulation Examples

To verify the correctness of Theorem 1, four single-link flexible manipulators (labeled 1–4) (1) and (2) and one virtual leader (labeled 0) are taken as the research objects.

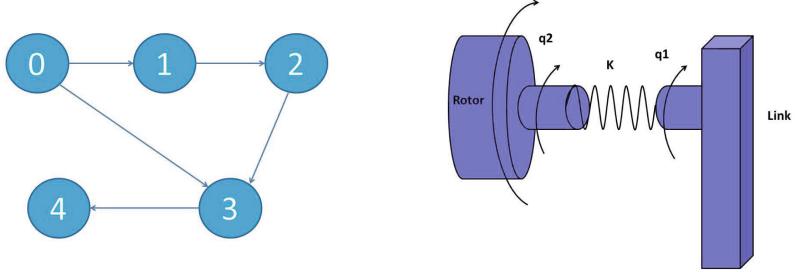


Fig. 1. The position communication topology graph of the leader and four followers nonidentical two-link flexible joint manipulators.

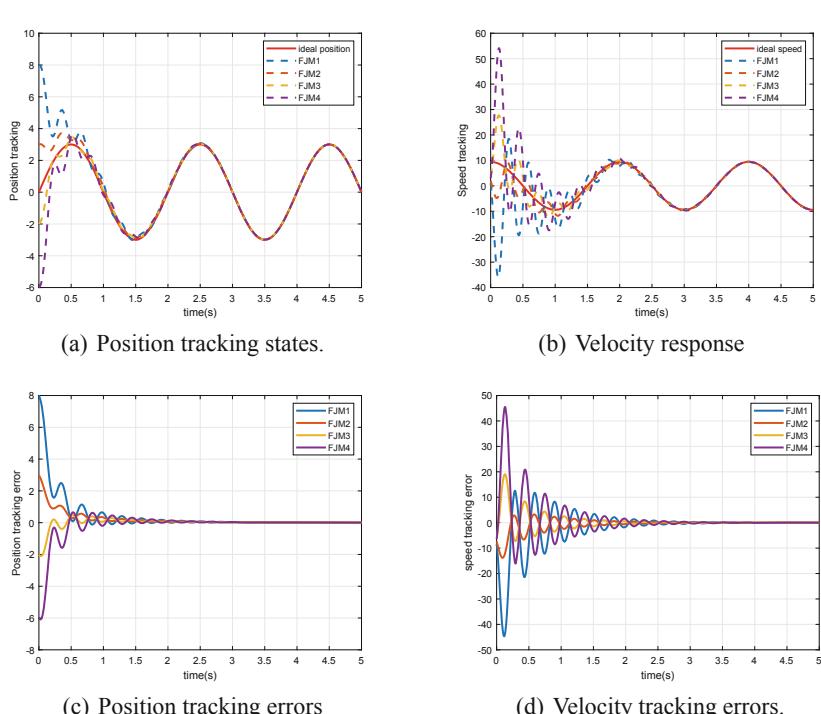


Fig. 2. Simulation results of the four flexible manipulators under communication topology \mathcal{G}

The detailed system parameters are $D_i = I_i$, $C_i = 0$, $G_i = M_i g l_i \sin q_{1i}$, $I_i = 1$, $J_i = 1$, $M_i = 0.5$, $K_i = 100$, $g = 9.8$, $l_i = 1$.

The position communication topology graph of the leader and two-link flexible joint manipulator model are shown in Fig. 1, which satisfies Assumption 1. Besides, the desired trajectory is chosen as $q_{1d} = 1.5 \sin(\pi t)$. The rest control parameters are selected as $a_{ij} = 2$, $c_{id} = 2$, $c_{i1} = 5$, $c_{i2} = 5$, $K_{si} = 3$, and the initial conditions are: $q_{11}(0) = 6$, $q_{12}(0) = 3$, $q_{13}(0) = -2$, $q_{14}(0) = -5$, $\dot{q}_{1i}(0) = 0$, $q_{21}(0) = 1$, $q_{22}(0) = 2$, $q_{23}(0) = 0$, $q_{24}(0) = 3$. It can be observed that all the conditions required in the Theorem 1 are satisfied. Then we perform the numerical simulations by using the consensus tracking control inputs (19) of the Theorem 1, the corresponding simulation results are shown in Fig. 2.

Figures 2(a) and 2(b) depict the trajectory tracking performance and velocity response, respectively, which shows that all the flexible manipulators can complete tracking the desired trajectory. Figure 2(c) and 2(d) show that the tracking errors of all the flexible manipulators tend to zero. Accordingly, the consensus tracking control strategy proposed in Theorem 1 is correctness and effectiveness.

5 Conclusions

This paper has studied the distributed consensus backstepping control problem for networked flexible-joint robotic systems. A unified distributed consensus tracking algorithm has been developed for the cases of the directed network topology by systematically combining backstepping approach and sliding mode technique. It has been shown that the proposed cooperative algorithms can always guarantee the desired tracking consensus of networked nonidentical flexible-joint manipulator systems, and it also has robustness to some unmatched uncertainties. Finally, the simulation illustrative examples have been provided to verify the proposed control algorithms.

Funding. This work is supported by the National Science Foundation of China (Nos. 51875331).

References

1. Cao, Y., Yu, W., Ren, W., Chen, G.: An overview of recent progress in the study of distributed multi-agent coordination. *IEEE Trans. Ind. Inform.* **9**(1), 427–438 (2013)
2. Zhi, Y., Jouandeau, N., Cherif, A.A.: A survey and analysis of multi-robot coordination. *Int. J. Adv. Robot. Syst.* **10**(12), 399 (2013)
3. Wu, X., Lan, X., Jin, Z.: Distributed adaptive tracking backstepping control in networked nonidentical Lagrange systems. *Nonlinear Dyn.* **78**(2), 1137–1148 (2014)
4. Spong, M.W.: Modeling and control of elastic joint robots. *J. Dyn. Syst. Measur. Control* **109**(4), 310–318 (1987)
5. Rsetam, K., Cao, Z., Man, Z.: Hierarchical sliding mode control applied to a single-link flexible joint robot manipulator. In: *IEEE International Conference on Advanced Mechatronic Systems*, pp. 476–481 (2016)
6. Nuno, E., Valle, D., Sarras, I., Basanez, L.: Leader-follower and leaderless consensus in networks of flexible-joint manipulators. *Eur. J. Control* **20**(5), 249–258 (2014)

7. Avilabecerril, S., Espinosaperez, G., Panteley, E., Ortega, R.: Consensus control of flexible-joint robots. *Int. J. Control.* **88**(6), 1201–1208 (2015)
8. Oh, J.H., Lee, J.S.: Control of flexible joint robot system by backstepping design approach. *Intell. Autom. Soft Comput.* **5**(4), 267–278 (1999)
9. Brogliato, B., Ortega, R., Lozano, R.: Global tracking controllers for flexible-joint manipulators: a comparative study. *Automatica* **31**(7), 941–956 (1995)
10. Avilabecerril, S., Espinosaperez, G.: Consensus control of flexible joint robots with uncertain communication delays. In: *Advances in Computing and Communications*, pp. 8–13 (2012)



Event-Triggered Stochastic Consensus for Networked Lagrangian Systems

Suying Pan¹, Zhiyong Ye², Lan Xiang^{3(✉)}, and Jin Zhou¹

¹ Shanghai Institute of Applied Mathematics and Mechanics,
School of Mechanics and Engineering Science, Shanghai University,
Shanghai 200072, People's Republic of China

² School of Science, Chongqing University of Technology,
Chongqing 400054, People's Republic of China

³ Department of Physics, School of Science, Shanghai University,
Shanghai 200444, People's Republic of China
lxiang@shu.edu.cn

Abstract. The problem of distributed event-triggered mean square consensus for networked Lagrangian systems with communication delays and semi-Markov switching topologies is considered in this brief. By considering event-triggered sampling control with stochastic disturbances, the distributed stochastic consensus protocol is proposed. Furthermore, by using stochastic analysis theory, with a combination of both algebraic graph tools and linear matrices inequalities (LMIs), a general delay-dependent criterion is presented for solving stochastic consensus problems in terms of mean square under the semi-Markov switching network topology structure with communication delays. Meanwhile, Zeno behavior of the triggering time sequences is also avoided. Subsequently, a typical example of four manipulators with two links is provided to demonstrate the developed consensus methodology in this brief.

Keywords: Lagrangian systems · Stochastic consensus · Event-triggered · Time varying communication delays · Semi-Markov switching topologies

1 Introduction

The cooperative behavior and its control of networked Lagrangian systems have received substantive attention [1]. There exist two elementary factors drive the research. Firstly, Lagrangian system can model the dynamical behavior of a large number of physical and mechanical objects such as robotic manipulator, electrical systems, flying spacecrafts, and walking robots. Secondly, the cooperative control of networked Lagrangian system has the advantage of performing the complex task more effectively than their individual counterparts. As a result, a

National Science Foundation of China (Nos. 51875331 and 11672169).

large amount of cooperative control protocols have been devised to achieve the consensus of networked Lagrangian system [2,3].

In networked control systems, as is known to all, the unnecessary communication may lead to low efficient performance implementations including communication bandwidth and processor usage, energy consumption, network congestion. In general, in order to reduce unnecessary communication and save computation resources, the event-triggered controller is proposed and developed, which updates occur only if a predefined triggering condition is violated, and so it helps to reduce the amount of actuator update. As a result, some excellent research works have been reported on distributed consensus problems based-event-triggered for networked multi-agent systems. However, the relevant most results were mainly focused on the agents with simple single and double-integrator linear dynamics [4]. For the nonlinear agent dynamics, especially for the Lagrangian system, there exists a few works [5]. For multiple Euler-Lagrange systems, the distributed consensus problem with the event-triggered information was addressed [5]. It should be noted that most existing event-triggered control strategies are addressed only in deterministic setting. To our best knowledge, for the stochastic setting, there exists few researches on the distributed event-based control for networked Euler-Lagrange systems subject to input random disturbance.

This brief is focused on the distributed event-triggered consensus problem for networked Lagrangian systems in the sense of stochastic setting with semi-Markov switching network topologies. The main contributions of this work in comparison with the previous results lie in three aspects: Firstly, the mean square consensus problem based-event-triggered for the networked Lagrangian system with both semi-Markov switching topologies and time-varying communication delays are taken into account. Secondly, the decentralized adaptive control scheme under input random disturbance is designed. Finally, the convergence analysis of the proposed event-triggered consensus scheme is performed by the use of the Lyapunov-Krasovskii functional and linear matrices inequalities (LMIs). To this end, one representative example of a team of four manipulators with two links is presented to show the effectiveness of the developed methodology of event-triggered stochastic consensus.

2 Preliminaries

Some necessary preliminaries related to basic notations and concepts of algebraic graph theory can refer to [6], and the relative knowledge for the semi-Markov process can be seen in [7].

Let the networked dynamics of N agent systems be represented by the equations of motion of the from [8]

$$M_i(q_i)\ddot{q}_i + C_i(q_i, \dot{q}_i)\dot{q}_i + g_i(q_i) = \tau_i, \quad i = 1, 2, \dots, N, \quad (1)$$

in which the vector of generalized coordinate is denoted by $q_i \in \mathbf{R}^n$, the inertia matrix $M_i(q_i) \in \mathbf{R}^{n \times n}$ is a symmetric positive-definition, the Coriolis and centripetal torques vector is represented by $C_i(q_i, \dot{q}_i) \in \mathbf{R}^n$, $g_i(q_i)$ is used to stand

for the vector of gravitational torque, and the control input vector for the i th agent system is denoted by $\tau_i \in \mathbf{R}^n$.

For subsequent analysis, the following key properties for the networked Lagrangian systems (1) are assumed to hold:

Property 1. There exists four positive constants k_M , k_m , k_c , k_g possessing $\|g_i(q_i)\| < k_g$, and $\|C_i(q_i, \dot{q}_i)\| \leq k_c \|\dot{q}_i\|$, $0 < k_m I_n \leq M_i(q_i) \leq k_M I_n$.

Property 2. The matrix $M_i(q_i) - 2C_i(q_i, \dot{q}_i)$ is skew symmetric satisfying $r^T(M_i(q_i) - 2C_i(q_i, \dot{q}_i))r = 0$ for a given vector $r \in \mathbf{R}^n$.

Property 3. For a unknown vector Θ_i but constant parameters, and the regressor matrix $Y_i(q_i, \dot{q}_i, z, y)$, this linearly parameterizable property is introduced as follows $M_i(q_i)z + C_i(q_i, \dot{q}_i)y + g_i(q_i) = Y_i(q_i, \dot{q}_i, x, y)\Theta_i$ for all vectors $z, y \in \mathbf{R}^n$.

Definition 1. The networked Lagrangian system (1) with Semi-Markov switching topologies is said to realize the stochastic mean square consensus under the designed event-triggered protocols, if $\lim_{t \rightarrow 0} E[q_i - q_j]^2 = 0$, $\lim_{t \rightarrow 0} E[\dot{q}_i - \dot{q}_j]^2 = 0$ for all $i, j \in \mathcal{V}$.

Lemma 1 [2]. If a directed spanning tree exists in the network topology graph $\mathcal{G}(r(t))$, then the matrix $CLC^{-1} = \text{diag}\{0, L_r\}$, where C is invertible, and $-L_r$ is Hurwitz stable.

3 Design of Event-Based Adaptive Controllers

First of all, the following auxiliary variables will be introduced as

$$q_{ri} = -\frac{\alpha}{\sum_{j=1}^N a_{ij}(r(t))} \sum_{j=1}^N a_{ij}(r(t))(q_i(t) - q_j(t - \tau(t))), \quad i = 1, 2, \dots, N, \quad (2)$$

and

$$s_i = \dot{q}_i + \frac{\alpha}{\sum_{j=1}^N a_{ij}(r(t))} \sum_{j=1}^N a_{ij}(r(t))(q_i(t) - q_j(t - \tau(t))), \quad i = 1, 2, \dots, N, \quad (3)$$

in which the term of $\tau(t)$ is time varying communication delay and $0 < \tau(t) \leq \bar{\tau}$, and α is a suitably chosen positive constant.

We also assume that every agent has at least one neighbour, which means that $\sum_{j=1}^N a_{ij}(r(t)) \neq 0$, $\forall i$. For agent i with an unknown but constant vector Θ_i , the estimate of Θ_i at time t is denoted by $\hat{\Theta}_i(t)$. Accordingly, the adaption law of $\hat{\Theta}_i(t)$ is defined as

$$\dot{\hat{\Theta}}_i(t) = -\Lambda_i Y_i^T s_i(t), \quad i = 1, 2, \dots, N, \quad (4)$$

in which Λ_i is a designed symmetric positive-definite matrix. Thus, the event-triggered controller with input random disturbance for agent i is designed in the following form

$$\tau_i(t) = -K_i s_i(t) + Y_i(t) \hat{\Theta}_i + \hat{\alpha}(t) \eta_i, \quad i = 1, 2, \dots, N, \quad (5)$$

where the feedback gain matrix K_i is symmetric positive definite one, and η_i stands for standard white noise.

Consequently, by Property 3, the resulting closed loop controlled system for (1) could be obtained

$$\begin{aligned} M_i(q_i)\dot{s}_i + C_i(q_i, \dot{q}_i)s_i - Y_i(q_i, \dot{q}_i, \dot{q}_{ri}, q_{ri})\tilde{\Theta}_i \\ = -K_i e_i(t) - K_i s_i(t) + \xi_i(t) + \hat{\alpha}(t)\eta_i, \end{aligned} \quad (6)$$

with the parameter estimation error $\tilde{\Theta}_i = \hat{\Theta}_i - \Theta_i$. In order to describe the triggered function, we first present the measurement state errors as follows

$$e_i(t) = s_i(t_k^i) - s_i(t), \quad \xi_i(t) = Y_i(t_k^i)\hat{\Theta}_i(t_k^i) - Y_i(t)\hat{\Theta}_i(t), \quad t \in [t_k^i, t_{k+1}^i], \quad (7)$$

furthermore, the triggered function is defined as

$$\begin{aligned} f_i(e_i(t), \xi_i(t)) &= \|\xi_i(t)\| + \lambda_{\max}(K_i)\|e_i(t)\| \\ &- \frac{\gamma_i}{2}\lambda_{\min}(K_i)\|s_i(t_k^i)\| - \delta_i(t), \text{ for } t \in [t_k^i, t_{k+1}^i], \end{aligned} \quad (8)$$

where $t_{k+1}^i = \inf\{t > t_k^i \mid f_i(e_i(t), \xi_i(t)) > 0\}$, $0 < \gamma_i < 1$, $\delta_i(t) = \beta_i(t)\sqrt{\lambda_{\min}(K_i)}e^{-b_i t}$ with $\beta_i(t) > 0$ and $0 < b_i < 1$, $0 < \beta_{i,1} < \beta_{i,2} < 1$.

4 Event-Triggered Mean Square Consensus

Theorem 1. *The networked Lagrangian system (1) with semi-Markov switching topologies having a directed spanning tree can reach the distributed mean square consensus under the designed event-triggered protocols (5), if for the allowable $\bar{\tau} > 0$, there exist the positive definite symmetric matrices P , P_v , W , K_i , Λ_i and positive constants α , β , $\tilde{\beta}$, β_i , δ , $0 < \gamma_i, \varepsilon_v, \varpi < 1$ such that*

$$\hat{\Xi}_v = \begin{bmatrix} \hat{\Xi}_{v1} & \hat{\Xi}_{v2} \\ * & \hat{\Xi}_{v3} \end{bmatrix} < 0, \Phi = \begin{bmatrix} -2\alpha P + p\beta P + P & \alpha P \\ * & -\beta P \end{bmatrix} < 0 \quad (9)$$

for all $\nu \in S = \{1, 2, \dots, \omega\}$. Where $\hat{\Xi}_{v1} = \begin{bmatrix} \hat{\Xi}_{v1}(1, 1) & \hat{\Xi}_{v1}(1, 2) \\ * & -\frac{1}{\bar{\tau}}W \end{bmatrix}$, $\hat{\Xi}_{v2} = \begin{bmatrix} -\alpha W(3I_{(N-1)n} + L_{rv} \otimes I_n) \\ -\alpha W(2I_{(N-1)n} - L_{rv} \otimes I_n) \end{bmatrix}$, $\hat{\Xi}_{v3} = -\frac{1}{\bar{\tau}(1+\varpi)}W$, $\hat{\Xi}_{v1}(1, 1) = \sum_{u=1}^{\omega} \bar{\lambda}_{vu} P_u - 2\alpha P_v(L_{rv} \otimes I_n) + \varepsilon_v P_v$, $\hat{\Xi}_{v1}(1, 2) = -2\alpha P_v + 2\alpha P_v(L_{rv} \otimes I_n)$.

Proof. First step: Consider a following Lyapunov function

$$V_i(t) = \frac{1}{2} \sum_{i=1}^N s_i^T M_i(q_i) s_i + \frac{1}{2} \sum_{i=1}^N \tilde{\Theta}_i^T \Lambda_i^{-1} \tilde{\Theta}_i. \quad (10)$$

According to Property 2, and by the triggered function in (8) and the absolute continuity of $E\{V_i(t)\}$ for $t \geq 0$, one can obtain

$$\begin{aligned} E\{\mathcal{L}V_i(t)\} &= E\{-\sum_{i=1}^N s_i(t)^T K_i s_i(t) + \sum_{i=1}^N s_i(t)^T \xi_i(t) \\ &- \sum_{i=1}^N s_i(t)^T K_i e_i(t) + \sum_{i=1}^N s_i^T(t) \hat{\alpha}(t) \eta_i\}. \text{ a.e. } \end{aligned} \quad (11)$$

Due to $E\{\sum_{i=1}^N s_i^T \hat{\alpha}(t) \eta_i\} = 0$, and on the basic of definition of δ_i , one can obtain

$$\begin{aligned} \sum_{i=1}^N (1 - \gamma_i) \lambda_{\min}(K_i) \int_0^t E\{\|s_i(\phi)\|^2\} d\phi + E\{V_i(t)\} \\ \leq \sum_{i=1}^N \frac{\beta_i^2}{4b_i\gamma_i} + E\{V_i(0)\}, \end{aligned} \quad (12)$$

which implies that $E\{V_i\}$ is bounded. From (10), $E\{\tilde{\Theta}_i\}$ and $E\{s_i\}$ are bounded obtained. In addition, by Property 3, the following inequality holds

$$\|Y_i \Theta_i\| \leq \|M_i\| \|\ddot{q}_{ri}\| + \|C_i\| \|\dot{q}_{ri}\| + \|g_i\|. \quad (13)$$

By (12), the inequality is derived

$$\sum_{i=1}^N (1 - \gamma_i) \lambda_{\min}(K_i) \int_0^t E\{\|s_i(\phi)\|^2\} d\phi \leq E\{V_i(0)\} + \sum_{i=1}^N \frac{\beta_i^2}{4b_i\gamma_i}, \quad (14)$$

which means $\int_0^t E\{\|s_i(\phi)\|^2\} d\phi$ is bounded and thus $E\{s_i\} \in \mathcal{L}_2$.

Second step: Here, based on the assumption of a directed spanning tree in $\mathcal{G}(r(t))$, for (3), we have

$$\dot{q}(t) = -\alpha q(t) + \alpha q(t - \tau(t)) - \alpha(\tilde{L}(r(t)) \otimes I_n)q(t - \tau(t)) + s(t), \quad (15)$$

where $\tilde{A}(r(t)) = [\tilde{a}_{ij}(r(t))] \in \mathbf{R}^{N \times N}$ is defined as $\tilde{a}_{ij}(r(t)) = \frac{a_{ij}(r(t))}{\sum_{j \in N_i} a_{ij}(r(t))}$, $i, j = 1, 2, \dots, N$ and $\tilde{L}(r(t)) = I_N - \tilde{A}(r(t))$. Because of $\mathcal{G}(r(t))$ with a directed a spanning tree, further, we can conclude that $\tilde{L}(r(t))$ exists a simple zero eigenvalue, and all other eigenvalues of $\tilde{L}(r(t))$ are on the open right half-plan [2]. By the same decomposition way in [1], (15) is decoupled as

$$\dot{x}_e = -\alpha x_e(t) + \alpha x_e(t - \tau(t)) - \alpha(L_{rv} \otimes I_n)x_e(t - \tau(t)) + B_e, \quad (16)$$

$$\dot{x}_1 = -\alpha x_1(t) + \alpha x_1(t - \tau(t)) + \sum_{j=1}^N \mu_j s_j. \quad (17)$$

Further, system (16) can be written as

$$\begin{aligned} \dot{x}_e &= (-\alpha I_{(N-1)n} + \alpha(L_{rv} \otimes I_n)) \int_{t-\tau(t)}^t \dot{x}_e(\varphi) d\varphi \\ &\quad - \alpha(L_{rv} \otimes I_n)x_e(t) + B_e. \end{aligned} \quad (18)$$

Here, the semi-Markov-based Lyapunov-Krasovskii functional is selected in the following:

$$\bar{V}(x_e(t), r(t), t) = \bar{V}_1(x_e(t), r(t), t) + \bar{V}_2(x_e(t), t), \quad (19)$$

where $\bar{V}_1(x_e(t), r(t), t) = x_e(t)^T P(r(t))x_e(t)$ and $\bar{V}_2(x_e(t), t) = \int_{t-\bar{\tau}}^t \int_\varphi^t \dot{x}_e(\psi)^T W \dot{x}_e(\psi) d\psi d\varphi$. Then, similar to the derivation of (11), it derives

$$\begin{aligned} E\{\mathcal{L}\{\bar{V}_1(x_e(t), r(t), t) + \bar{V}_2(x_e(t), t)\}\} &= E\{2x_e^T(t)P_v\dot{x}_e(t) + \bar{\tau}\dot{x}_e^T(t) \\ &\quad W \dot{x}_e(t) + x_e^T(t)(\Sigma_{u \in S} \lambda_{vu}(h)P_u)x_e(t) - \int_{t-\bar{\tau}}^t \dot{x}_e^T(\varphi)W \dot{x}_e(\varphi)d\varphi\}, \text{ a.e.} \end{aligned} \quad (20)$$

where $\bar{\lambda}_{vu} = E\{\lambda_{vu}(h)\} = \int_0^\infty \lambda_{vu}(h)f_v(h)dh$. Further, by combining Young's inequality, and Jensen inequality, with the similar analysis procedure [3], then

we can derived that $E\{\mathcal{L}\{\bar{V}_1(x_e(t), r(t), t) + \bar{V}_2(x_e(t), t)\}\} < 0$, which means that there exists a scalae $\beta > 0$ such that $E\{\mathcal{L}\{\bar{V}_1(x_e(t), r(t), t) + \bar{V}_2(x_e(t), t)\}\} < -\beta E\|x_e(t)\|^2$. Using Dynkin's formula, one can get $E\{\bar{V}_1(x_e(t), r(t), t) + \bar{V}_2(x_e(t), t)\} - E\{\bar{V}_1(x_e(0), r(0), 0) + \bar{V}_2(x_e(0), 0)\} < -\tilde{\beta}E\int_0^T\|x_e(t)\|dt$.

Further, letting $T \rightarrow \infty$, $E\int_0^\infty\|x_e(t)\|^2dt < \infty$. Therefore, one has $\lim_{t \rightarrow \infty}E[x_e]^2 = 0$. For system (17), we take $\bar{V}_3(x_1(t), t) = x_1^T(t)Px_1(t)$. Based on Young's inequality, assume $\bar{V}_3(x_1(t - \tau(t)), t) < p\bar{V}_3(x_1(t), t)$ with $p > 1$, for any $\beta > 0$, we have $E\{\mathcal{L}\bar{V}_3(x_1(t), t)\} \leq E\{\varsigma_1^T(t)\Phi\varsigma_1(t)\}$, where $\varsigma_1^T(t) = [x_1^T(t) \ x_1^T(t - \tau(t))]$. For some sufficiently small $\delta > 0$, $p = 1 + \delta$, $\Phi < 0$ holds in (9), which implies that the Razumikhin derivation condition $E\{\mathcal{L}\bar{V}_3(x_1(t), t)\} \leq -\varepsilon E\{\|x_1(t)\|^2\}$ is satisfied. According to the Razumikhin stability theorem in [9], the system is mean square asymptotically stable. Thus, it means that (15) is input to state stable. Since $E\{s_i\}$ is bounded, it yields that $E\{q_i\}$ and $E\{\dot{q}_i\}$ are both bounded. Further, we can get $E\{\ddot{q}_ri\}$ and $E\{\dot{q}_ri\}$ are bounded. $\|Y_i\|$ is upper bounded by a positive constant on the basic of (15). From (6), M_i, C_i and $E\{s_i\}$ are bounded, one gets that $E\{\dot{s}_i\}$ is bounded. Therefore $E\{\dot{\Theta}_i(t)\}$, Y_i , $E\{\dot{\Theta}_i(t)\}$, $E\{s_i\}$ and $E\{\dot{s}_i\} \in \mathcal{L}_\infty$, which means $E\{\ddot{V}_i(t)\}$ is bounded. By Barbalat's lemma [10], we conclude that $\lim_{t \rightarrow \infty}E\{\dot{V}_i(t)\} = 0$, further, $\lim_{t \rightarrow \infty}E\{s_i\} = 0$. Since (15) is input to state stable and $\lim_{t \rightarrow \infty}E\{s_i\} = 0$, and $\lim_{t \rightarrow \infty}E\{q(t)\} = 0$. Thus, $\lim_{t \rightarrow \infty}E[q_i(t) - q_j(t)]^2 = 0$, one derives $\lim_{t \rightarrow \infty}E[\dot{q}_i(t) - \dot{q}_j(t)]^2 = 0$ for $\forall i, j$.

Third step: Let $T = t_{k+1}^i - t_k^i$. From (7), we can get the Dini derivative of $e_i(t)$ with respect to time $t \in [t_k^i, t_{k+1}^i]$: $D^+\|e_i(t)\| \leq \|\dot{s}_i(t)\| \leq U_e$ where $U_e = D^+\|s_i(t)\|$ for all $i \in 1, 2, \dots, N$. By the similar way, we can also obtain $D^+\|\xi_i(t)\| \leq U_\xi$ where $U_\xi = D^+\|\dot{Y}_i(t)\dot{\Theta}_i(t) + Y_i(t)\dot{\Theta}_i(t)\|$ for all $i \in 1, 2, \dots, N$. The following inequality can be got

$$\|\xi_i(t)\| + \lambda_{\max}(K_i)\|e_i(t)\| \leq 2U(t - t_k^i), \quad U = \max\{U_\xi, \lambda_{\max}(K_i)U_e\}. \quad (21)$$

Combining (8), it can be seen $\|\xi_i(t)\| + \lambda_{\max}(K_i)\|e_i(t)\| = \frac{\gamma_i}{2}\lambda_{\min}(K_i)\|s_i(t_k^i)\| + \delta_i(t)$ holds at t_{k+1}^i . From the first two steps, we may have $E\{s_i(t)\} = 0$ at t_{k+1}^i due to the fact that we have noted $E\{s_i\} \rightarrow 0$ as $t \rightarrow \infty$. But, it does not means the consensus is realized as $\dot{s}_i(t)$ can be nonzero as t_{k+1}^i . So, the event can only be triggered in the following two situations at t_{k+1}^i :

(1) If $\|s_i(t_{k+1}^i)\| \neq 0$, then $\|\xi_i(t_{k+1}^i)\| + \lambda_{\max}(K_i)\|e_i(t_{k+1}^i)\| = \frac{\gamma_i}{2}\lambda_{\min}(K_i)\|s_i(t_k^i)\| + \delta_i(t_{k+1}^i)$ holds. (2) If $\|s_i(t_{k+1}^i)\| = 0$, then $\|\xi_i(t_{k+1}^i)\| + \lambda_{\max}(K_i)\|e_i(t_{k+1}^i)\| = \delta_i(t_{k+1}^i)$ holds. From the above two cases, it can be noted that $\|s_i(t_{k+1}^i)\| > 0$ for any $\|s_i(t_{k+1}^i)\| \neq 0$. From the inequality (21), we have $2U\delta_{(2)} \geq \delta_i(t) = \beta_i(t)\sqrt{\lambda_{\min}(K_i)} e^{-b_i(t_k^i + \delta_{(2)})}$. This implies that there exists a event time $\delta_{(2)}$ with a lower bounded satisfying the following equation $2U\delta_{(2)} = \beta_i(t)\sqrt{\lambda_{\min}(K_i)} e^{-b_i(t_k^i + \delta_{(2)})}$. Because U is upper bounded, so it can be concluded that at any finite time the solution is positive and time-dependent. As a result, we can see for every agent the zeno behavior is avoided.

5 Illustrative Examples

Example 1. In order to illustrate the applicability of the designed control strategy, a numerical simulation for a group of four two-link manipulators will be presented. Firstly, the dynamics of four two-link manipulators are given as [8]

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \ddot{q}_{i1} \\ \ddot{q}_{i2} \end{bmatrix} + \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & 0 \end{bmatrix} \begin{bmatrix} \dot{q}_{i1} \\ \dot{q}_{i2} \end{bmatrix} = \begin{bmatrix} \tau_{i1} \\ \tau_{i2} \end{bmatrix},$$

where $M_{11} = (m_1 + m_2)d_1^2 + m_2d_2^2 + 2m_2d_1d_2\cos(q_{i2})$, $M_{12} = M_{21} = m_2(d_1^2 + d_1d_2\cos(q_{i2}))$, $M_{22} = m_2d_2^2$, $C_{11} = -m_2d_1d_2\sin(q_{i2})\dot{q}_{i2}$, $C_{12} = -m_2d_1d_2\sin(q_{i2})\dot{q}_{i2} - m_2d_1d_2\sin(q_{i2})\dot{q}_{i1}$, $C_{21} = m_2d_1d_2\sin(q_{i2})\dot{q}_{i1}$, $g_{i1} = (m_1 + m_2)gd_1\sin(q_{i1}) + m_2gd_2\sin(q_{i1} + q_{i2})$, and $g_{i2} = m_2gd_2\sin(q_{i1} + q_{i2})$, in which g denotes the acceleration because of gravity, d_1 and d_2 denote the lengths of the first and second links of the manipulator, respectively; m_1 and m_2 represent the mass of the first and second of the manipulator. The corresponding physical parameters are chosen as $g = 9.8 \text{ m/s}^2$, $d_1 = 1.5 \text{ m}$, $d_2 = 1 \text{ m}$, $m_1 = 1 \text{ kg}$, $m_2 = 2 \text{ kg}$.

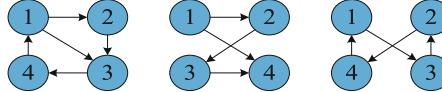


Fig. 1. The semi-Markov switching network topologies with respect to $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3$.

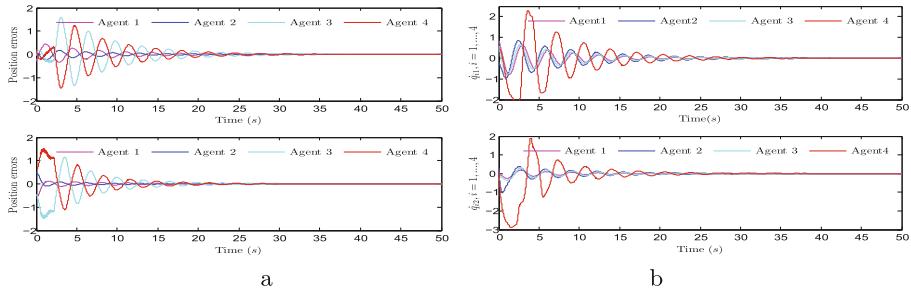


Fig. 2. a. The errors of position of the manipulators; b. The joint angels derivative of the manipulators.

The position initial values of each manipulator are presented as $q_1(0) = [\pi/15 \ \pi/16]^T$, $q_2(0) = [\pi/6 \ \pi/4]^T$, $q_3(0) = [\pi/9 \ \pi/9]^T$, $q_4(0) = [\pi/8 \ \pi/4]^T$. The joint angels derivative initial values are given as $\dot{q}_1(0) = [0.8 \ 0.2]^T$, $\dot{q}_2(0) = [-0.2 \ 0.32]^T$, $\dot{q}_3(0) = [0.6 \ -0.4]^T$, $\dot{q}_4(0) = [-0.5 \ 0.1]^T$. The control

gain matrices are selected as $K_1 = 5.2051I_2$, $K_2 = 3.9439I_2$, $K_3 = 4.2966I_2$, $K_4 = 5.2283I_2$, $\Lambda_1 = 1.8786I_5$, $\Lambda_2 = 2.3111I_5$, $\Lambda_3 = 3.6453I_5$, $\Lambda_4 = 1.5049I_5$.

The information exchange among the four manipulator is shown in Fig. 1. Figure 2 a shows that the errors of joint angels of four manipulators tend to zero, which means mean square consensus can be obtained. The joint angels derivative of the four manipulators also tend to zero in Fig. 2 b. The parameters estimation and the control input are showed in a and b of Fig. 3, respectively. From Fig. 3 a, we can see the estimate of $\hat{\theta}_i$, ($i = 1, 2, 3, 4$) will tend to stable. The control input will tend to zero when the manuscripts achieve consensus.

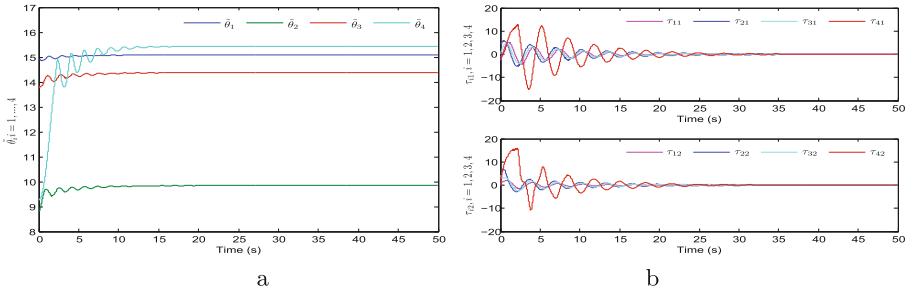


Fig. 3. a. The parameters estimation; b. The control input of each manipulator.

6 Conclusion

The stochastic event-triggered consensus problem for networked Lagrangian systems with time-varying communication delays under semi-Markov switching topologies has been tackled in this brief. By employing appropriate Lyapunov-Krasovskii functional and LMIs technologies, the stochastic event-triggered consensus conditions are established. Finally, a numerical example for four manipulators with two links is presented to manifest the effectiveness of the designed methods.

References

1. Liu, J., Ji, J.C., Zhou, J., Xiang, L., Zhao, L.Y.: Adaptive group consensus in uncertain networked Euler-Lagrange systems under directed topology. *Nonlinear Dyn.* **82**(3), 1145–1157 (2015)
2. Yu, J.W., Ji, J.C., Miao, Z.H., Zhou, J.: Formation control with collision avoidance for uncertain networked Lagrangian systems via adaptive gain techniques. *IET Control Theory A.* **12**(10), 1393–1401 (2018)
3. Miao, Z.H., Liu, J., Wang, G.Q., Zhou, J.: Group consensus in uncertain networked Euler-Lagrange systems with stochastic disturbances. *Scientia Sinica Informationis* **46**(11), 1608–1620 (2016)

4. Dimarogonas, D.V., Frazzoli, E.F., Johansson, K.H.: Distributed event-triggered control for multi-agent systems. *IEEE Trans. Autom. Control* **57**(5), 1291–1297 (2012)
5. Huang, N., Duan, Z.S., Zhao, Y.: Distributed consensus for multiple Euler-Lagrange systems: an event-triggered approach. *Sci. China Technol. Sci.* **59**(1), 33–44 (2016)
6. Zhou, Z., Wang, H.B., Hu, Z.Q., Xue, X.J.: Event-triggered finite-time consensus of multiple Euler Lagrange systems under Markovian switching topologies. *Int. J. Sys. Sci.* **49**(8), 1641–1653 (2018)
7. Xu, Z.W., Su, H.Y., Shi, P., Wu, Z.G.: Asynchronous H_∞ control of semi-Markov jump linear systems. *Appl. Math. Comput.* **349**, 270–280 (2019)
8. Kelly, R., Davila, V.S., Perez, J.A.L.: Control of Robot Manipulators in Joint Space. Springer, Heidelberg (2006)
9. Huang, L.R., Deng, F.Q.: Razumikhin-type theorems on stability of stochastic retarded systems. *Int. J. Syst. Sci.* **40**(1), 73–80 (2009)
10. Yao, X.Y., Ding, H.F., Ge, M.F.: Fully distributed control for task-space formation tracking of nonlinear heterogeneous robotic systems. *Nonlinear Dyn.* **96**(1), 87–105 (2019)



Autonomous Operation of Elevator Buttons for Multi-floor Navigation

Zhen Zhao, Junxin Zhao, and Yunjiang Lou^(✉)

Harbin Institute of Technology Shenzhen, Shenzhen 518000, China
louyj@hit.edu.cn

Abstract. This paper focuses on a service robot equipped with a manipulator to take the elevator autonomously. When the robot has no communication with the elevator, the robot multi-floor navigation needs to detect and operate the unknown elevator button like a human. This paper researches on the elevator button detection and the elevator button operation. First, a button detection algorithm is proposed based on stroke width transform for regions extraction, and then these regions are recognized by the convolutional neural network. Then, this paper proposes an algorithm to plan the operation pose of the manipulator based on the plane of the target button. Finally, the experiments show that the proposed algorithm can detect and operate the unknown elevator panel button.

Keywords: Service robot · Elevator button detection · Elevator button manipulation · Multi-floor navigation

1 Introduction

Service robots have begun to be used in the building environment, and have produced a series of practical applications, such as express, document distribution, and welcome reception. These applications all require robots to navigate across floors in a building. The common method on the market now is that service robots control elevators by transmitting signals. But this method needs to retrofit the elevator. Therefore, this paper proposes a method such that the service robot equipped with a manipulator is controlled to press elevator buttons autonomously. This paper studies service robots autonomously pressing elevator buttons.

In the task of service robot autonomously pressing elevator buttons, it can be divided into two parts. First, based on the characteristics of elevator buttons, an elevator button detection algorithm based on the stroke width transform and convolutional neural network is proposed. Second, according to the point cloud information of the target button area, the operation poses planning algorithm of the manipulator operating elevator buttons is proposed. Experiments with the proposed algorithms verify that the manipulator can operate the target button of unknown elevator panels safely and effectively.

2 Related Works

The operation of elevator buttons requires accurate recognition of the target button and spatial positioning. The accuracy of button detection and recognition results will determine whether the operation is effective. Now, elevator button detection algorithms can be divided into three ways according to different processes.

The first way is to use a sliding window to traverse image sub-regions, and identify the image in the window during the traversal process. Troniak D et al. [1] used sliding windows and fast template matching to complete the detection of buttons on the known elevator outer panel. The second way is to extract candidate regions by artificially designed features for recognition. Kim HH et al. [2] proposed the edge feature extraction method of square buttons for known elevator buttons to obtain candidate regions, and then used template matching to identify buttons. Dong Z et al. [3] used the Canny algorithm to obtain candidate regions through button edge features, collected elevator panel images, and trained a convolutional neural network to identify elevator buttons. The third way is the convolutional neural network based on target detection. This method shows good accuracy and generalization when training samples are sufficient. Yang PY et al. [4] collected and labeled a total of 260,560 images of eight kinds of elevator inner and outer panels, used labeled data to train YOLO-v2 target detection network, and post-processed the network results to achieve high accuracy. Zhu D et al. [5] collected 2100 panel images and labeled 21767 elevator buttons, and proposed a detection network based on Faster-RCNN. By adding a character recognition branch, the problem of imbalance of the button samples is solved.

The algorithm proposed in this paper is based on candidate regions combined with a post-recognition method. This algorithm does not need to collect and label elevator button images and has a certain generalization ability.

3 Elevator Button Detection Algorithm

3.1 Segmentation of the Character Area

This section samples elevator panels of different buildings on campus. The sampled elevator buttons conform to the recommended standards of national standard documents. The prompt characters of buttons are located above the moving parts, and they are in sharp contrast with the background. Therefore, the characters can be used to identify and locate the position of elevator buttons.

The characters of elevator buttons usually have similar sizes and the same style. Therefore, the stroke width transform [6] method is used to achieve the division of independent characters. The stroke width transform needs to be carried out at the edge of the image, so it's necessary to extract the edge of the image. This paper uses the Canny algorithm [7] to extract the image edges. Through the Canny algorithm, the edge pixels of the image are extracted and

the gradient is obtained. The stroke width can be extracted from the information. Let a point $p(x_o, y_o)$ be the pixel on the edge in the edge feature image. From this point, the pixel $p(x_o + \Delta d \cos \theta(x, y), y_o + \Delta d \sin \theta(x, y))$ is traversed in the gradient direction. Iterate Δd successively to the next edge point $p(x_e, y_e)$ and then terminate the traversal. If the gradient directions of the edge points $\theta(x_e, y_e)$ and $\theta(x_o, y_o)$ are almost opposite, that is $\theta(x_o, y_o) - \pi/6 \leq \theta(x_e, y_e) \leq \theta(x_o, y_o) + \pi/6$. It is judged that the ray is valid. Calculate the width $s = \|p(x_o, y_o) - p(x_e, y_e)\|$ of the starting point and ending point of the ray, and assign it to each pixel in the traversal process. If a point in the ray has been traversed and assigned, the smaller width value is retained. To deal with the abnormal stroke width caused by the right-angle area in the character, the effective ray in the above traversal process needs to be traversed twice. Use the median width m on each ray to assign a value to each pixel in the ray.

To further extract the candidate regions of the image, it is necessary to connect regions with similar widths. But after setting the stroke width threshold, there are still some wrong rays and broken similar areas as shown in Fig. 1(a). So this paper adds morphological opening and closing operations based on the original stroke width transform. Then the connected area is extracted again. This paper uses a second-stage connected domain labeling algorithm. The criterion for connectivity is that the ratio of the current pixel width to the adjacent pixel width is less than 3. The first stage marks the adjacent regions, and the second stage updates the connected regions according to the inheritance relationship of the labels. According to the geometric characteristics of elevator buttons, the filtering rules of candidate regions are proposed: character width and height usually present a relatively fixed ratio; eliminate areas where the stroke width is too small; eliminate areas that are too large and too small; eliminate the outer regions where there is a semi-enclosed or fully-enclosed relationship area. The filtered candidate area image is shown in Fig. 1(b).

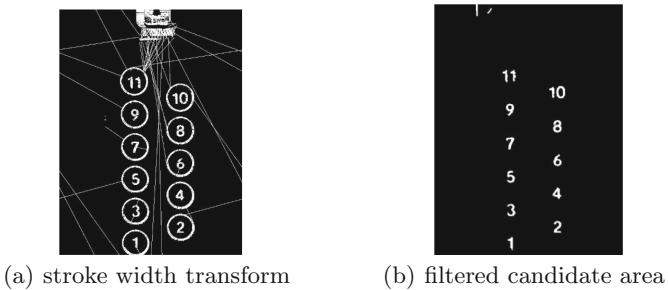


Fig. 1. (a)The image extracted by stroke width transform. (b)The filtered candidate area image. It can be seen that the character area is effectively divided.

3.2 Candidate Region Identification and Post-Processing

Since characters on buttons are usually printed, this section extracts characters from text font, then normalizes and cleans characters. This avoids manual labeling and reduces manpower during data collection. The cleaned data is used for neural network training, and post-processing is proposed to filter results. According to keywords, a total of 3864 character pictures were extracted, of which about 351 were in each category. This paper sets the image size of the network input to 28×28 . After binarizing, perform horizontal and vertical histogram projection on the image, locate the character area, and delete the blank area around the character. After extracting the character area, the image will be scaled by the letterbox scaling method [8] to maintain the aspect ratio of the image. Due to the small image size and the small number of prediction output results, a small convolutional neural network concerning [9] was proposed. The structure of the convolutional neural network is shown in Fig. 2. The loss function is shown in (1). Use multi-label cross-entropy loss and L2 regularization to avoid overfitting. The weight of the L2 regularization is λ_{nor} . In the formula, γ is used to adjust the attention to difficult-to-divide samples, p_i is the true label of prediction category i , and \hat{y}_i is the prediction output. Then the sigmoid function is used to normalize the output result, indicating the predicted probability of the corresponding category.

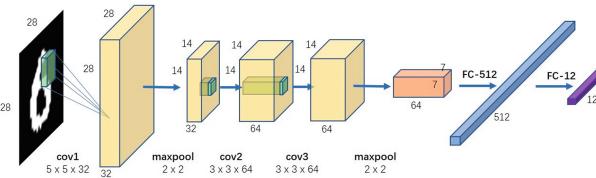


Fig. 2. The structure of the neural network for character recognition.

$$L(p) = - \sum_{i \in \text{classes}} ((1 - \hat{p}_i)^\gamma p_i \log \hat{p}_i + \hat{p}_i^\gamma (1 - p_i) \log (1 - \hat{p}_i)) + \lambda_{\text{nor}} \sum_j w_j^2 \quad (1)$$

$$\hat{p}_i (\hat{y}_i) = 1 / (1 + e^{-\hat{y}_i})$$

Since the data set is obtained from a standard font, the number of image samples in up and down categories is significantly less than that in the digital category. To avoid category imbalance affecting the recognition effect, this section refers to [10] to classify training samples. At the same time, to avoid network fitting abnormal data, the character images were manually cleaned, and the indistinguishable character images were deleted. The network completes convergence at 15,000 steps, and the correct rate obtained through the verification set test is 99.1837%. The results show that the trained network model can accurately recognize characters of each category.

After extracting candidate regions, there is a part of the region that does not contain characters. Such areas are also output predicted probability to form a false-positive result. Each character detected is independent, and adjacent detection results need to be composed to realize the detection of high-floor buttons. Therefore, it's necessary to post-process: filter low-confidence targets; associate adjacent independent characters according to the spatial position relationship; delete low-confidence targets in the same recognition results; filter the results with the highest floor as the threshold. The final result of button detection is shown in Fig. 3.



Fig. 3. The detection results of different elevator panels.

4 Target Button Operation

After character detection, the manipulator needs to obtain operation pose from the depth image according to detection information. To press the button accurately, it's necessary to perform hand-eye calibration to obtain the transformation between camera coordinate and manipulator coordinate. In this paper, the camera is installed on the robot body. It belongs to Eye-to-Hand calibration.

Considering that the moving direction is usually perpendicular to the button plane, this paper proposes a second-stage button operation algorithm. In the first stage, the manipulator adjusts end pose to be perpendicular to target button plane and maintain a certain distance. In the second stage, the manipulator moves along normal vector of the button plane until the end reaches center area of target button. Considering that part of noise in the point cloud data may cause least squares fitting error, this paper uses random sampling consistency algorithm [11] to achieve point cloud plane fitting. After fitting, the plane equation in camera coordinate is $ax_c + by_c + cz_c + d = 0$, and normal vector of the plane is $\mathbf{V} = (a, b, c)^T$. In camera coordinate, the coordinate of center point of target button is ${}^c\mathbf{P}_{ctr}(u_{ctr}, v_{ctr})$, and this point is set as the end point ${}^c\mathbf{P}_{es}$ of second stage operation pose. The end point ${}^c\mathbf{P}_{eo}$ of the operation pose in first stage can be set along the normal vector. Assuming that the distance between

the end point of first stage operation and the button is Δd , the end point of first-stage operation pose can be calculated by (2).

$$\begin{aligned} {}^c\mathbf{V} &= \mathbf{V}/\|\mathbf{V}\| = (a_c, b_c, c_c)^T \\ {}^c\mathbf{P}_{es} &= {}^c\mathbf{P}_{ctr} (u_{ctr}, v_{ctr}) = (x_{ctr}, y_{ctr}, z_{ctr})^T \\ {}^c\mathbf{P}_{eo} &= ((x_{ctr} - a_c \Delta d), (y_{ctr} - b_c \Delta d), (z_{ctr} - c_c \Delta d))^T \end{aligned} \quad (2)$$

There is a transformation relationship ${}^b\mathbf{T}$ between camera coordinate F_c and manipulator base coordinate F_b . Therefore, the end point of second-stage operation and the unit normal vector of target button plane measured by camera coordinate can be transformed into manipulator coordinate, as shown in (3).

$$\begin{aligned} {}^b\mathbf{T} &= \begin{bmatrix} {}^b\mathbf{R} & {}^b\mathbf{t} \\ 0 & 1 \end{bmatrix} \\ {}^b\mathbf{P}_{eo} &= {}^c\mathbf{P}_{eo} + {}^b\mathbf{t} \\ {}^b\mathbf{P}_{es} &= {}^c\mathbf{P}_{es} + {}^b\mathbf{t} \\ {}^b\mathbf{V} &= {}^b\mathbf{R} {}^c\mathbf{V} \end{aligned} \quad (3)$$

In this paper, the end pose of the manipulator is defined by the ZYZ Euler angle. Since the unit normal vector of the button plane only has a constraint on the z-axis, it is sufficient to define the pose with the ZY Euler angle, as is shown in (4).

$$\begin{aligned} {}^b\mathbf{V} &= (a_b, b_b, c_b)^T = \mathbf{R}_{ZY}(0, 0, 1)^T \\ &= \begin{bmatrix} \cos \alpha \cos \beta & -\sin \alpha \cos \alpha \sin \beta \\ \sin \alpha \cos \beta & \cos \alpha \sin \alpha \sin \beta \\ -\sin \beta & 0 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= (\cos \alpha \sin \beta, \sin \alpha \sin \beta, \cos \beta)^T \end{aligned} \quad (4)$$

According to (4), Euler angles α and γ can be solved as (5). Due to the singularity of Euler angles, two sets of solutions can be finally obtained. Select the set of solutions closest to the current end pose as the target pose.

$$\begin{aligned} \alpha &= \arctan 2(b_b, a_b) \\ \beta &= \arccos c_b \\ \arctan 2(y, x) &\triangleq \begin{cases} 2 \arctan \left(\frac{y}{\sqrt{x^2+y^2+x}} \right) & x > 0 \\ 2 \arctan \left(\frac{\sqrt{x^2+y^2}-x}{y} \right) & x \leq 0, y \neq 0 \\ \pi & x < 0, y = 0 \\ \text{undefined} & x = 0, y = 0 \end{cases} \end{aligned} \quad (5)$$

Then the target poses \mathbf{H}_{eo} and \mathbf{H}_{es} of two stages can be obtained, as shown in (6). The manipulator obtains the joint position by inverse kinematics, and linear interpolation is used to obtain the joint trajectory. Finally, the manipulator

moves along the trajectory to realize effective operation.

$$\begin{aligned}\mathbf{H}_{eo} &= \begin{bmatrix} \mathbf{R}_v & \mathbf{P}_{eo} \\ \mathbf{0} & 1 \end{bmatrix} \\ \mathbf{H}_{es} &= \begin{bmatrix} \mathbf{R}_v & \mathbf{P}_{es} \\ \mathbf{0} & 1 \end{bmatrix} \\ \mathbf{R}_v &= \begin{bmatrix} \cos \alpha \cos \beta & -\sin \alpha \cos \alpha \sin \beta \\ \sin \alpha \cos \beta & \cos \alpha & \sin \alpha \sin \beta \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}\end{aligned}\tag{6}$$

5 Experiments

5.1 Elevator Button Detection Experiments

We collected images of different elevator panels in different buildings on campus to test the elevator button detection algorithm. A total of 18 images of the inner and outer elevator panels were collected. Each panel is shot at three angles, with a total of 90 elevator buttons. The experimental results are shown in Table 1 and Fig. 3. In the experiment, the elevator button detection algorithm successfully locates the button positions of different elevators and completes the character recognition on the buttons. Experimental results show that the detection algorithm performs better in frontal viewing angles and is less robust to offset viewing angles.

Table 1. The results of the elevator button detection experiment.

Elevator panel	Number of images	Number of buttons	Number of successful tests	Accuracy	Recall rate	F1-score
The 1st elevator outer panel	3	6	4	1	0.6667	0.80
The 1st elevator inner panel	3	33	32	1	0.9697	0.9846
The 2nd elevator outer panel	3	6	6	1	1	1
The 2nd elevator inner panel	3	15	11	0.84	0.7333	0.7830
The 3rd elevator outer panel	3	6	0	0	0	0
The 3rd elevator inner panel	3	24	17	1	0.70	0.8235
Total	18	90	70	0.8067	0.7778	0.7920

5.2 Target Button Operations Experiments

In this paper, experiments were conducted on the panels inside and outside the elevator. Because the robot has a large moving space in front of the elevator outer panel, the operation experiments are carried out on the left, front, and right sides respectively, and the operation targets are the up and down buttons. In the elevator, since the elevator inner panel is installed on the right side of the elevator door and the space is limited, the robot can only experiment on the left side of the panel. In this experiment, the operation targets of the inner panel of the elevator are buttons from 1 to 9 floors. In this experiment, the manipulator's gripper grabs the foam and touches the elevator button. Table 2 is the result of elevator panel operations in the operation process.

Table 2. The result of the button operation process of the elevator panel.

Position	Button	Test times	First-stage times	Second-stage times	Successful trigger times
Frontal side	Up	3	3	3	2
Frontal side	Down	3	3	3	3
Left side	Up	3	3	3	3
Left side	Down	3	3	3	3
Right side	Up	3	1	1	0
Right side	Down	3	2	2	0
Outer panel total		18	15	15	11
Left side	1	3	3	3	3
Left side	2	3	3	3	1
Left side	3	3	3	3	3
Left side	4	3	1	1	1
Left side	5	3	3	3	3
Left side	6	3	3	3	1
Left side	7	3	3	3	3
Left side	8	3	2	2	1
Left side	9	3	3	3	3
Inner panel total		27	24	24	19

The experimental results show that the button operation algorithm proposed in this paper can realize the effective operation of buttons on the inner and outer elevator panels. During the operation of the inner panel, there may be two reasons for the failure of the operation. One is the instability of the random tree planner used in trajectory planning. The second is the occasional large deviation in the operation end pose. The error may come from the control and positioning

error of the end of the manipulator, the calibration error, and the error of the depth image data. For the button operation on the outer elevator panel, the manipulator can effectively operate in front of the panel and on the left side, but cannot complete the operation when it's on the right side of the panel. The reason is that we only use the right manipulator of the robot for the experiment. When the robot is located on the right side of the elevator panel, the robot body and the wall occupy a large amount of operation space of the manipulator, so that the manipulator cannot be effectively planned to the target button. In practical applications, the outer elevator panel is usually on the right side of the elevator, so this type of failure will not affect the operation of the manipulator.

6 Conclusion

This paper has solved the problem of service robots taking the elevator of multi-floor navigation when the elevator is not retrofitted. We propose an elevator button detection algorithm to detect characters on the button, and propose an operation pose planning algorithm to achieve pressing elevator buttons by the manipulator. Experiments verify that the algorithm can detect buttons of unknown elevator panels. Compared with previous studies, the detection method proposed in this paper does not require the collection and labeling of elevator button images. The region segmentation algorithm based on stroke width transform is interpretable and easy to transplant. By using the generalization of the neural networks, unknown elevator buttons can be detected. After completing the hand-eye calibration of the robot, it was verified that the operation pose planning algorithm can perform safe and effective operations on target buttons of unknown elevator panels.

Acknowledgment. This work was supported partially by the NSFC-Shenzhen Robotics Basic Research Center Program (No. U1713202) and partially by the Shenzhen Science and Technology Program (No. JCYJ20180508152226630).

References

- Troniak, D., et al.: Charlie rides the elevator—integrating vision, navigation and manipulation towards multi-floor robot locomotion. In: Proceedings of 2013 International Conference on Computer and Robot Vision, pp. 1–8. IEEE (2013)
- Kim, H.-H., Kim, D.-J., Park, K.-H.: Robust elevator button recognition in the presence of partial occlusion and clutter by specular reflections. *IEEE Trans. Ind. Electron.* **59**(3), 1597–1611 (2011)
- Dong, Z., Zhu, D., Meng, M.Q.-H.: An autonomous elevator button recognition system based on convolutional neural networks. In: Proceedings of 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 2533–2539. IEEE (2017)
- Yang, P.-Y., Chang, T.-H., Chang, Y.-H., Wu, B.-F.: Intelligent mobile robot controller design for hotel room service with deep learning arm-based elevator manipulator. In: Proceedings of 2018 International Conference on System Science and Engineering (ICSSE), pp. 1–6. IEEE (2018)

5. Zhu, D., Li, T., Ho, D., Zhou, T., Meng, M.Q.-H.: A novel OCR-RCNN for elevator button recognition. In: Proceedings of 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3626–3631. IEEE (2018)
6. Epshtain, B., Ofek, E., Wexler, Y.: Detecting text in natural scenes with stroke width transform. In: Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2963–2970. IEEE (2010)
7. Bao, P., Zhang, L., Xiaolin, W.: Canny edge detection enhancement by scale multiplication. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(9), 1485–1490 (2005)
8. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2016)
9. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014)
10. Zhong, Q., Li, C., Zhang, Y., Sun, H., Yang, S., Xie, D., Pu, S.: Towards good practices for recognition & detection. In: Proceedings of CVPR Workshops, vol. 1 (2016)
11. Chum, O., Matas, J., Kittler, J.: Locally optimized RANSAC. In: Proceedings of Joint Pattern Recognition Symposium, pp. 236–243. Springer (2003)



Optimal Design of a Five-Bar Leg Mechanism for a Quadruped Robot

Chenguang Bai, Guangzeng Chen, Jiangtao Ran, and Yunjiang Lou^(✉)

Harbin Institute of Technology Shenzhen, Shenzhen 518000, China
louyj@hit.edu.cn

Abstract. In this paper, a five-bar leg mechanism is designed and optimized in the context of the quadruped robot. The kinematics and statics of the five-bar leg mechanism is analyzed. The goal of this paper is to design a leg with a load capacity of 30 kg and a walking speed of 0.5 m/s. Under the given conditions, it's necessary to obtain optimal link length parameters so that the maximum torque and angular velocity required for the motors are minimum. And the genetic algorithm is used to obtain the optimal solutions of the overall performances. In addition, comparative study verifies the correctness and effectiveness of the optimization.

Keywords: Quadruped robots · Five-bar mechanism · Optimal design

1 Introduction

For years, quadruped robots have been a research hotspot in the field of mobile robots because they have brilliant mobile performance on unstructured terrains. Quantities of researchers pay attention to the development of quadruped robots. And quite a few quadruped robots, like BigDog [1], ANYmal [2], Cheetah [3], HyQ [4], StarlETH [5], use the animal-like serial mechanism to construct legs. However, quadruped robots with serial legs mechanism tend to have poor load capacity, so that their applications are very limited. This paper propose a quadruped robot leg based on five-bar mechanism. Compared with serial leg mechanism, the parallel five-bar leg mechanism can carry much more weight since there are two motors mounted on the torso to support the body in company. Besides, leg's inertia gets lower because no actuators or sensors fixed on the leg.

The contribution of this paper is the optimization methods for the quadruped robots. Although a great many researchers have put forward lots of mechanism optimization methods [6, 7], few of them are in the context of quadruped robots. This paper propose an original optimal design method to improve the performance of quadruped robots.

This paper aims to reduce the peak torque and peak angular velocity required for the motors under the same circumstances, which is related to the design of

legs. And a multi-objective optimal model is built to achieve the goal of improving the performance of it. The rest of the paper is organized as follows. In Sect. 2, the kinematics and statics of the leg is analysed based on the quadruped robot model presented. In Sect. 3, a multi-objective optimal model is constructed to obtain the link length to have better performances. Section 4 shows the optimal results and verifies the correctness and effectiveness. Finally, the conclusions are drawn in Sect. 5.

2 Kinematics and Statics of the Five-Bar Mechanism

2.1 Inverse Kinematics

The quadruped robot whose leg is composed of five-bar mechanism is presented in Fig. 1(a). It's a planar five-bar structure with two degree of freedoms. The leg is driven by the two motors mounted on the torso.

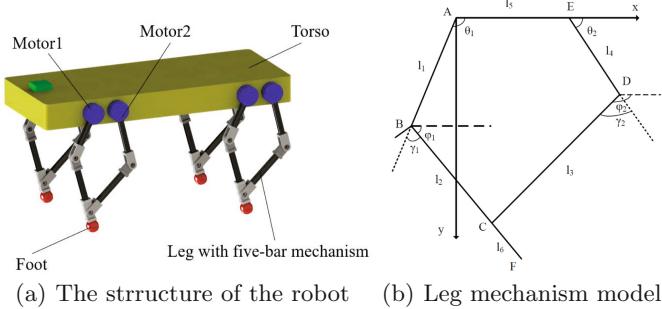


Fig. 1. The model of quadruped robot

Each leg could be modelled as five-bar mechanism in Fig. 1(b). The coordinate's origin is fixed on the center of point A. The positive direction of the x axis points to the center of the body in the horizontal direction. The positive direction of the y axis points to the foot in the vertical direction. Point B, C, D denote passive joints that move along with active joints. Point F, on behalf of the center of ball-shaped foot, is the ultimate objective we want to control.

Based on the characters shown in Fig. 1(b), the horizontal coordinate x_F and vertical coordinate y_F of point F could be obtained:

$$\begin{cases} x_F = l_1 \cos \theta_1 + (l_2 + l_6) \cos \varphi_1 \\ y_F = l_1 \sin \theta_1 + (l_2 + l_6) \sin \varphi_1 \end{cases} \quad (1)$$

$$\begin{cases} x_F = l_5 + l_4 \cos \theta_2 + l_3 \cos \varphi_2 + l_6 \cos \varphi_1 \\ y_F = l_4 \sin \theta_2 + l_3 \sin \varphi_2 + l_6 \sin \varphi_1 \end{cases} \quad (2)$$

After eliminating φ_1 and φ_2 , the equations could be expressed as

$$(l_2 + l_6)^2 + 2l_1 \cos\theta_1 x_F + 2l_1 \sin\theta_1 y_F - l_1^2 - x_F^2 - y_F^2 = 0 \quad (3)$$

$$x_F^2 + y_F^2 - 2(l_5 + l_6 \cos\varphi_1)x_F + 2l_5 l_6 \cos\varphi_1 - 2l_6 y_F \sin\varphi_1 + l_4^2 l_5^2 l_6^2 - l_3^2 + 2(l_4 l_5 - l_4 x_F + l_4 l_6 \cos\varphi_1) \cos\theta_2 + 2(l_4 l_6 \sin\varphi_1 - l_4 y_F) \sin\theta_2 = 0 \quad (4)$$

And θ_1 , θ_2 could be calculated as

$$\tan \frac{\theta_1}{2} = \frac{U + \sqrt{U^2 + V^2 - W^2}}{V - W} \quad (5)$$

$$\tan \frac{\theta_2}{2} = \frac{u + \sqrt{u^2 + v^2 - w^2}}{v - w} \quad (6)$$

where

$$U = 2l_1 y_F, \quad V = 2l_1 x_F, \quad W = (l_2 + l_6)^2 - l_1^2 - x_F^2 - y_F^2.$$

$$u = 2(l_4 l_6 \sin\varphi_1 - l_4 y_F), \quad v = 2(l_4 l_5 - l_4 x_F + l_4 l_6 \cos\varphi_1),$$

$$w = x_F^2 + y_F^2 - 2(l_5 + l_6 \cos\varphi_1)x_F + 2l_5 l_6 \cos\varphi_1 - 2l_6 \sin\varphi_1 y_F - l_3^2 + l_4^2 + l_5^2 + l_6^2.$$

And the expression of φ_1 is

$$\varphi_1 = \arccos \frac{l_1^2 + (l_2 + l_6)^2 - l_{AF}^2}{2l_1(l_2 + l_6)} + \theta_1 - \pi \quad (7)$$

where l_{AF} represents the length between the origin A and the point F.

Hence, given any desired position of point F in the coordinate, corresponding joint angles θ_1 , θ_2 can be determined by (5) and (6).

2.2 Statics Analysis

It's necessary to obtain the mechanism's Jacobian matrix. The derivation of (1) is shown as:

$$\begin{cases} \dot{x}_F = -l_1 \sin\theta_1 \dot{\theta}_1 - (l_2 + l_6) \sin\varphi_1 \dot{\varphi}_1 \\ \dot{y}_F = l_1 \cos\theta_1 \dot{\theta}_1 + (l_2 + l_6) \cos\varphi_1 \dot{\varphi}_1 \end{cases} \quad (8)$$

According to geometry characters of five-bar mechanism shown in Fig. 1(b), it could be got

$$\begin{cases} l_1 \cos\theta_1 + l_2 \cos\varphi_1 = l_3 \cos\varphi_2 + l_4 \cos\theta_2 + l_5 \\ l_1 \sin\theta_1 + l_2 \sin\varphi_1 = l_3 \sin\varphi_2 + l_4 \sin\theta_2 \end{cases} \quad (9)$$

The derivation form of (9) is

$$\begin{cases} l_1 \sin\theta_1 \dot{\theta}_1 + l_2 \sin\varphi_1 \dot{\varphi}_1 = l_3 \sin\varphi_2 \dot{\varphi}_2 + l_4 \sin\theta_2 \dot{\theta}_2 \\ l_1 \cos\theta_1 \dot{\theta}_1 + l_2 \cos\varphi_1 \dot{\varphi}_1 = l_3 \cos\varphi_2 \dot{\varphi}_2 + l_4 \cos\theta_2 \dot{\theta}_2 \end{cases} \quad (10)$$

After eliminating $\dot{\varphi}_2$, (10) could be written as

$$\dot{\varphi}_1 = \frac{l_1 \sin(\varphi_2 - \theta_2) \dot{\theta}_2 - l_1 \sin(\varphi_2 - \theta_1) \dot{\theta}_1}{l_2 \sin(\varphi_2 - \varphi_1)} \quad (11)$$

Bring (11) into (8), we would obtain

$$\begin{bmatrix} \dot{x}_F \\ \dot{y}_F \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} \quad (12)$$

where

$$\begin{cases} J_{11} = \frac{-l_1 \sin \theta_1 + (l_2 + l_6) \sin \varphi_1 l_1 \sin(\varphi_2 - \theta_1)}{l_2 \sin(\varphi_2 - \varphi_1)} \\ J_{12} = \frac{-(l_2 + l_6) \sin \varphi_1 l_4 \sin(\varphi_2 - \theta_2)}{l_2 \sin(\varphi_2 - \varphi_1)} \\ J_{21} = \frac{l_1 \cos \theta_1 - (l_2 + l_6) \cos \varphi_1 l_1 \sin(\varphi_2 - \theta_1)}{l_2 \sin(\varphi_2 - \varphi_1)} \\ J_{22} = \frac{(l_2 + l_6) \cos \varphi_1 l_4 \sin(\varphi_2 - \theta_2)}{l_2 \sin(\varphi_2 - \varphi_1)} \end{cases} \quad (13)$$

$J_{11}, J_{12}, J_{21}, J_{22}$ constitute the complete Jacobian matrix J . F_x is horizontal friction force and F_y is used to support the mass of the whole robot. τ_{F1}, τ_{F2} denote the torque required for the motors due to F_x and F_y . They could be calculated by (14):

$$\begin{bmatrix} \tau_{F1} \\ \tau_{F2} \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}^T \begin{bmatrix} F_x \\ F_y \end{bmatrix} \quad (14)$$

As the quadruped robot leg, the five-bar mechanism is going to move as designed. When the leg run in a gait cycle, the torque τ required for the motors could be expressed as:

$$\tau = \tau_F + \tau_{swing} \quad (15)$$

τ_F denotes the torque due to friction force F_x and support force F_y on the foot and it could be calculated by (14). τ_{swing} means the torque of legs swing and it's expressed as:

$$\tau_{swing} = D(m)\ddot{\theta} + C(m)\dot{\theta} + G(m) \quad (16)$$

where $\dot{\theta}$ and $\ddot{\theta}$ are the angular velocity and angular acceleration required for the motors. m is the mass of a leg. $D(m)\ddot{\theta}$ means the torque due to inertia and $C(m)\dot{\theta}$ means the torque due to Coriolis force. $G(m)$ represents the torque due to gravity. And the quadruped robot leg is so light compared to the whole robot that it can be ignored in comparison:

$$m \ll M$$

where M is the mass of a total quadruped robot. And

$$\begin{cases} \tau_{swing} \propto m \\ \tau_F \propto M \end{cases}$$

According to (15), an approximation could be got

$$\tau \approx \tau_F \quad (17)$$

And this approximation will be verified in the following section.

3 Model of Optimization

The goal of this paper is to design a leg with a load capacity of 30 kg and a walking speed of 0.5 m/s. Under the given conditions, it's necessary to obtain optimal link length parameters so that the maximum torque and angular velocity required for the motors are minimum. So, the link length parameters, $l_1, l_2, l_3, l_4, l_5, l_6$, are set as optimization variables:

$$x = (l_1, l_2, l_3, l_4, l_5, l_6) \quad (18)$$

In order to obtain the optimal link length parameters, the functions between peak torque, peak angular velocity required for the robot and link length parameters will be established. And a multi-objective optimization model for link length parameters is built to achieve the goal of improving the performance of quadruped robots.

3.1 Constraints of Optimization

Taking into account the sizes of the components and the assembly requirements, each link's length has a boundary constraint to avoid being too long or too short. So the boundary constraints are established as:

$$l_{imin} \leq l_i \leq l_{imax} (i = 1, 2, 3, 4, 5, 6) \quad (19)$$

where l_i denotes each link's length and l_{imin}, l_{imax} denote its lower bound and upper bound. And it's shown in Table 1:

Table 1. Constraints of link length

Items	l_1	l_2	l_3	l_4	l_5	l_6
Lower bound	200 mm	200 mm	200 mm	200 mm	100 mm	80 mm
Upper bound	350 mm	350 mm	350 mm	350 mm	120 mm	150 mm

To guarantee good force transmission performance of the five-bar mechanism, transmission angle in mechanism is introduced. If the transmission angle gets larger, the effective component force gets larger. Therefore, it should be ensured that the minimum transmission angle γ_{imin} is greater than the allowable transmission angle $[\gamma]$:

$$[\gamma] \leq \gamma_{imin} (i = 1, 2) \quad (20)$$

Here, γ_{imin} is set as 40° .

3.2 Sub-objective Function

(i) Torque performance

Under the same circumstance, the torque required for the motors should be kept as small as possible, which reduces the cost and makes it easy to select appropriate motors.

$$\begin{cases} f_1(x) = \min \tau_{1max} \\ f_2(x) = \min \tau_{2max} \end{cases} \quad (21)$$

The torque τ required for the robot could be calculated by (17).

(ii) Velocity performance

Besides, the general motors can not achieve the output of both high velocity and great torque. So, the angular speed required by the motors should be kept as small as possible:

$$\begin{cases} f_3(x) = \min \dot{\theta}_{1max} \\ f_4(x) = \min \dot{\theta}_{2max} \end{cases} \quad (22)$$

Given any desired trajectories of the quadruped robot's foot point F, the corresponding joint angle of the motors can be calculated by (5), (6) respectively. In this way, the angular velocity of the motors can be obtained easily.

(iii) Total length of links

In addition, the total length of the leg mechanism should be remained as small as possible. So that it can reduce not only the mass of leg, but also the space occupied by the mechanism:

$$f_5(x) = \min (l_1 + l_2 + l_3 + l_4 + l_5 + l_6) \quad (23)$$

3.3 Normalization and Weight Assignment

Based on the constraints and sub-objective functions above, an optimization model can be obtained:

$$\begin{aligned} \min F(x) &= \min [f_1(x), f_2(x), f_3(x), f_4(x), f_5(x)] \\ \text{s.t. } &200 \text{ mm} \leq l_1 \leq 350 \text{ mm} \\ &200 \text{ mm} \leq l_2 \leq 350 \text{ mm} \\ &200 \text{ mm} \leq l_3 \leq 350 \text{ mm} \\ &200 \text{ mm} \leq l_4 \leq 350 \text{ mm} \\ &100 \text{ mm} \leq l_5 \leq 120 \text{ mm} \\ &80 \text{ mm} \leq l_6 \leq 150 \text{ mm} \\ &40^\circ \leq \gamma_1 \text{ min} \\ &40^\circ \leq \gamma_2 \text{ min} \end{aligned}$$

As each objective function has different dimensions, we need to make them normalization:

$$f_i^*(x) = \frac{f_i(x) - \min f_i(x)}{\max f_i(x) - \min f_i(x)} \quad (24)$$

Among them, $\max f_i(x)$ is the maximum value of the objective function $f_i(x)$ and $\min f_i(x)$ is its minimum value. In order to measure the relative importance of different sub-objective functions, each sub-objective function should be given a corresponding weight W_i :

$$\min F^*(x) = \sum_{i=1}^5 W_i f_i^*(x) \quad (25)$$

In that case, the multi-objective optimization model is transformed into a single-objective optimization model. $F^*(x)$ is the ultimate object to be optimized. We just need to optimize $F^*(x)$ to minimize its value.

4 Optimal Results and Verification

In this paper, genetic algorithm is used to solve the nonlinear optimization problem. The optimization process is shown in the Fig. 2. It can be seen that $F^*(x)$, the value to be optimized, converges to a constant quickly. And the final optimal results are shown in Table 2.

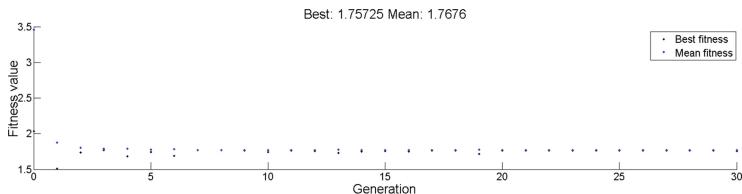
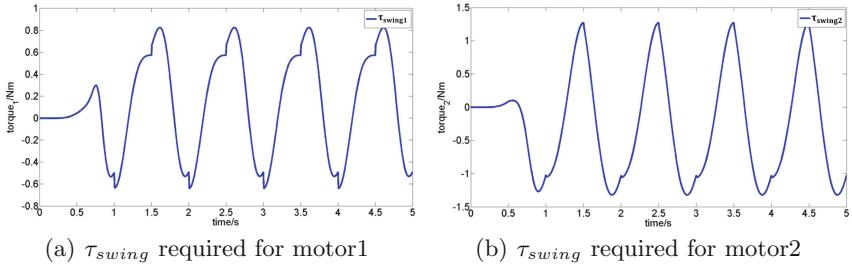


Fig. 2. Genetic algorithm optimization process

Table 2. Optimal results

Items	l_1	l_2	l_3	l_4	l_5	l_6
Optimal results	216 mm	269 mm	286 mm	263 mm	120 mm	100 mm
Items	γ_{1min}	γ_{2min}	$\dot{\theta}_{1max}$	$\dot{\theta}_{2max}$	τ_{1max}	τ_{2max}
Optimal results	40.6°	41.5°	82.47°/s	134.87°/s	44.67 Nm	41.54 Nm

As the results shown in Table 2, all the link length is between the constrained lower bound and upper bound. The minimum transmission angle γ_{1min} is greater than, the allowable one, 40°. In the light of the optimal link size, a model is built in simulation environments. And simulation results in Fig. 3 illustrate that the torque due to τ_{swing} is no more than 1.27Nm, which can be ignored in τ_{imax} . The approximation in (17) proves to be correct.

**Fig. 3.** τ_{swing} required for two motors

And compared with previous design, $l_1 = 200$ mm, $l_2 = 293$ mm, $l_3 = 314$ mm, $l_4 = 263$ mm, $l_5 = 152$ mm, $l_6 = 90$ mm, all indicators have been greatly improved. The specific performance comparison is shown in Table 3. The peak torque of motor1 is reduced by 5.38 Nm, accounting for 10.7% of the torque before optimization. The peak torque of motor2 is reduced by 7.02 Nm, accounting for 14.4% of the torque before optimization. The peak velocity of motor1 is reduced by $18.09^\circ/\text{s}$, accounting for 17.9% of the velocity before optimization. The peak velocity of motor2 is reduced by $6.06^\circ/\text{s}$, accounting for 4.3% of the velocity before optimization.

Table 3. Performance comparison before and after optimization

Items	Before optimization	After optimization
Peak torque required for motor1	50.05 Nm	44.67 Nm
Peak torque required for motor2	48.56 Nm	41.54 Nm
Peak velocity required for motor1	$100.56^\circ/\text{s}$	$82.47^\circ/\text{s}$
Peak velocity required for motor2	$140.93^\circ/\text{s}$	$134.87^\circ/\text{s}$

5 Conclusions

This paper presents the kinematics and statics analysis on the five-bar mechanism leg. On the basis of it, an optimal design method is proposed to design the length of the five-bar mechanism leg to have better performances. And comparative study proves that the peak torque and peak angular velocity required for the motors reduce a lot by the optimization method.

Acknowledgment. This work was supported partially by the NSFC-Shenzhen Robotics Basic Research Center Program (No. U1713202) and partially by the Shenzhen Science and Technology Program (No. JCYJ20180508152226630).

References

1. Raibert, M., Blankespoor, K., Nelson, G., Playter, R.: BigDog, the rough-terrain quadruped robot. IFAC Proc. Vol. **41**(2), 10822–10825 (2008)
2. Hutter, M., Gehring, C., Lauber, A., Gunther, F., Bellicoso, C.D., Tsounis, V., Fankhauser, P., Diethelm, R., Bachmann, S., Blösch, M., et al.: Anymal-toward legged robots for harsh environments. Adv. Robot. **31**(17), 918–931 (2017)
3. Bledt, G., Powell, M.J., Katz, B., Di Carlo, J., Wensing, P.M., Kim, S.: MIT Cheetah 3: design and control of a robust, dynamic quadruped robot. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2245–2252. IEEE (2018)
4. Semini, C., Tsagarakis, N.G., Guglielmino, E., Focchi, M., Cannella, F., Caldwell, D.G.: Design of HyQ—a hydraulically and electrically actuated quadruped robot. Proc. Inst. Mech. Eng. Part I: J. Syst. Control Eng. **225**(6), 831–849 (2011)
5. Hutter, M., Gehring, C., Bloesch, M., Hoepflinger, M.A., Remy, C.D., Siegwart, R.: Starleth: a compliant quadrupedal robot for fast, efficient, and versatile locomotion. In: Adaptive Mobile Robotics, pp. 483–490. World Scientific (2012)
6. Briot, S., Goldsztejn, A.: Topology optimization of industrial robots: application to a five-bar mechanism. Mech. Mach. Theory **120**, 30–56 (2018)
7. McDougall, R., Nokleby, S.: Grashof mechanism synthesis using multi-objective parallel asynchronous particle swarm optimization. In: Proceedings of the Canadian Society for Mechanical Engineering Forum, pp. 7–9 (2010)



Research on Fault Diagnosis Method of Train Wheelset Based on Deep Learning and Big Data Analysis

WanQing OuYang, Jian Zhang^(✉), Rongchao Fang, HengBo Jiang, Yujie Luo, and Xiaoyu Lu

Hunan University of Science and Technology, Xiangtan 411100, Hunan, China
oywq0617@163.com, jzhang@hnust.edu.cn

Abstract. With the continuous upgrading of the intelligent train wheelset system, the number and types of the wheelset monitoring data have gradually increased, the research on the fault diagnosis of train wheelset is moving towards the era of big data. It is a great challenge to collect and sort out the useful fault diagnosis information in the chaotic monitoring data of train wheelset. This paper proposed an improved fault diagnosis method of train wheelset based on the LSTM network and MapReduce framework, which takes advantage of deep learning and big data analysis. This novel method can directly process the raw monitoring data without any preprocessing or traditional feature extraction, also can process the large-scale data quickly and get a higher accuracy of diagnosis results.

Keywords: Fault diagnosis · Big data analysis · Deep learning · Train wheelset

1 Introduction

The wheelset system is a critical part of the train mechanical equipment. It is also one of the most likely fault-occurring part of the running gear. An effective fault diagnosis method could obtain the healthy condition of wheelset and probe the fault patterns, which are also the most challenging tasks in fault diagnosis [8].

The traditional method of fault diagnosis in train wheelset mainly adopts vibration analysis which, based on the vibration characteristics of the wheelset itself, adopts the vibration signal to on-line monitoring and fault diagnosis, and has become the main means in the management and maintenance of train equipment in the past few decades. Then the vibration signals are preprocessed, feature extracted and mode classified to achieve the goal of fault detection. Classical time-frequency domain analysis methods are applied to feature extraction, such as Fast Fourier Transform(FFT) [2], Empirical Mode Decomposition (EMD) [1], etc. Next, the extracted features are fed into classifiers such as a support vector machine (SVM) [9], a BP neural network [17], etc.

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 180–192, 2021.

https://doi.org/10.1007/978-981-15-8458-9_20

However, all these fault methods usually has the following two defects: On the one hand, it does not consider the inherent sequence features behind the monitoring data [12]; On the other hand, it mainly uses single data to diagnose directly without considering the operation monitoring data, historical monitoring data, and maintenance data in train wheelset system as a whole, which leads to the lower efficiency and the inaccurate diagnosis results [13].

Aiming at the shortcomings of the traditional algorithms mentioned above, deep learning is taken into consideration [6]. This method is thought capable of discovering useful high-order feature representations, as well as the relevance of raw signals [5], which better suit the fault diagnosis task and hence lead to improved performance. Recurrent Neural Network (RNN) is a deep learning model for processing sequence data, which have exhibited state-of-the-art performance on a variety of sequence classification and prediction problems. Long Short-Term Memory (LSTM) is a significant branch of RNN, capable of learning long-term dependencies [14, 16], which also proves to be effective in the research field of machine health assessment. In recent years, Big data analysis technology has developed rapidly and has been widely used in various industries. Applying big data analysis to fault diagnosis can not only extract useful information for fault diagnosis from chaotic monitoring data but also effectively integrate datum collected from multiple locations.

This paper takes advantage of both Deep Learning and Big Data Analysis, combined with data acquisition, big data fusion, and other related technologies, proposed an improved fault diagnosis method based on the LSTM network and big data MapReduce framework. This method effectively realizes the timely processing of massive amounts of disordered train wheelset monitoring data, so as to facilitate the timely processing of large-scale data sets and meet the real-time requirements of train wheelset fault diagnosis.

2 Data Acquisition and Fusion

The multi-sensor system is the source of the train wheelset status data. The accuracy of sensor measurement data directly affects the operation of the system and the correctness of analysis and decision-making [11], so the multi-point measurement is usually used to acquire the data. Figure 1 depicts the process of collecting the monitoring data of the train wheelset system. Ten sensors from

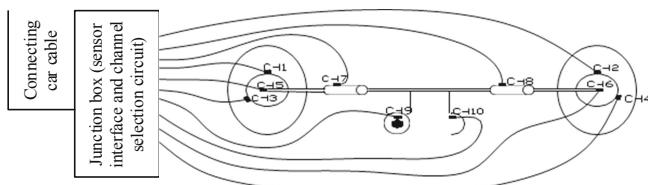


Fig. 1. The schematic diagram of monitoring data acquisition

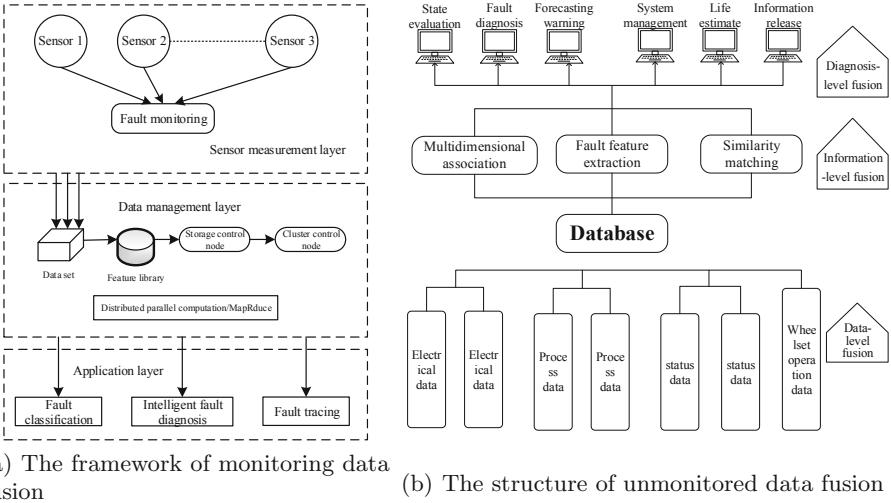


Fig. 2. The performance of LSTM network

CH1 to CH10 are installed at each measuring point of the train wheelset to prepare for the data fusion and fault diagnosis.

Data fusion is to appropriately fuse the sensor monitoring data into useful information for fault diagnosis, which can accurately reflect the running state of train wheelset. According to Fig. 2(a), the framework of the monitoring data fusion has a three-tier structure which can realize the fault diagnosis, fault classification, and fault tracing of train wheelset by sharing, fusing, and integrating the monitoring data of various parts.

The structure of the parts in the train wheelset is complex, and the state data of some parts cannot be directly monitored by sensors. So it is necessary to integrate the other data that can be monitored, and then obtain the state quantity of the parts according to the feature vector. In this paper, an unmonitored data fusion structure of the train wheelset is designed, as shown in Fig. 2(b).

3 Fault Diagnosis Algorithms

The fault diagnosis method of train wheelset proposed in this paper combines the LSTM network with the big data MapReduce framework.

3.1 Long Short-Term Memory Network

As a typical neural network of deep learning, RNN is able to process arbitrary-length sequences data of input patterns at a time, and the neurons of RNN have a recursive structure that can pass the information of the previous state to the current state. In this way, RNN can adaptively model dynamic information of sequence data on multiple scales [15].

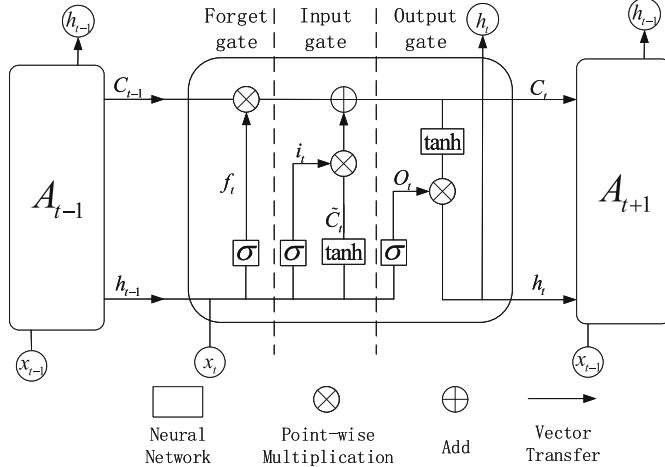


Fig. 3. The architecture of Vanilla LSTM

However, with the memory distance increases, the effect of the simple RNN will decline and cause the problem of gradients vanishing or exploding. Therefore, a LSTM architecture that involves a memory cell was constructed [3]. As the most basic variant of LSTM, Vanilla LSTM has been widely used in the sequential data processing. The architecture of Vanilla LSTM (hereinafter referred to as LSTM) can be described as Fig. 3.

The LSTM has four neural network layers, the cell state is the core of it, which adopt the gates to remove or add information to the neural network, and selectively allow information to pass, thus solving the problem of long-term dependence. The picture clearly illustrates that the cell of LSTM consists of three gates (forget gate, input gate, and output gate) and two flows (cell flow and hidden flow).

The forget gate is used to filter sequences by passing the previous information and removing the others, the output can be calculated as follows:

$$f_t = \sigma(w_{xf}x_t + w_{hf}h_{t-1} + b_f) \quad (1)$$

where σ is a Sigmoid function, and w_{xf} , w_{hf} and b_f are the weights and bias of forget gate, x_t is current time input and h_{t-1} is hidden layer output of the previous moment.

The input gate is used to determine which new information should be added in the cell state, which could be expressed as follows:

$$i_t = \sigma(w_{xi}x_t + w_{hi}h_{t-1} + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(w_{xc}x_t + w_{hc}h_{t-1} + b_c) \quad (3)$$

Where w_{xi} , w_{hi} and b_i are the weights and bias of input gate, \tilde{C}_t is the temporary unit state, w_{xc} , w_{hc} and b_c are the weight and bias of it. Then update the cell flow and obtain the current cell state:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{C}_t \quad (4)$$

The output gate is used to determine the output information of current cell status, the formula is as follows:

$$O_t = \sigma(w_{xo}x_t + w_{ho}h_{t-1} + b_o) \quad (5)$$

Where w_{xo} , w_{ho} and b_o are the weights and bias of input gate. Finally, update the hidden flow and calculate the hidden layer output of the current moment h_t :

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

In addition, the output value of tanh function is between -1 to $+1$, and \odot denotes the point-wise multiplication.

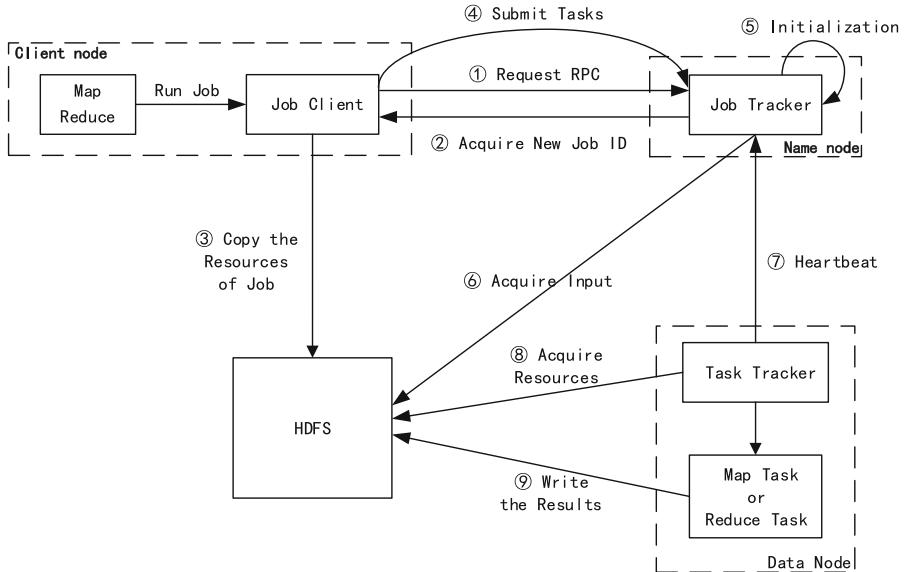


Fig. 4. The data process model of MapReduce

3.2 Big Data MapReduce Framework

The algorithm based on LSTM realizes extracting the fault feature parameters of disordered monitoring data and diagnose the fault of train wheelset. However, when using the algorithm to process a large quantity of the monitoring data, it has the defect of slow convergence speed. That cannot meet the real-time requirements for train wheelset fault diagnosis [4].

Therefore, this paper proposes an improved algorithm that combines the LSTM-based fault diagnosis model with the MapReduce distributed computing

framework to realize parallel processing of large-scale data sets. MapReduce is the core of big data analysis, which provides a simple distributed computing program to process big data by dividing it into small pieces of data to achieve the purpose of analysis and calculation. The data process model of MapReduce is shown in Fig. 4.

As shown in Fig. 4, the MapReduce distributed computing framework is composed of four independent parts: Client, Job Tracker, Task Tracker, and HDFS. In this way, MapReduce can improve the efficiency and speed of the process for large-scale data set. It also detailedly illustrates the 10 steps of MapReduce execution tasks.

3.3 The Method Proposed in This Paper

This paper proposed an improved train wheelset fault diagnosis method based on LSTM network and MapReduce distributed computing framework. As the faults of train wheelset are relatively complex, and it is necessary to consider multiple sensors comprehensively to obtain the fault type accurately. Therefore, the monitoring data of sensors should be integrated and preprocessed before being input into the LSTM neural network.

For LSTM, outputs can be obtain in each time step, or in any time steps. The method proposed in this paper uses the many-to-one type of LSTM, which compute outputs after certain time steps.

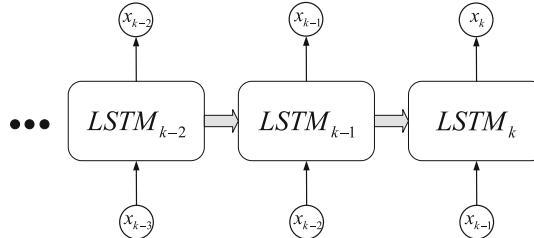


Fig. 5. The architecture of LSTM for fault diagnosis

As can be seen from Fig. 5, the predictive value of the output can be obtained by multi-step recursion with time goes up. For example, the predictive value x_k of the sensors at the time of k is produced based on the time of $k - m$ to the time of $k - 1$ [10]. The formula for the predictive value of output at time k is expressed as

$$x^*(k) = f_{LSTM}[x(k-m), x(k-m+1), \dots, x(k-1)] \quad (7)$$

Where f_{LSTM} is the function of the LSTM cell chain model, from input to output; the value of k is mainly determined according to the actual situation.

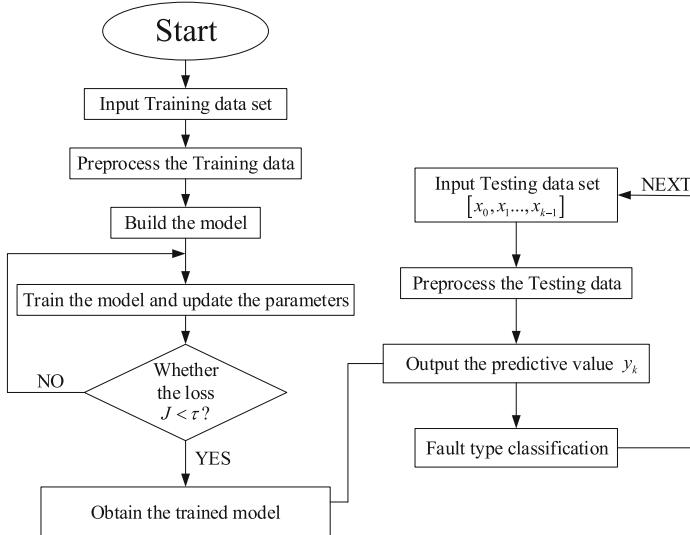


Fig. 6. The flowchart of the LSTM for fault diagnosis

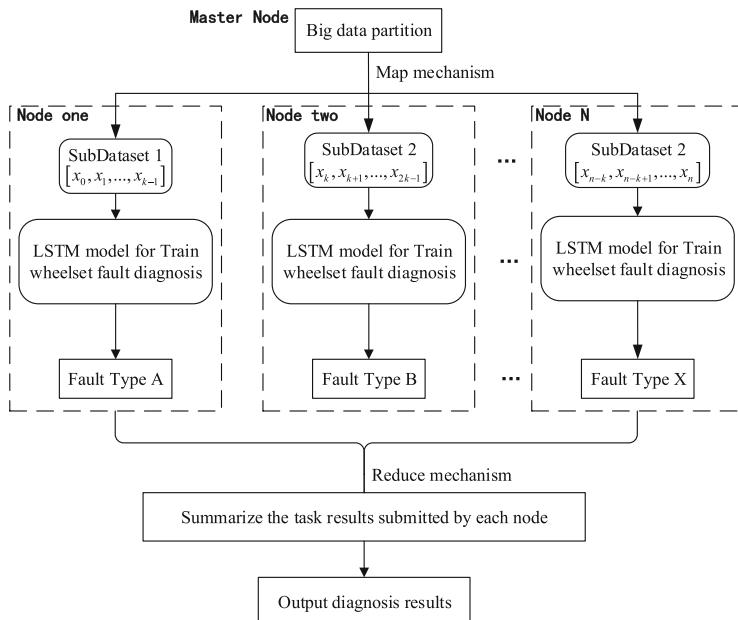


Fig. 7. The flowchart of Big data parallel processing

Besides, behind the LSTM layer, softmax function is added as the activation function of the output layer for fault classification, the formula is as follows:

$$\text{soft max}(y_i) = \frac{e^{u_i}}{\sum e^{u_i}} \quad (8)$$

Where u_i is the i -th output of the former layer.

Before fault diagnosis, the network is trained and the model parameters are continuously optimized during the training. When the loss function of the network is dropped to τ (stable oscillation up and down at τ), it means that the model learning effect is better, and then stopped training. After getting the predictive value of the sensor data fusion at the time of k , it is compared with the true output value to determine the type of failure at that time. The flowchart of the LSTM for fault diagnosis is shown in Fig. 6.

Table 1. Types of sensor monitoring data

Type	Symbol
Traction motor voltage	$U(V)$
Traction motor current	$I(A)$
Train speed	$V(\text{km/h})$
Outdoor temperature	$T_1 (\text{ }^{\circ}\text{C})$
Wheelset horizontal vibration	$F_1(m/\text{s}^2)$
Wheelset vertical vibration	$F_2(m/\text{s}^2)$
Wheelset axial vibration	$F_3(m/\text{s}^2)$

Table 2. Big data cluster software configuration

Configuration	Version
Operating System	Linux CentOS 7
Hadoop	2.2.6-cdh5.7.0
Hive	1.1.0-cdh5.7.0
JDK	1.7.0-79
flume	1.6.0
Hbase	1.2.0-cdh5.7.0
Mysql	5.5.32
Storm	0.9.4

Next, the LSTM-based fault diagnosis network is built in the MapReduce distributed computing framework. Firstly, the large-scale train wheelset monitoring

data is divided into several sub-datasets by period, then the cluster master node Job Tracker assigns the calculation task to each sub-node Task Tracker, in this way, each sub-data set is input into the LSTM fault diagnosis network respectively [7]. The next step is to integrate the diagnosis results of each sub-node, and finally output the fault results. The flowchart of big data parallel processing of train wheelset fault diagnosis in MapReduce distributed computing model is shown in Fig. 7.

According to the task requirements, the framework can set multiple cluster nodes to process the train wheelset monitoring data in parallel, which greatly improves the calculation efficiency, realize the rapid diagnosis, and shorten the fault time.

Table 3. The training parameters of LSTM fault diagnosis network

Parameter	Learning rate	Batch size	Training iteration	Optimizer
Value	0.001	1500	400	AdamOptimizer

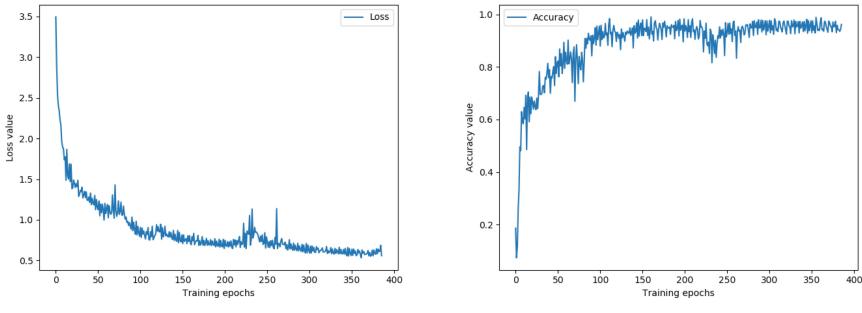


Fig. 8. The performance of LSTM network

4 Experiment Analysis

4.1 Experimental Data and Parameters

In this paper, the historical monitoring data for a certain type of train (which had a slight locomotive axle failure), in June 2019 was collected in the maintenance workshop of the Changsha depot. Then, according to the change characteristics of various monitoring data, fitting and expanding the historical monitoring data, the experimental sample data in this paper are obtained. There are four fault

types of train wheelset: Traction gear fault, Wheelset tread fault, Axle box fault, and Locomotive axle fault. The types of train wheel monitoring data are shown in Table 1.

The software configuration of the big data cluster built in this paper is shown in the following Table 2.

4.2 The Result of Experiment

In order to verify the feasibility and superiority of the algorithm in this paper, the fault diagnosis results of this algorithm and the BP neural network algorithm are compared in the experiment. The training parameters used in the fault diagnosis network based on LSTM are shown in the Table 3.

Table 4. The performance of BP neural network

	Traction gear fault	Wheelset tread fault	Axle box fault	Locomotive axle fault
F1-score	0.76	0.76	0.77	0.77
Precision rate	0.83	0.83	0.84	0.83
Recall rate	0.70	0.71	0.70	0.71

Table 5. RMSE and Testing accuracy of two algorithms

Algorithm type	Standard errors (RMSE)	Testing accuracy
Traditional method	2.6728	0.8170
Method in this paper	1.2536	0.9612

To show the performance of the LSTM network model more intuitively, the Matplotlib library is used to describe the loss decline curve and training accuracy curve of the model training process (shown in Fig. 8).

As shown in Fig. 8, the network loss curve has declined continuously with the increase of the training rounds, the loss value of the network has reached a minimum of 0.41425 when training to about 400 rounds. Although there is a slight vibration around 200 rounds, it tends to be stable until the end of the training. The accuracy curve shows an upward trend as the number of training rounds increases and reached the highest accuracy of 0.9768 when training to about 400 rounds.

Then take the F1-score, precision rate and recall rate as evaluation indicators to comprehensively evaluate the effect of traditional BP neural network

algorithm in train wheelset fault diagnosis. These evaluation indicators can be calculated with the following formula:

$$F1 = 2 \times \frac{PRE \times RE}{PRE + RE} \quad (9)$$

$$precision = \frac{TP}{TP + FP} \quad (10)$$

$$recall = \frac{TP}{TP + FN} \quad (11)$$

Where TP indicates the number of correct diagnosis in the train wheelset fault data; FP indicates the number of wrong diagnosis in the train wheelset fault data; TN represents the number of correct diagnosis in the train wheelset fault-free data; FN represents the number of wrong diagnosis in the train wheelset fault-free data.

The performance of BP neural network in four kinds of fault diagnosis is shown in the Table 4.

To further compare the pros and cons of the two methods, the mean absolute percentage error (MAPE), root mean square error (RMSE), and testing accuracy rate is introduced here to evaluate the performance of the improved LSTM network and BP neural network based on the testing dataset. The formulas are as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{pi} - x_{mi})^2}{n}} \quad (12)$$

$$MAPE = \frac{\sum_{i=1}^n \left(\frac{|X_{pi} - x_{mi}|}{x_{mi}} \right)}{n} \times 100\% \quad (13)$$

Where the x_{mi} is true value (historical data), the X_{pi} is forecasting value, n is the sample scope.

The root mean square errors (RMSE) and the Testing Accuracy Rate of the two methods are shown in the Table 5. The simulation diagram of the MAPE is as shown in Fig. 9(a).

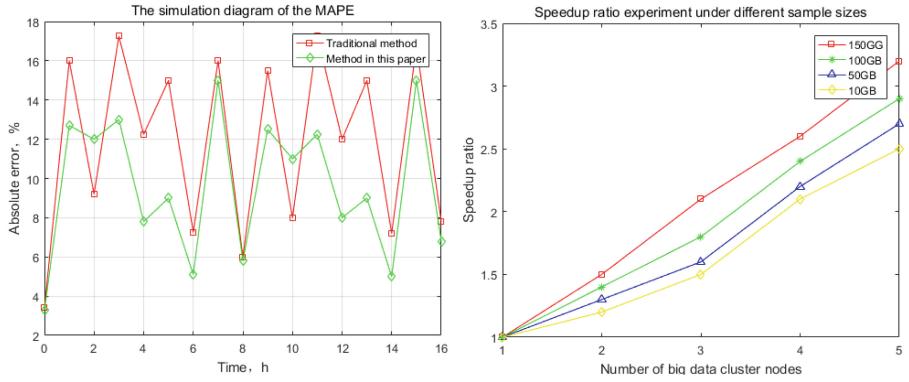
As can be seen from the results, the standard error of this method is less, the MAPE value is lower in different periods, but the testing accuracy rate is higher, which means that the method in this paper is more stable and accurate than the traditional method in the fault diagnosis.

In order to verify the advantages of the fault diagnosis method based on big data analysis proposed in this paper when processing massive train wheelset monitoring data, it is compared with the serial calculation, and the Speedup is introduced as an evaluation index. The formula is as follows:

$$SP = \frac{T_1}{T_P} \quad (14)$$

Where Sp is the speedup, T_1 is the data processing time of single machine, T_P is the data processing time of big data cluster.

By fitting and expanding the train wheelset monitoring data set, it is divided into four groups of data samples with different capacities: 10G (raw data), 50G, 100G, and 150G. The Speedup of four groups of different data capacities is shown in the Fig. 9(b), which indicated that with the increase of big data nodes, the data processing efficiency has been higher, and the number of cluster nodes can be determined according to the amount of monitoring data during train wheelset fault diagnosis.



(a) The simulation diagram of the MAPE (b) The training accuracy curve of LSTM

Fig. 9. The Speed ratio experiment at different sample sizes

To sum up, the method proposed in this paper is more efficient and effective in processing massive monitoring data, and more suitable for fault diagnosis of train wheelset.

5 Conclusion

The fault diagnosis of train wheelset is the core technology to obtain the healthy condition of wheelset and ensure the reliability and safety of the actuation system. This paper proposed an improved fault diagnosis method of train wheelset based on the LSTM network and big data analysis technology. Considering the superiority of LSTM in adaptively modeling of the dynamic information of sequence data and the advantage of MapReduce framework in parallel processing of large-scale data sets, the method this paper proposed can achieve the effect of high efficiency and accuracy in the fault diagnosis of train wheelset.

The feasibility of the proposed method is verified by experiments, the train wheelset fault diagnosis method based on LSTM network and big data has less misdiagnosis, higher fault diagnosis efficiency, and is suitable for processing massive monitoring data.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Grant Nos. 51577057 and 11972157), and the National Defense Foundation Of China (Grant No. 41413050502).

References

1. Do, V.T., Nguyen, L.C.: Adaptive empirical mode decomposition for bearing fault detection. *Strojnicki Vestnik-J. Mech. Eng.* **62**(5), 281–290 (2016)
2. Gan, C., Wu, J., Yang, S., Hu, Y., Cao, W., Si, J.: Fault diagnosis scheme for open-circuit faults in switched reluctance motor drives using fast Fourier transform algorithm with bus current detection. *IET Power Electron.* **9**(1), 20–30 (2016)
3. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
4. Hu, H., Tang, B., Gong, X., Wei, W., Wang, H.: Intelligent fault diagnosis of the high-speed train with big data based on deep neural networks. *IEEE Trans. Ind. Inform.* **13**(4), 2106–2116 (2017)
5. Jia, F., Lei, Y., Lin, J., Zhou, X., Lu, N.: Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech. Syst. Signal Process.* **72**, 303–315 (2016)
6. Lecun, Y., Bengio, Y., Hinton, G.E.: Deep learning. *Nature* **521**(7553), 436–444 (2015)
7. Lei, Y., Jia, F., Lin, J., Xing, S., Ding, S.X.: An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Trans. Ind. Electron.* **63**(5), 3137–3147 (2016)
8. Pan, H., He, X., Tang, S., Meng, F.: An improved bearing fault diagnosis method using one-dimensional CNN and LSTM. *J. Mech. Eng.* **64**, 443–452 (2018)
9. Pan, S., Han, T., Tan, A.C.C., Lin, T.R.: Fault diagnosis system of induction motors based on multiscale entropy and support vector machine with mutual information algorithm. *Shock Vibr.* **2016**, 1–12 (2016)
10. Saeed, H.A., Wang, H., Peng, M., Hussain, A., Nawaz, A.: Online fault monitoring based on deep neural network and sliding window technique. *Prog. Nucl. Energy* **121**, 103236 (2020)
11. Xing, L.: Research on condition monitoring and fault diagnosis of shock absorber in high speed train. Ph.D. thesis, China Academy of Railway Sciences (2018)
12. Xiao, D., Huang, Y., Zhang, X., Shi, H., Li, Y.: Fault diagnosis of asynchronous motors based on LSTM neural network. In: 2018 Prognostics and System Health Management Conference (PHM-Chongqing) (2018)
13. Zhang, J., Fan, X.P., Huang, C.L., Chen, T.F.: Fusion monitoring system of locomotive wheelset state. *J. Traffic Transp. Eng.* **8**(6), 14–19 (2008)
14. Zhang, S., Zhang, S., Wang, B., Habetler, T.G.: Deep learning algorithms for bearing fault diagnostics-a comprehensive review. *IEEE Access* **8**, 29857–29881 (2020)
15. Zhao, H., Sun, S., Jin, B.: Sequential fault diagnosis based on LSTM neural network. *IEEE Access* **6**, 12929–12939 (2018)
16. Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., Gao, R.X.: Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **115**, 213–237 (2019)
17. Zhao, Z., Xu, Q., Jia, M.: Improved shuffled frog leaping algorithm-based BP neural network and its application in bearing early fault diagnosis. *Neural Comput. Appl.* **27**(2), 375–385 (2016)



MEMS Gyroscope Noise Analysis and Calibration Using Allan Variance and Improved Hann Filter

Zhanyi Yan, Senchun Chai^(✉), Liu Yang, Baihai Zhang, and Lingguo Cui

Beijing Institute of Technology, Beijing 100081, People's Republic of China
1104784839@qq.com

Abstract. MEMS gyroscope, as a component of low-cost ins devices, has been widely used in unmanned platforms and intelligent terminals. However, it suffers from interference from deterministic and random errors. In order to solve this problem, a noise analysis and calibration method of MEMS gyroscope based on Allan variance and improved Hann filter is proposed in this paper, and the effectiveness of the method is verified by experiments. The results show that the output is closer to the true value by improved method. After applying this method to the accelerometer as well, the absolute errors of the calculated pitch and roll angles decreased by 0.6825° and 0.3861° , respectively.

Keywords: Improved Hann filter · Allan variance · MEMS gyroscope · Noise analysis · Calibration

1 Introduction

With the development of semiconductor technology and precision manufacturing industry, Micro-Electro-Mechanical System (MEMS) has been widely used in industrial production and daily life. An important application of this technology is MEMS-based inertial sensors. Compared with traditional inertial devices, MEMS inertial devices have the advantages of low cost, low power consumption, and high cost performance [1–3]. These advantages make MEMS devices widely used in unmanned platforms, INS, smart phones and IoT devices [3–6].

MEMS gyroscope is mainly used to measure the dynamic angular velocity information to estimate the attitude and position of the system. However, the physical characteristics of the gyroscope make its performance poor under the conditions of low speed, vibration and noise interference. Therefore, the noise analysis and error calibration of MEMS gyroscope is of great significance.

Generally, the errors of MEMS gyroscopes can be divided into two categories: deterministic errors and random noise [7]. There are two commonly used gyroscope signal noise analysis methods: Power Spectral Density [8] (PSD) and Allan variance methods. The former analyzes the periodic signal by transform in frequency domain. The Allan variance is a method of representing the root means

square (RMS) random-drift error as a function of averaging time [9], which provides a basis for estimating the characteristics of random noise and guides the design of filters based on these quantified indicators.

To suppress the noise of MEMS inertial devices, Zhang Yanshun [10] proposed an adaptive Kalman filtering approach based on the dynamic variance model. And an adaptive multiscale Savitzky-Golay filter (AMSGF) based on adaptive moving average (AMA) was proposed by He Jingjing [11]. In ref. [12], A Kalman filter designed based on the Allan variance identification results is proposed. These methods have achieved certain effects, but most of them require a lot of calculation and analysis, which is not convenient in practical use of embedded systems. Cui [6] proposed a low-pass filter based on Kaiser window function and got good results.

Based on the above analysis, this paper proposes a method of analyzing the noise of MEMS gyroscope through Allan variance, and using the improved Hann window function to design the filter, to achieve the noise suppression and error calibration of MEMS gyroscope.

The structure of this article is as follows: Sect. 2 introduces the error model of MEMS gyroscope, the principle of Allan variance, Hann window function filter and calibration method. Section 3 provides experimental details and results analysis. Section 4 is the conclusion.

2 Method

2.1 Error Model of MEMS Gyroscope

A typical MEMS gyroscope has two main sources of error, which are random and deterministic errors. Random error is a Gaussian noise that can removed by the mean. Regardless of temperature effects, the deterministic error can be divided into three parts, constant bias (zero-drift), scale error and non-orthogonal error. In summary, a widely accepted simplified error model [13–15] of MEMS gyroscope can be introduced as follow.

$$\begin{bmatrix} \Omega_x \\ \Omega_y \\ \Omega_z \end{bmatrix} = \begin{bmatrix} \omega_{x_0} \\ \omega_{y_0} \\ \omega_{z_0} \end{bmatrix} + \begin{bmatrix} K_{xx} & S_{xy} & S_{xz} \\ S_{yx} & K_{yy} & S_{yz} \\ S_{zx} & S_{zy} & K_{zz} \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} + \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \end{bmatrix} \quad (1)$$

where Ω_x , Ω_y and Ω_z are the measured value of tri-axial gyroscope. ω_x , ω_y and ω_z are the true angular velocity of gyroscope that we expect. ω_{x_0} , ω_{y_0} and ω_{z_0} are the constant bias of the gyroscope, also called zero-drift. K_{xx} , K_{yy} , K_{zz} on the main diagonal of the matrix, are the scale coefficients. S_{xy} , S_{xz} , S_{yx} , S_{yz} , S_{zx} and S_{zy} represent non-orthogonal errors due to misalignment. ε_x , ε_y and ε_z are the random error which long-term measurement mean is zero.

2.2 Allan Variance

Since David Allan [16] first proposed the phase and frequency instability of oscillators using Allan variance analysis in 1966, it has been gradually applied

to identify the random errors of inertial devices. It reflects the contribution of the five main noise sources to the total static noise by calculating the change in the average frequency difference in adjacent sampling segments in the time domain. And the five main noise are quantization noise, angle random walk, bias instability, rate random walk and rate ramp [9].

Assume that the sampling time with the gyroscope the original output angular rate is τ_0 , and the number of sampling points is N . These sampling points are divided into N clusters, and each cluster contains M (with $M \leq (N - 1)/2$) consecutive sampling points. Therefore, the duration / τ_0 of each cluster is equal to M/τ_0 . Then the average angular velocity $\bar{\Omega}_k$ of each group can be defined as

$$\bar{\Omega}_k = \frac{1}{M} \sum_{n=1}^N \omega_{(k-1)M+i} \quad (2)$$

where ω is the original angular velocity value output by the gyroscope, and the upper bound of k is related to K , $k = 1, 2, \dots, K$.

Thus, the Allan variance in time period τ_0 can be defined as

$$\sigma_{Allan}^2 = \frac{1}{2(K-1)} \sum_{k=1}^{K-1} (\bar{\Omega}_{k+1}(M) - \bar{\Omega}_k(M))^2 \quad (3)$$

2.3 Improved Hann Filter Design

The improved Hann filter proposed in this paper is a FIR low-pass filter based on Hann window function. Compared with the traditional Kalman filter, it not only ensures the effectiveness of signal filtering, but also reduces the computational complexity. Therefore, it is very suitable for application in embedded system.

For aperiodic signals, the truncation effect in the sampling process will lead to signal leakage, which leads to the distortion of filtering results. Spectrum leakage can only be suppressed, not eliminated. However, it can be improved by adding a window function with low sidelobe and relatively concentrated energy. In addition, avoiding signal leakage also includes sidelobe suppression.

Due to the large side lobe attenuation rate of Hann window, better filtering effect can be obtained. And the following Eq. (4) gives the mathematical expression of Hann window.

$$w(n) = \begin{cases} \frac{1}{2} \left[1 - \cos \left(\frac{2\pi(n-1)}{N} \right) \right] & 0 \leq n \leq N \\ 0 & otherwise \end{cases} \quad (4)$$

where N represents the length of the original data sample sequence.

However, considering that MEMS gyroscopes have a wide range of noise in the sampling frequency range, it is necessary to improve Hann window to speed up the drop rate of the filter side lobes. Therefore, in this paper, the exponential coefficients are introduced to improve the Hann window function. The improved expression is shown in Eq. (5).

$$w(n) = \begin{cases} \frac{1}{2}e^{an} \left[1 - b \cos\left(\frac{2\pi(n-1)}{N}\right) \right] & 0 \leq n \leq N \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where a and b are adjustable parameters. a characterizes the decay rate, b is used for further adjustment.

The time domain and frequency domain characteristics of the original and improved Hann window function are shown in Fig. 1.

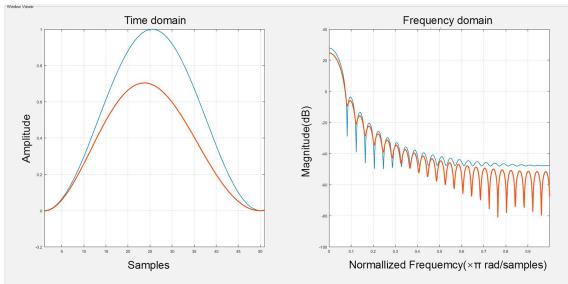


Fig. 1. The time-domain and frequency-domain characteristic of the original (blue line) and improved Hann window

2.4 Calibration Method

Due to the physical characteristics of MEMS gyroscope, the measurement error is large at low speed. The commonly used calibration method is mainly 6-DOF calibration method. This method takes the rotation speed of the turntable as the true value, and the angular velocity sensed by gyroscope is taken as the measurement value. This method has some difficulties in measurement and calculation, and requires high experimental conditions. Therefore, this paper proposes a method to obtain MEMS gyroscope data at a specific angular velocity, and directly uses the angular velocity for least square fitting to obtain the error model.

The above Eq. (1) can be expressed as follows

$$\omega_m = K_g \omega_t + B_g \quad (6)$$

where ω_m is the output measurement, ω_t is the expected true value. And K_g is a three-dimensional square matrix, representing scale coefficients and non-orthogonal errors. B_g is a three-dimensional column vector that represents the constant bias.

Let $[K_g | B_g]$, hence, the above formula (6) can be expressed in augmented matrix form as

$$\omega_m = P_g \omega_t I = P_g \Omega_t \quad (7)$$

$$\Omega_t = \left[\frac{\omega_t}{I} \right] \quad (8)$$

Where I is a row vector with the same dimensions as the matrix ω_t .

Therefore, the calibration process of the gyroscopes original output data is actually to obtain the augmentation matrix P_g . By Least squares method, P_g can be solved as

$$P_g = \omega_m \Omega_t^T (\omega_m \Omega_t^T)^{-1} \quad (9)$$

3 Experiment and Result

3.1 Hardware Platform

The experimental platform is a self-developed flight control board. The board is equipped with MEMS accelerometer, gyros, a magnetometer, RTK, temperature and static pressure sensor and wireless communication unit. The Data processing unit is FPGA, and the MCU is STM32. The physical photo of flight control board is shown in Fig. 2.

Gyroscope dynamic calibration requires turntable. The experimental turntable is a 2-DOF turntable, and both axes can realize 360° rotation around the rotation axis. The experimental turntable and control software interface are shown in Fig. 3 and Fig. 4 respectively.

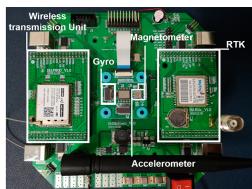


Fig. 2. Flight control board



Fig. 3. Experimental turntable



Fig. 4. Control software interface

3.2 Experiment Scheme and Result Analysis

The flight control board was placed on the main rotation axis of the turntable as shown in Fig. 3. In this paper, the flight control board had three attitudes of X, Y, and Z axes vertically downward, respectively, with the main axis of the turntable at rate of 0°/s, ±1°/s, ±2.5°/s, ±5° s, ±10°/s, ±16°/s, ±25°/s, ±36°/s, ±45°/s. Substituting the collected data into Eqs. (7)–(9), a coefficient matrix can be obtained. Then the Allan variance and FFT method were used to analyze signal characteristics by MATLAB. Next, the improved Hann filter is determined based on the previous analysis results. Finally, the filter and calibration method were applied to the raw data processing.

Figure 5 represents the spectrum after fast Fourier transform of raw data. Obviously, it is easily to be disturbed by noise at low rate. In addition, Fig. 5 also indicates that the cut-off frequency of the filter can be 0.1–1 Hz.

As illustrated in Fig. 6, the Allan Varian is used to analyze the gyroscope noise. And the five main error coefficients are shown in Table 1. It is worth noting that the parameters of the Allan variance fitted by least squares are more suitable as a reference, rather than being calculated as exact values, especially for the MEMS gyroscope, the reliability of the last two coefficients of its Allan variance is poor.

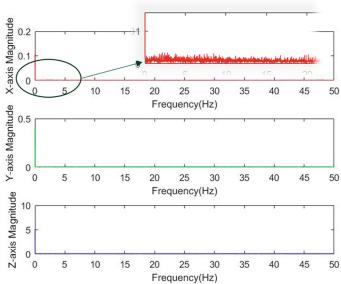


Fig. 5. Spectrogram of triaxial raw data

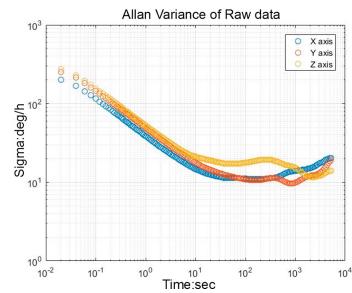


Fig. 6. Allan variance of triaxial raw data

Table 1. Coefficient of Allan variance for gyroscope

Parameters	X-axis	Y-axis	Z-axis
Quantization Noise (deg)	0.9108	1.1805	1.0825
Angle Random Walk ($\text{deg}/h^{1/2}$)	0.6702	0.8922	0.9873
Bias Instability (deg/h)	1.2525	1.5033	5.9373
Rate Random Walk ($\text{deg}/h^{3/2}$)	52.7397	36.8411	65.5568
Rate Ramp (deg/h^2)	19.4150	11.6546	42.4865

Similar to the traditional six-position calibration method for accelerometer [17], this paper proposes a new calibration algorithm. From all the raw data collected, 12 sets of data are randomly selected and brought into Eqs. (6)–(9) for calculation, and calibration coefficients are illustrated in Table 2.

Table 2. Coefficient of Allan variance for gyroscope

Coefficients	Scale factor		Non-orthogonal error						Bias			
Parameter	K_{xx}	K_{yy}	K_{zz}	S_{xy}	S_{xz}	S_{yx}	S_{yz}	S_{zx}	S_{zy}	ω_x	ω_y	ω_z
Value	1.0121	1.0202	1.0104	0.0091	0.0018	-0.0108	0.0089	0.0006	0.0081	0.1421	0.4761	0.1298

After introducing the previous calibration coefficients, the output data of the gyroscope is compared with that before the calibration as shown in Fig. 7. In the case where the turntable rotates around the Y-axis at $-5^\circ/\text{s}$, the average value of the three-axis output of the gyroscope before calibration is $0.2016^\circ/\text{s}$, $-4.3760^\circ/\text{s}$, $0.1630^\circ/\text{s}$. The corrected triaxial angular rate is $0.0162^\circ/\text{s}$, $-4.9506^\circ/\text{s}$, $-0.0054^\circ/\text{s}$. Obviously, the calibration method reduces the error of the gyroscope.

The signal in Fig. 7 has obviously glitches and oscillations. Therefore, it is necessary to use a reasonable filter for signal processing. The improved Hann filter was proposed to perform low-pass filtering on the gyroscope's raw data. In Fig. 8, the FIR low-pass filter and Hann filter are compared with improved Hann filter. The experiment collected 20,000-second static data at a sampling frequency 50 Hz. And results show that improved Hann filter can effectively reduce the chatter and glitch and improve the smoothness.

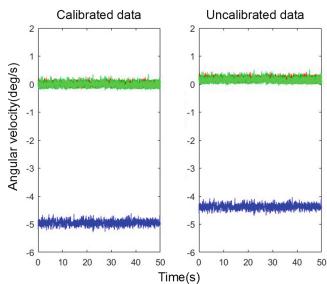


Fig. 7. The calibrated and uncalibrated data of gyroscope

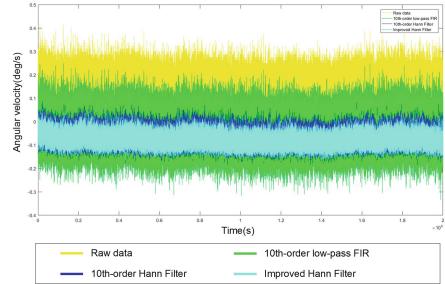


Fig. 8. Comparison of static data filtering effect

Figure 9 shows the Allan variance by FIR low-pass filtering and improved Hann filtering. When using Allan variance to analyze filtered data, the analysis results reflect the performance of the filter. It can be seen from Fig. 11 that the data processed by the improved Hann filter has smaller variance and lower fluctuation degree. The fitting results show that the random walk noise of triaxial angle after FIR filter and improved Hann filter are $[0.7121 \ 0.9123 \ 0.9243]^T$ and $[0.0.6898 \ 0.8819 \ 0.8952]^T$, respectively. And the Bias Instability coefficients are $[2.0948 \ 2.6116 \ 7.4571]^T$ and $[3.4234 \ 4.4024 \ 9.0349]^T$, respectively. The comparison of these two sets of parameters and Table 2 shows that improved Hann filter can suppress the random noise of the sensor to some extent, but it may exacerbate the bias instability.

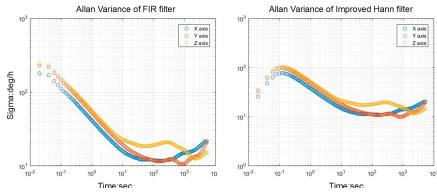


Fig. 9. Allan variance of FIR and improved Hann Filter

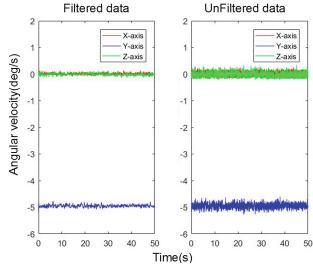


Fig. 10. Comparison of dynamic data filtering effect

The dynamic performance of MEMS gyroscope is also one of the points that need to be focused on in practical applications. In Fig. 10, we use the data on the left side of Fig. 7 to compare the effect of adding the improved Hann filter. The results show that the spikes and glitches of the filtered data are significantly reduced, and the overall signal is smoother. In addition, Fig. 11 illustrates that the improved Hann filter has a small and acceptable time delay.

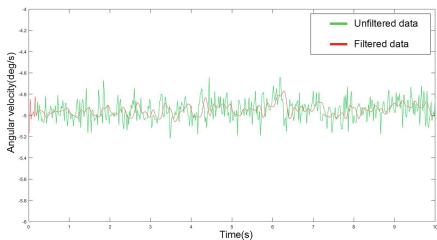


Fig. 11. Time delay caused by filter

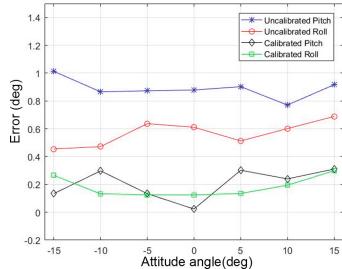


Fig. 12. Pitch and roll errors

The MEMS inertial device studied in this paper is equipped on a quadrotor drone flight control system, self-developed by our laboratory. Its main application scenario is the estimation of attitude and position of drones. In this paper, the improved Hann filter is also applied to the processing of accelerometer data. The six-axis fusion Mahony complementary filtering method [18] is used to estimate the attitude angle. Figure 12 shows the comparison of pitch and roll angle errors before and after using improved Hann filter.

The average of unfiltered pitch angle error is 0.8885° , and roll angle error is 0.5683° . For comparison, the average values of the filtered pitch and roll angle errors are 0.2060° and 0.1822° , respectively. At least 67% reduction in pitch and roll angle estimates.

4 Conclusion

A noise analysis method of MEMS gyroscope based on FFT and Allan variance analysis is proposed in this paper. According to the analysis results, an improved Hann window function filter is designed. Both dynamic and static experimental results indicate that, this method can effectively reduce the spikes and chattering of the raw data, improve the smoothness of the results, and reduce errors. In the attitude angle estimation experiment, the average errors of the pitch and roll angles are reduced by 0.6825° and 0.3861° , respectively. Compared with the traditional method, the error of the MEMS inertial device is greatly reduced. In addition, the calibration and filtering methods proposed in this paper have the advantages of simple design and small amount of calculation. All these advantages show that this method is an effective calibration method for MEMS gyroscope.

References

1. Xue, L., Jiang, C.-Y., Chang, H.-L.: A novel Kalman Filter for combining outputs of MEMS gyroscope array. *Measurement* **45**, 745–754 (2012)
2. Nadeau, A., Dinesh, K., Sharma, G., et al.: In-situ calibration of accelerometers in body-worn sensors using quiescent gravity. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2192–2196. IEEE (2017)
3. Eldesoky, A., Kamel, A.M., Elhabiby, M., et al.: Performance enhancement of low-cost MEMS inertial sensors using extensive calibration technique. In: 2017 34th National Radio Science Conference (NRSC), pp. 415–424. IEEE (2017)
4. Jiang, X., Zong, Y., Wang, X.: Precision improving solutions based on ARMA model and modified self-adapted Kalman filter for MEMS Gyro. In: 2020 SPIE CCC (2010)
5. Wang, L., Wang, F.: Intelligent calibration method of low cost MEMS inertial measurement unit for an FPGA-based navigation system. *Int. J. Intell. Eng. Syst.* **4**(2), 32–41 (2011)
6. SiTime Corporation: SiTime and Bosch accelerate innovation in MEMS timing for 5G and IoT. *J. Eng.* (2018)
7. Naseri, H., Homaeinezhad, M.R.: Improving measurement quality of a MEMS-based gyro-free inertial navigation system. *Sens. Actuators A: Phys.* **207**, 10–19 (2014)
8. Zhang, Y., Guo, S., Chen, Q., et al.: Noise identification and analysis in MEMS sensors using an optimized variable step Allan variance. In: 2019 Chinese Control Conference (CCC), pp. 6309–6314. IEEE (2019)
9. El-Sheimy, N., Hou, H., Niu, X.: Analysis and modeling of inertial sensors using Allan variance. *IEEE Trans. Instrum. Meas.* **57**(1), 140–149 (2008)
10. Zhang, Y., Peng, C., Mou, D., Li, M., Quan, W.: An adaptive filtering approach based on the dynamic variance model for reducing MEMS gyroscope random error. *Sensors (Basel, Switzerland)* **18**(11) (2018)
11. He, J., Sun, C., Wang, P.: Noise reduction for MEMS gyroscope signal: a novel method combining ACMP with adaptive multiscale SG filter based on AMA. *Sensors (Basel, Switzerland)* **19**(20) (2019)

12. Yanning, G., et al.: Performance analysis of MEMS gyro and improvement using Kalman filter. In: 2015 34th Chinese Control Conference (CCC), pp. 4789–4794. IEEE (2015)
13. Fang, J., Sun, H., Cao, J., et al.: A novel calibration method of magnetic compass based on ellipsoid fitting. *IEEE Trans. Instrum. Meas.* **60**(6), 2053–2061 (2011)
14. Liu, Y., Xiang, G., Lu, Y., et al.: Calibration of MEMS accelerometer based on kalman filter and the improved six position method. *J. Commun.* **11**(5), 516–522 (2016)
15. Cui, S., et al.: Calibration of MEMS accelerometer using Kaiser filter and the ellipsoid fitting method. In: IEEE 37th Chinese Control Conference (CCC), pp. 4679–4684 (2018)
16. Allan, D.W., Barnes, J.A.: A modified Allan variance with increased oscillator characterization ability. In: Thirty Fifth Frequency Control Symposium, pp. 470–475 (2011)
17. Won, S.P., Golnaraghi, F.: A triaxial accelerometer calibration method using a mathematical model. *IEEE Trans. Instrum. Meas.* **59**(8), 2144–2153 (2010)
18. Euston, M., Coote, P., Mahony, R., et al.: A complementary filter for attitude estimation of a fixed-wing UAV. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 340–345 (2008)



Monocular Dense 3D Reconstruction Algorithm Based on Inverse Depth Filter

Jingyun Duo¹ and Long Zhao^{1,2(✉)}

¹ School of Automation Science and Electrical Engineering, Beihang University,
Beijing 100191, China
flylong@buaa.edu.cn

² Science and Technology on Aircraft Control Laboratory, Beihang University,
Beijing 100191, China

Abstract. In this paper, an inverse depth filter based real-time monocular 3D dense reconstruction algorithm is proposed. We construct the Gaussian-uniformly mixed probability distribution model by fusing the right and wrong observations in inverse depth filter, which effectively solves the “long tail” problem in conventional depth filter. Then a fast smoothing method is adopted to reduce the effects of noises and false observations, which greatly improves the accuracy and robustness of our method. Experiments are designed to evaluate the accuracy and effectiveness of the proposed algorithm, and the results show that our algorithm has good 3D dense reconstruction performance and high computational accuracy.

Keywords: Monocular camera · 3D reconstruction · Inverse depth · De-noising

1 Introduction

With the development of robot technology and the enhancement of computer information processing capability, intelligent robot has become an important part of intelligent manufacturing. While 3D map construction is one of the focus researches for intelligent robot environment perception, which provides important information for navigation, obstacle avoidance and path planning.

Among common vision sensors, RGBD camera can directly measure scene depth, which has been widely used in 3D reconstruction researches [1]. However, RGBD camera is greatly affected by illumination, and the measurement range is limited, which restricts the application in a large number of scenes. Monocular mobile camera estimates pixel depth through multiple view geometry method, which can be used in outdoor and large scenes. So the research on monocular camera dense 3D reconstruction continues to heat up and attracts much attention [2]. However, most of the existing algorithms are based on the strategy of SFM (Structure from Motion) and build accurate maps offline, which

cannot meet the real-time requirement. The difficulties of monocular real-time dense 3D reconstruction are mainly the following two aspects. Firstly, because of the enormous computation burden, the algorithm needs to be highly parallel computation and independently estimate the depth of each pixel to meet the real-time requirement. Secondly, factors such as noises and false observations will affect the accuracy of the algorithm, and how to smooth noises and false observations is another difficulty to be solved in the dense reconstruction of monocular camera. Monocular visual SLAM (Simultaneous Localization and Mapping) [3,4] can construct sparse or semi-dense map in real time, however sparse or semi-dense map can reflect the environment to a certain extent, the applications is limited, and it is difficult to apply this method to obstacle avoidance and path planning. Literature [5] points out that multi-image fusion can effectively reduce the impact of noises, and the above viewpoint is proved by experiments. In literature [6], the cost function is constructed by regularization of the energy function, and multiple depth maps are fused to realize 3D reconstruction. However, the above two algorithms have two disadvantages. Firstly, due to different camera motions and scene textures, it is not possible to confirm how many images are required to realize accurate depth estimation. Secondly, false observations cannot be detected from a large number of depth observations. Literature [7] proposes a probabilistic method to deal with measurement uncertainty, which effectively overcomes the above defects. This method constructs a Gaussian-uniform mixed probability distribution model, and the depth information is estimated by a depth filter based on Bayes formula. However, pixels corresponding to objects which are far away from camera lens often have “long tail” problem in depth filter. In this paper, a real-time dense reconstruction algorithm based on inverse depth estimation is proposed, which combines the bayesian probability model with the convex optimization method to estimate the scene depth map accurately. There are two contributions in this paper. Firstly, an inverse depth filter is introduced to solve the “long tail” problem in conventional depth filter, which greatly improves the accuracy and robustness of depth estimation. Secondly, a fast smoothing method is adopted to reduce the effects of noises and false observations.

This paper is organized as follows: In Sect. 1, the research background and actuality are presented. In Sect. 2, we introduce our proposed depth estimation method. Experimental results are presented in Sect. 3. Finally, we draw some conclusions in Sect. 4.

2 Method

The pixel depth can be solved by multiple view geometry for moving image sequence, however due to the influence of noise, illumination change and image blur, a single measurement may exist a large error or even a wrong calculation. Depth filter fuses multiple measurements can improve the accuracy of depth estimation. However, pixels corresponding to objects which are far away from camera lens often have “long tail” problem in depth filter. In order to effectively solve the above problems, an inverse depth filter based pixel depth estimation

method is introduced in this paper, which greatly improves the accuracy and robustness of depth estimation.

2.1 Pixel Depth Observation

The pixel depth solved by multiple view geometry is shown in Fig. 1. For pixel \mathbf{u} in image I_r , we can find the correspondence pixel \mathbf{u}' in image I_k by epipolar searching. In our method, 5×5 pixel-sized image block around pixel \mathbf{u} is used for template matching, and NCC (Normalized Cross Correlation) [8] is used to measure the correlation between two image blocks. After completing the epipolar

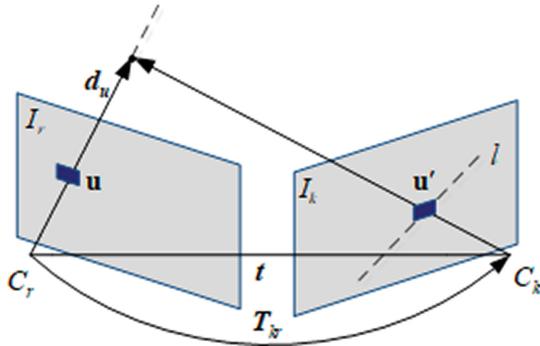


Fig. 1. An illustration of pixel depth solved by multiple view geometry

searching, the depth d_u can be calculated by triangulation. Assuming \mathbf{x}_r and \mathbf{x}_k denote the unit vector of the projection in camera coordinates C_r and C_k respectively, it satisfies the following relationship

$$s_r \mathbf{x}_r = s_k \mathbf{R} \mathbf{x}_k + \mathbf{t} \quad (1)$$

where \mathbf{R} and \mathbf{t} denote the rotation matrix and translation vector form coordinate C_r to C_k , s_r and s_k denote the modulus of depth observations in camera coordinates C_r and C_k respectively. Multiply the left and right sides of Eq. (1) by \mathbf{x}_r^T , we get

$$s_r \mathbf{x}_r^T \mathbf{x}_r = s_k \mathbf{x}_r^T \mathbf{R} \mathbf{x}_k + \mathbf{x}_r^T \mathbf{t} \quad (2)$$

Multiply the left and right sides of Eq. (1) by $(\mathbf{R} \mathbf{x}_k)^T$, we get

$$s_r (\mathbf{R} \mathbf{x}_k)^T \mathbf{x}_r = s_k (\mathbf{R} \mathbf{x}_k)^T \mathbf{R} \mathbf{x}_k + (\mathbf{R} \mathbf{x}_k)^T \mathbf{t} \quad (3)$$

Equations (2) and (3) are set up as

$$\begin{bmatrix} \mathbf{x}_r^T \mathbf{x}_r & -\mathbf{x}_r^T \mathbf{R} \mathbf{x}_k \\ (\mathbf{R} \mathbf{x}_k)^T \mathbf{x}_r & -(\mathbf{R} \mathbf{x}_k)^T \mathbf{R} \mathbf{x}_k \end{bmatrix} \begin{bmatrix} s_r \\ s_k \end{bmatrix} = \begin{bmatrix} \mathbf{x}_r^T \mathbf{t} \\ (\mathbf{R} \mathbf{x}_k)^T \mathbf{t} \end{bmatrix} \quad (4)$$

s_r and s_k can be obtained by solving the Eq. (4), and the depth d_u can be expressed as

$$d_u = [(s_r \mathbf{x}_r) + (s_k \mathbf{R} \mathbf{x}_k + \mathbf{t})]/2 \quad (5)$$

2.2 Uncertainty of Inverse Depth Observation

As shown in Fig. 2, the depth observation of pixel \mathbf{u} along its optical axis is d_u . Define \mathbf{t} denotes the translation vector from coordinate C_r to C_k , we get

$$\mathbf{a} = d_u \cdot \mathbf{t} \quad (6)$$

$$\alpha = \arccos\left(\frac{\mathbf{d}_u \cdot \mathbf{t}}{\|\mathbf{d}_u\| \cdot \|\mathbf{t}\|}\right) \quad (7)$$

$$\beta = \arccos\left(-\frac{\mathbf{a} \cdot \mathbf{t}}{\|\mathbf{a}\| \cdot \|\mathbf{t}\|}\right) \quad (8)$$

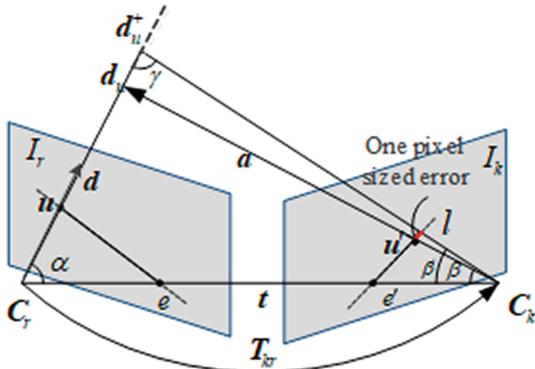


Fig. 2. An illustration of uncertainty of inverse depth observation

Define the focal length of the camera is f , when the corresponding pixel \mathbf{u}' in the image I_k produces one pixel sized error along the epipolar \mathbf{l} , we get

$$\beta^+ = \beta + 2\tan^{-1}\left(\frac{1}{2f}\right) \quad (9)$$

$$\gamma = \pi - \alpha - \beta^+ \quad (10)$$

$$\|\mathbf{d}_u^+\| = \|\mathbf{d}_u\| \frac{\sin\beta^+}{\sin\gamma} \quad (11)$$

the uncertainty of inverse depth observation can be expressed as

$$\tau^2 = (1/\|d_u\| - 1/\|d_u^+\|)^2 \quad (12)$$

2.3 Inverse Depth Estimation

In order to effectively solve the “long tail” problem in depth filter, an inverse depth filter based pixel depth estimation method is introduced in this paper. Assuming $\mathbf{z}_{u_1}, \mathbf{z}_{u_2}, \dots, \mathbf{z}_{u_n}$ are the inverse depth observations of the same pixel \mathbf{u} calculated by n different images, we construct the Gaussian-uniformly mixed probability distribution model by fusing the right and wrong inverse depth observation models. The probability density function of the inverse depth observation of pixel \mathbf{u} caculated by the k -th image can be expressed as

$$p(\mathbf{z}_{u_k} | \mu, \rho) = \rho N(\mathbf{z}_{u_k} | \mu, \tau_k^2) + (1 - \rho) U(\mathbf{z}_{u_k} | \mu_{min}, \mu_{max}) \quad (13)$$

where $N(\mathbf{z}_{u_k} | \mu, \tau_k^2)$ denotes the Gaussian distribution with the inverse depth truth μ as the expectation and the uncertainty of correct inverse depth measurement τ_k^2 as the variance, which is used to construct the probability distribution model of correct inverse depth measurements. $U(\mathbf{z}_{u_k} | \mu_{min}, \mu_{max})$ denotes a uniform distribution in the interval $[\mu_{min}, \mu_{max}]$, which is used to construct the probability distribution model of error inverse depth measurements. ρ and $1 - \rho$ denote the probabilities for correct and wrong measurements respectively. Assuming $\mathbf{z}_{u_1}, \mathbf{z}_{u_2}, \dots, \mathbf{z}_{u_n}$ are independent of each other, from Bayes formula we get

$$p(\mu, \rho | \mathbf{z}_{u_1}, \mathbf{z}_{u_2}, \dots, \mathbf{z}_{u_n}) \propto p(\mu, \rho) \prod_{k=1}^n p(\mathbf{z}_{u_k} | \mu, \rho) \quad (14)$$

where $p(\mu, \rho)$ denotes a prior distribution of the inverse depth truth μ and the correct measurements probability ρ . We update \mathbf{z}_{u_k} by fusing multiple measurements until τ_k^2 converges iteratively.

2.4 Depth Map Smoothing

In our method, GPU (Graphics Processing Unit) is used to estimate the depth of each pixel independently, which can realize efficient parallel computation. However due to noises and false observations, the depth estimations of adjacent pixels may change greatly, which affects the accuracy of the 3D reconstruction. In order to effectively solve the above problems, a de-noising method [9] is adopted to construct an optimization model, which can be expressed as

$$\min_{F(\mathbf{u})} (G(\mathbf{u}) \|\nabla F(\mathbf{u})\|_1 + \lambda \|F(\mathbf{u}) - D(\mathbf{u})\|_1) \quad (15)$$

where $D(\mathbf{u})$ denotes the depth estimation of pixel \mathbf{u} , $F(\mathbf{u})$ denotes the depth of pixel \mathbf{u} after de-noising, $\nabla F(\mathbf{u})$ denotes the gradient of $F(\mathbf{u})$, $\|\cdot\|_1$ denotes 1-norm, λ denotes the regularization parameter, and in this paper, we set $\lambda = 0.2$,

$G(\mathbf{u})$ is the weight function .In formula (15), $\|F(\mathbf{u}) - D(\mathbf{u})\|_1$ is the fidelity item, which is used to maintain the similarity between the de-noised image and the original image, $(G(\mathbf{u}) \|\nabla F(\mathbf{u})\|_1$ is the regularization item, which plays the role of suppressing noise.

3 Experimental Results

In this section, two experiments are designed to evaluate the accuracy and effectiveness of the proposed algorithm. In the first experiment, the image depth estimation test is designed on public dataset, and we compare the proposed algorithm with the algorithm in literature[7] and verify the superiority of our algorithm through quantitative analysis. In the second experiment, the proposed algorithm is applied to 3D reconstruction of indoor scene in order to verify the effectiveness of our algorithm.

3.1 Experiment on Public Dataset

The dataset provided by literature[10] is selected to evaluate the accuracy of the algorithm, which provides the image sequences, the corresponding depth image sequences and camera poses. Our experiment selects two image sequences of “Fast Motion” and “Slow Motion”. In the “Fast Motion” image sequence, the camera moves quickly with a large view-point change, but in the “Slow Motion” image sequence, the camera moves slowly, and the trajectory is smooth. In this experiment, the first image in the image sequence is used as the reference frame. Our algorithm and the algorithm in literature[7] are respectively used to estimate the pixel depth, and the results are shown in Fig. 3 and Fig. 4. The statistical results of depth error are shown in Table 1.

In the depth error image (see Fig. 3 and Fig. 4), the higher the pixel brightness is, the greater the estimation error of the pixel depth value is. So both the algorithm can be used for scene depth estimation, but the depth map obtained by our algorithm is more accurate and smooth. As can be seen from Table 1, for algorithm in literature [7], the root mean square errors of the estimated depth on “Slow Motion” and “Fast Motion” dataset are 0.086998 m and 0.073840 m respectively, while the corresponding results of our algorithm are 0.065820 m

Table 1. Statistical results of depth estimation absolute error

Data file	Algorithm	Root mean square error (m)
Slow Motion	Algorithm in Literature [7]	0.086998
	Our algorithm	0.065820
Fast Motion	Algorithm in Literature [7]	0.073840
	Our algorithm	0.054938

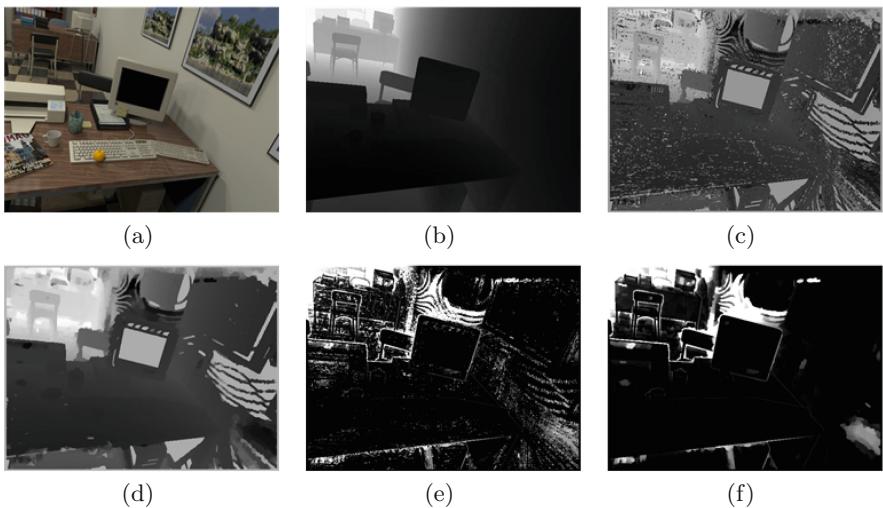


Fig. 3. Experimental results on “Slow Motion” dataset: (a) reference image, (b) groundtruth, (c) depth estimation result of Literature [7], (d) depth estimation result of our algorithm, (e) depth estimation error of Literature[7], (f) depth estimation error of our algorithm

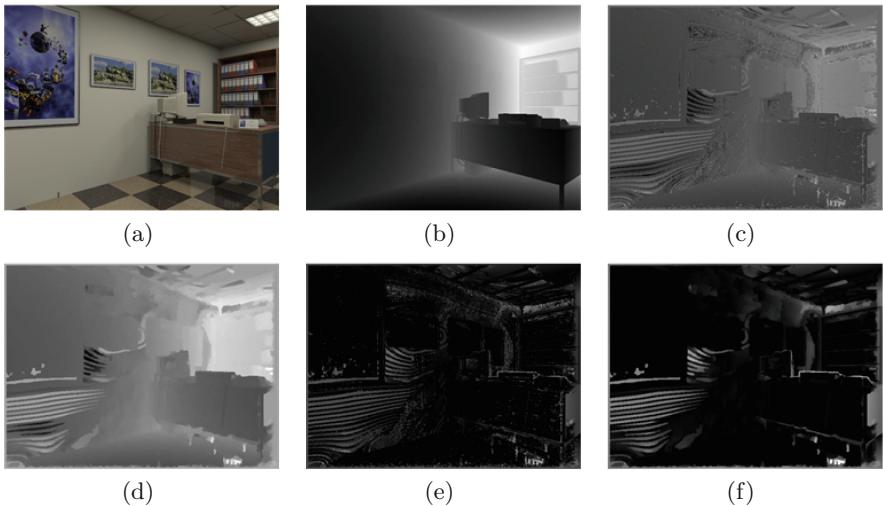


Fig. 4. Experimental results on “Fast Motion” dataset: (a) reference image, (b) groundtruth, (c) depth estimation result of Literature [7], (d) depth estimation result of our algorithm, (e) depth estimation error of Literature[7], (f) depth estimation error of our algorithm

and 0.054938 m respectively. The depth estimation accuracy of our algorithm is obviously better than that in the literature [7], which is mainly due to the inverse depth estimation and depth map smoothing strategy, which effectively deals with pseudo-observation and “long tail” problem in depth estimation.

3.2 3D Reconstruction of Indoor Scene

In order to further verify the validity of the proposed algorithm in real time 3D reconstruction, three small-scale indoor scenes are reconstructed, and the results are shown in Fig. 5.

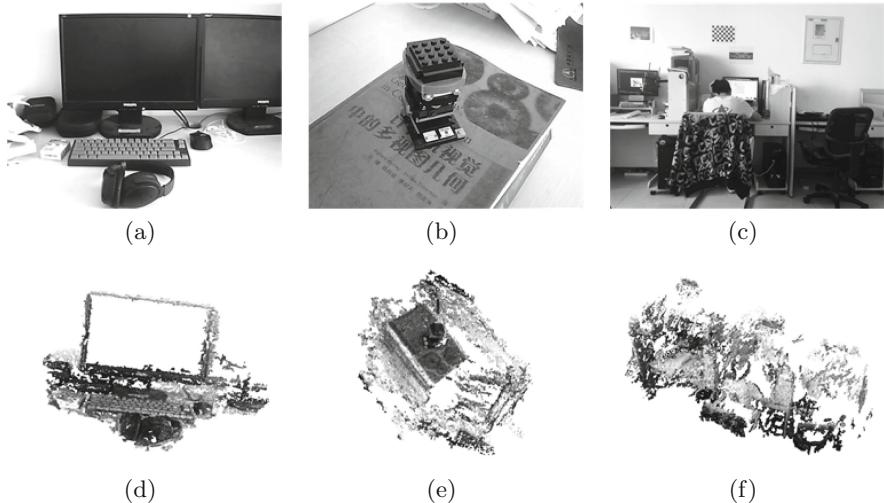


Fig. 5. Experimental results on 3D reconstruction of indoor scene: (a)–(c) three indoor scenes, (d)–(f) results on 3D reconstruction of three indoor scenes

It can be seen from Fig. 5 that our proposed algorithm has the ability of accurate 3D reconstruction and can effectively estimate the depth of the vast majority of pixels in the scene. However, the depth estimations for pixels of the computer monitor and the white wall are not convergence. The problem stems from the fact that the pixel depth measurements are obtained by epipolar searching. If lots of pixel blocks have the same gray information, the epipolar searching will fail, and the failure of estimation is directly caused by the inability to obtain accurate depth measurements.

4 Conclusion

In this paper, a dense reconstruction algorithm based on inverse depth estimation is proposed, which combines the bayesian probability model with the

convex optimization method to estimate the scene depth map accurately. In order to ensure computational efficiency, we adopts GPU to estimate the depth of each pixel in parallel, and then introduces a fast smoothing method to deal with pseudo-observations and observation noises. The experiment results show that the algorithm has high computational accuracy, and can be used for 3D reconstruction in small-scale indoor scenes.

Acknowledgements. This work is supported by the National Natural Science Foundation of China (Grant No. 41874034), the National Science and Technology Major Project of the National Key R&D Program of China (Grant No. 2016YFB0502102), the Beijing Natural Science Foundation (Grant No. 4202041), the Aeronautical Science Foundation of China.

References

1. Shan, Z., Li, R., Sören, S.: RGBD-inertial trajectory estimation and mapping for ground robots. *Sensors* **19**(10), 2251–2279 (2019)
2. Wang, K., Ding, W., Shen, S.: Quadtree-accelerated real-time monocular dense mapping. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2018, Madrid, Spain, 1–5 October 2018
3. Mur-Artal, R., Montiel, J.M.M., Tardos, J.D.: ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Trans. Rob.* **31**(5), 1147–1163 (2015)
4. Engel, J., Schps, T., Cremers, D.: LSD-SLAM: large-scale direct monocular slam. In: 13th European Conference on Computer Vision, ECCV 2014, Zurich, Switzerland, 6–12 September 2014
5. Szeliski, R., Scharstein, D.: Sampling the disparity space image. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(3), 419–425 (2004)
6. Forster, C., Pizzoli, M., Scaramuzza, D.: Air-ground localization and map augmentation using monocular dense reconstruction. In: 26th IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2013, Tokyo, Japan, 3–8 November 2013
7. George, V., Carlos, H.: Video-based, real-time multi-view stereo. *Image Vis. Comput.* **29**(7), 434–441 (2011)
8. Kutulakos, K.N., Seitz, S.M.: A theory of shape by space carving. *Int. J. Comput. Vision* **38**(3), 199–218 (2000)
9. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **40**(1), 120–145 (2011)
10. Handa, A., Newcombe, R.A., Angeli, A., et al.: Real-time camera tracking: when is high frame-rate best? In: 12th European Conference on Computer Vision, ECCV 2012, Florence, Italy, 7–13 October 2012



Multi-scale Superpixel Based Homomorphic Filtering for the Component Image Enhancement

Pingting Song, Jianlin Wang^(✉), Yongqi Guo, Liqiang Zhao, and Kepeng Qiu

Beijing University of Chemical Technology, Beijing 100029, China
wangjl@mail.buct.edu.cn

Abstract. Aiming at the low accuracy of visual measurement caused by the nonuniform illumination image of components during industrial processing and assembly, a multi-scale superpixel based homomorphic filtering method for the component image enhancement is proposed. Firstly, the image is segmented into multiple subblocks using superpixel segmentation in different scales. Secondly, different subblocks are homomorphically filtered at each scale. Finally, images of different scales which have been filtered are weighted according to the standard deviation and fused to get the enhanced image. The experimental results show that the proposed method can effectively improve the average value and average gradient value of the enhanced image.

Keywords: Superpixel segmentation · Homomorphic filtering · Image enhancement · Multi-scale fusion

1 Introduction

Visual measurement [1,2] is widely used in industrial processing and assembly because of its non-contact, high accuracy, and good robustness. Industrial components which are produced and assembled are non-textured, smooth, and reflective. The component images' nonuniform illumination resulting from the lighting characteristics and reflective characteristics causes a low visual measurement accuracy in industrial processing and assembly. Therefore, enhancing the component image is indispensable in processing the optical image of the easily reflective component.

At present, image enhancement methods mainly include histogram equalization method, wavelet transform method, Retinex-based method, homomorphic filtering method, etc. [3].

Histogram equalization method [4–6] expands the dynamic range of gray values between pixels by transforming the original image's histogram into a uniformly distributed form. However, this method adopts a uniform cropping amplitude, which causes some subblocks are not enhanced.

Wavelet transform method [7,8] can enhance the low-frequency part of the image by nonlinear image enhancement, and denoise the high-frequency part by the threshold. However, it reduces the noise interference and the contrast of image detail information leading to the image gray value becomes larger.

Retinex-based method [9,10] enhances the image's original reflectance component by estimating and removing the illumination component of the light source. Funt et al. [11] proposed a Frankle-McCann Retinex method using a spiral structure path for pixel comparison. The spiral path utilizes the global light-dark relationship of the entire image and improves the problem of complex computation in the Retinex-based method based on "path". Nonetheless, there are still many halos at the edges of the enhanced image.

The homomorphic filtering method [12,13] uses a high-pass filter to suppress the illumination component and enhance the reflectance component by converting the image from the spatial domain to the frequency domain. This method does not take account of the color and spatial relationship between pixels at filtering, which leads to an over-enhancement effect; For solving the above problem, Zhang et al. [14] proposed a local homomorphic filtering method which divides the image into subblocks with the same size and performs homomorphic filtering on each subblock. But the blocking method is relatively fixed, which can cause the obvious blocking artifact at the edge of the subblocks, and it needs additional algorithms to deal with the blocking artifact problem.

Regarding to solving these problems, we propose a multi-scale superpixel based homomorphic filtering method for the component image enhancement. Firstly, the image is segmented into multiple subblocks using superpixel segmentation [15,16] with different scales in CIELAB color space. Then, different superpixels are homomorphically filtered at each scale. Finally, images of different scales which have been filtered are weighted according to the standard deviation and fused to get the enhanced image.

2 Local Homomorphic Filtering Method

Based on the illumination-reflectance model, the image $f(x, y)$ can be expressed as

$$f(x, y) = i(x, y)r(x, y) \quad (1)$$

where $i(x, y)$ and $r(x, y)$ are the illumination component and the reflectance component respectively.

The logarithmic function is applied to Eq. (1).

$$\log f(x, y) = \log i(x, y) + \log r(x, y) \quad (2)$$

The Eq. (3) can be obtained using Fourier transform (FFT).

$$F(u, v) = I(u, v) + R(u, v) \quad (3)$$

where $F(u, v)$, $I(u, v)$ and $R(u, v)$ are the Fourier transform of $\log f(x, y)$, $\log i(x, y)$ and $\log r(x, y)$ respectively.

To compress the low frequency and enhance the high frequency, we select a suitable homomorphic filter function $H(u, v)$ to filter $F(u, v)$ as follows.

$$S(u, v) = H(u, v)F(u, v) \quad (4)$$

A homomorphic filter is defined as

$$H(u, v) = (\gamma_H - \gamma_L)[1 - e^{-c(D^2(u, v)/D_0^2)}] + \gamma_L \quad (5)$$

where γ_H and γ_L are the filter coefficients which effect reflectance and illumination respectively, $\gamma_H > 1$ and $\gamma_L < 1$. c is a parameter that controls the slope sharpening of a homomorphic filter function. D_0 is the cut-off frequency.

$D(u, v)$ is the distance between the frequency point and the center point.

$$D(u, v) = \sqrt{(u - M/2)^2 + (v - N/2)^2} \quad (6)$$

where M and N are the rows and columns of the image respectively. $(M/2, N/2)$ is the center of image.

The filtered image is transformed using Inverse Fourier transform (IFFT) as

$$s(x, y) = F^{-1}S\{(u, v)\} \quad (7)$$

The enhanced image can be computed using exponential operation as

$$g(x, y) = \exp(s(x, y)) \quad (8)$$

Figure 1 shows the algorithm flowchart of homomorphic filtering method.

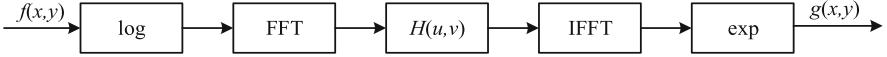


Fig. 1. The algorithm flowchart of homomorphic filtering method

The local homomorphic filtering method divides the image into a series of subblocks. Each subblock is processed by homomorphic filtering, and the processed subblocks are fused to form the enhanced image.

3 Multi-scale Superpixel Based Homomorphic Filtering for the Component Image Enhancement

3.1 Superpixel Segmentation

The image is converted to the CIELAB color space, its adjacent pixels with similar colors, brightness, and texture are clustered. These clustered pixels form the superpixels.

In terms of the approximately equally superpixels, there is a superpixel center at every grid interval S which is defined as

$$S = \sqrt{(M \times N)/K} \quad (9)$$

where K is the superpixel's size and $M \times N$ is the size of the image. $(M \times N)/K$ represents the approximate size of each superpixel.

Each superpixel has its own superpixel cluster center $C_k = [l_k, a_k, b_k, x_k, y_k]^T$, with $k \in [1, K]$. The spatial extent of each superpixel is approximately S^2 . For avoiding the situation that C_k is on the noise point of the pixel edge, K superpixel cluster centers are sampled and moved to 3×3 nearby.

According to Eq. (10), the pixel with minimum gradient value is calculated and it used as the new superpixel cluster center.

$$G(x, y) = \|I(x + 1, y) - I(x - 1, y)\|^2 + \|I(x, y + 1) - I(x, y - 1)\|^2 \quad (10)$$

where $I(x, y)$ is the LAB color vector corresponding to the pixel at (x, y) , and $\|\cdot\|$ represents the L_2 norm.

The transformed Euclidean norm is used to perform the clustering operation on the pixels in the 2×2 area after adjusting the clustering center as follows.

$$\begin{aligned} d_{lab} &= \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \\ d_{xy} &= \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \\ D_s &= d_{lab} + \frac{m}{S} d_{xy} \end{aligned} \quad (11)$$

where D_s is the sum of the LAB color distance and the normalized x, y plane distance. The variable m can control the compactness of superpixels, whose value range is $[1, 20]$. We choose $m = 10$ for all the results in this paper. Figure 2 shows the algorithm flowchart of the superpixel segmentation.

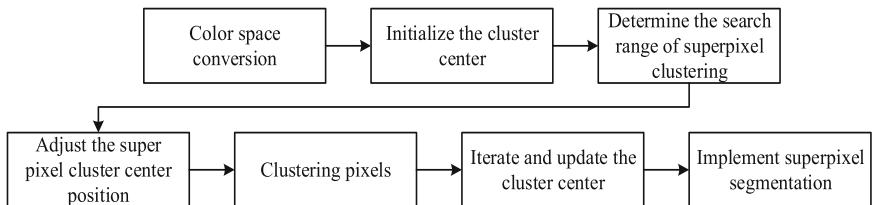


Fig. 2. The algorithm flowchart of superpixel segmentation

We label the pixels which is nearest to the cluster center and then recalculate the cluster center when all pixels belong to their nearest cluster center. The processes of associating pixels and recomputing the cluster center are iteratively repeated until convergence.

3.2 Multi-scale Superpixel Based Homomorphic Filtering for the Component Image Enhancement

In the CIELAB color space, the L color vector represents brightness information, and the number of the crests in its color histogram can represent the color categories' number in the image area. If the initial amount of superpixels is the image area number's multiple, it is possible to compensate for the problem of the same color distributed in multiple areas by shrinking the area proportionally. The initial number of superpixels is defined as

$$k = nq \quad (12)$$

where q represents the number of color histogram's crests belonging to L color vector, and n is the shrinkage ratio of the image area. We use four different contraction ratios to segment the image.

The image is converted to the CIELAB color space and multi-scale superpixel segmentation is performed according to the scale of Eq. (12).

Superpixels are extracted and V vector in the HSV color space is used homomorphic filtering.

Multiple superpixel homomorphic filtered images are obtained. The above images will be weighted according to the standard deviation and fused to obtain the enhanced image.

The standard deviation indicates the degree of dispersion between the image gray value and the average value, which is equivalent to representing the distribution of the image pixel values. The standard deviation is defined as

$$S_{n-1} = \left[\sum_{i=1}^n (x_i - \hat{\mu}) / (n-1) \right]^{1/2} \quad (13)$$

The larger image's standard deviation means the grey degree distributes more equally. The weighted average fusion of multiple images can reflect superpixels' advantages in homomorphic filtering at different scales as Eq. (15) shown.

$$g(x, y) = \sum_{p=1}^m W_p \sum H_{N_p}[i'(x, y)] \quad (14)$$

where $i'(x, y)$ are the superpixels after superpixel segmentation. H_{N_p} are the homomorphic filters for the superpixels under the scale of p . m is the scale in the multi-scale superpixel segmentation. W_p is the weight coefficient with $\sum_{p=1}^m W_p = 1$, it is determined by the ratio of each image's standard deviation to the sum of all images' standard deviations. Figure 3 shows the algorithm flowchart of our method.

4 Experiment

In this section, we compare the experimental results of this method with those of Funt's method [11] and Zhang's method [14] in natural image (Fig. 4(a)),

aerial image (Fig. 4(b)), indoor low-light image (Fig. 4(c)) and component image (Fig. 4(d)) respectively. The average value (AVG) and average gradient value (AVG-GRAD) are used to evaluate the performance of different methods.

All experiments are conducted in MATLAB 2018a on a PC running Windows 10 with 8G RAM and 1.80 GHz CPU.

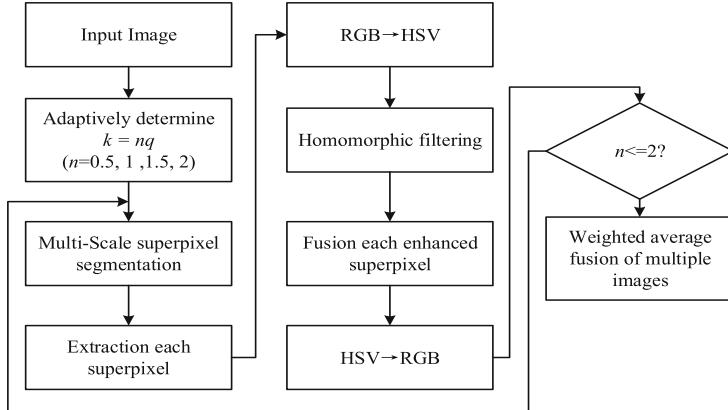


Fig. 3. The algorithm flowchart of our method

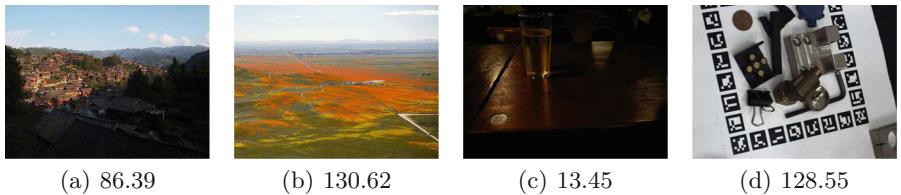


Fig. 4. Four different types of nonuniform illumination images

We use AVG to describe the average brightness of the image. The AVG of an image is defined as

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \quad (15)$$

where n is the total number of pixels, and x_i is the gray value of the i pixel. The larger image's AVG indicates higher image brightness. Figure 4 shows four different types of nonuniform illumination images. We can observe that the bottom of each image is the AVG.

Figure 5 shows the results of superpixel segmentation and homomorphic filtering on the natural image at four different scales. The bottom part shows the shrinkage ratio and the number of superpixels.

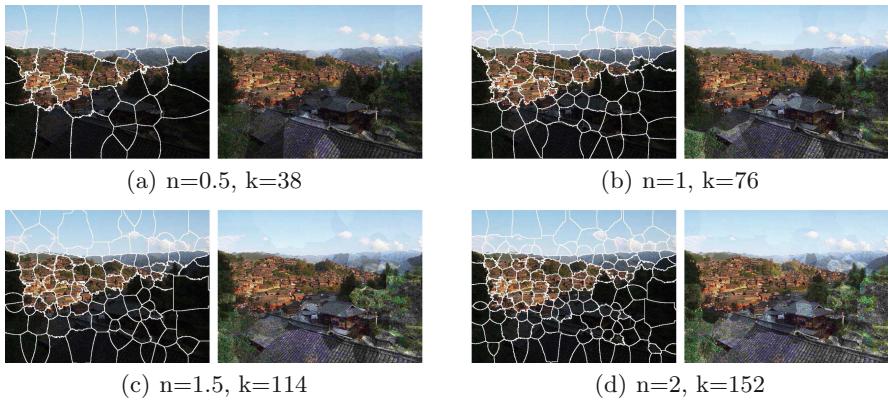


Fig. 5. Experimental results of superpixel segmentation and homomorphic filtering of natural image at different scales

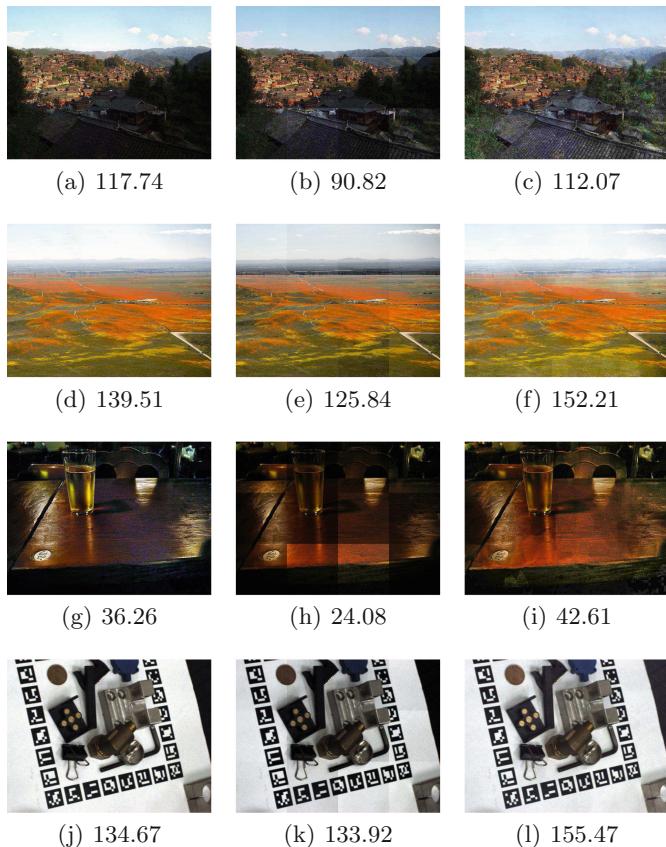


Fig. 6. Different experimental results of four images. (a)(d)(g)(j) Funt's method [11]. (b)(e)(h)(k) Zhang's method [14]. (c)(f)(i)(l) Our method

Figure 6 shows the experimental results with four different illumination images using three different image enhancement methods. Funt’s method [11] has the highest AVG, there are halos appearing at the edges of the image, and the color of the enhanced image is distorted. Zhang’s method [14] has a low AVG, and there are obviously blocking artifact at the edges of subblocks. Compared with the above method, our method has a better visual ability and higher AVG.

The AVG-GRAD is used as another evaluation index to judge the effectiveness and rationality of the method, it represents the sharpness of the image using the changing rate of the image’s grayscale. The definition of the AVG-GRAD is as follows.

$$\bar{g} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \sqrt{\frac{(F(i, j) - F(i + 1, j))^2 + (F(i, j) - F(i, j + 1))^2}{2}} \quad (16)$$

where $F(i, j)$ is the gray value at (i, j) . M and N are the rows and columns of the image respectively. The size of image’s AVG-GRAD has positive correlation with the quality and definition of images.

Table 1 shows the AVG-GRAD of different methods. Our method has a better performance on natural image, aerial image, and component image under the AVG-GRAD. Although the AVG-GRAD of Funt’s method [11] in indoor low-light image is higher than that of our method, its AVG is lower and there are still large dark areas in this image. The comparison shows that our method can improve AVG and AVG-GRAD in image enhancement.

Table 1. Comparison of different methods

Method	Natural image		Aerial image		Indoor low-light image		Component image	
	Avg	Avg-Grad	Avg	Avg-Grad	Avg	Avg-Grad	Avg	Avg-Grad
Funt et al.[11]	95.71	0.0246	142.15	0.0121	39.37	0.0293	150.19	0.0258
Zhang et al.[14]	90.81	0.0276	125.84	0.0208	24.08	0.0153	133.92	0.0260
Ours	112.07	0.0375	152.21	0.0218	42.61	0.0255	155.47	0.0265

5 Conclusions

This paper proposes a method to enhance component image using multi-scale superpixel based homomorphic filtering. We use the superpixel segmentation method to perform multi-scale segmentation on component image and homomorphic filtering on the superpixels. The images at different scales are weighted according to the standard deviation and fused to obtain the enhanced image. The effectiveness of this method has been shown by experimental results in natural image, aerial image, indoor low-light image, and component image. In the future, we plan to optimize our method, reduce the complexity of the algorithm, and improve its operation efficiency.

Acknowledgments. This work is supported by the National Key Research and Development Program of China [No. 2018YFC0808904].

References

1. Ye, S.H., Zhu, J.G., Wang, Z., Yang, X.Y.: Visual inspection technology and its application. *Eng. ENCE* **1**, 49–52 (1999). <https://doi.org/10.3969/j.issn.1009-1742.1999.01.011>
2. Guo, Y.Q., Wang, J.L., Qiu, K.P., Zhao, L.Q., Fu, X.S., Wang, R.T.: Specular reflection separation for a single image using IR-SIP. In: Chinese Intelligent Systems Conference, pp. 504–512 (2019). https://doi.org/10.1007/978-981-32-9686-2_57
3. Cai, B.L., Xu, X.M., Guo, K.L., Jia, K., Hu, B., Tao, D.C.: A Joint intrinsic-extrinsic prior model for retinex. In: 2017 IEEE International Conference on Computer Vision (ICCV), Venice, pp. 4020–4029 (2017). <https://doi.org/10.1109/ICCV.2017.431>
4. Pizer, S.M., Johnston, R.E., Erickson, J.P., Yankaskas, B.C., Muller, K.E.: Contrast-limited adaptive histogram equalization: speed and effectiveness. In: [1990] Proceedings of the First Conference on Visualization in Biomedical Computing, Atlanta, GA, USA, pp. 337–345 (1990). <https://doi.org/10.1109/VBC.1990.109340>
5. Reza, A.M.: Realization of the contrast limited adaptive histogram equalization (CLAHE) for realTime image enhancement. *J. VLSI Signal Process.* **38**(1), 35–44 (2004). <https://doi.org/10.1023/b:vlsi.0000028532.53893.82>
6. Wadud, M.A.A., Kabir, M.H., Dewan, M.A.A., Chae, O.: A dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* **53**(2), 593–600 (2007). <https://doi.org/10.1109/tce.2007.381734>
7. Wan, Y., Shi, D.B.: Joint exact histogram specification and image enhancement through the wavelet transform. *IEEE Trans. Image Process.* **16**(9), 2245–2250 (2007). <https://doi.org/10.1109/TIP.2007.902332>
8. Huang, L.D., Zhao, W., Wang, J., Sun, Z.B.: Combination of contrast limited adaptive histogram equalisation and discrete wavelet transform for image enhancement. *Image Process.* **9**(10), 908–915 (2015). <https://doi.org/10.1049/iet-ipr.2015.0150>
9. Jobson, D.J., Rahman, Z., Woodell, G.A.: Properties and performance of a center/surround retinex. *IEEE Trans. Image Process.* **6**(3), 451–462 (1997). <https://doi.org/10.1109/83.557356>
10. Rahman, Z., Jobson, D.J., Woodell, G.A.: Multi-scale retinex for color image enhancement. In: Proceedings of 3rd IEEE International Conference on Image Processing, Lausanne, Switzerland, pp. 1003–1006 (1996). <https://doi.org/10.1109/ICIP.1996.560995>
11. Funt, B.V., Ciurea, F., Mccann, J.J.: Retinex in MATLAB. *J. Electron. Imaging* **13**(1), 48–57 (2000). <https://doi.org/10.1117/1.1636761>
12. Jiao, Z.Q., Xu, B.G.: Color image illumination compensation based on homomorphic filtering. *Guangdianzi Jiguang/J. Optoelectron. Laser* **21**(4), 602–605 (2010). <https://doi.org/10.16136/j.joel.2010.04.027>
13. Jmal, M., Souiden, W., Attia, R.: New color image illumination enhancement technique based on homomorphic filtering. In: 2014 5th European Workshop on Visual Information Processing (EUVIP), Paris, pp. 1–6 (2015). <https://doi.org/10.1109/EUVIP.2014.7018406>
14. Zhang, Y.F., Xie, M.H.: Colour image enhancement algorithm based on HSI and local homomorphic filtering. *Comput. Appl. Softw.* **30**(12), 303–307 (2013). <https://doi.org/10.3969/j.issn.1000-386x.2013.12.080>

15. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süstrunk, S.: SLIC Superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(11), 2274–2282 (2012). <https://doi.org/10.1109/TPAMI.2012.120>
16. Zhang, Y.X., Li, X.M., Gao, X.F.: A simple algorithm of superpixel segmentation with boundary constraint. *IEEE Trans. Circ. Syst. Video Technol.* **27**(7), 1502–1514 (2017). <https://doi.org/10.1109/TCSVT.2016.2539839>



Multi-scale Visual Saliency Fusing with Spatiotemporal Features for Fire Detection

Huanyu Guo, Jianlin Wang^(✉), Yongqi Guo, Xinjie Zhou, and Kepeng Qiu

Beijing University of Chemical Technology, Beijing 100029, China
wangjl@mail.buct.edu.cn

Abstract. Aiming at the problem that the existing methods cannot realize the fire detection accurately, a multi-scale visual saliency fusing with spatiotemporal features for fire detection method is proposed. Firstly, the simple linear iterative clustering (SLIC) method and region contrast (RC) method are used to realize multi-scale visual saliency detection of fire images. Then the spatiotemporal features of fire are extracted in the time domain and space domain respectively. Finally, multi-scale visual saliency detection and the spatiotemporal features of fire are combined to realize the fire detection. The experimental results show the proposed method has a higher accuracy rate and a lower false rate.

Keywords: Superpixel segmentation · Region contrast · Visual saliency · Spatiotemporal features · Fire detection

1 Introduction

Fire detection is very important for fire warning and fire extinguishing. Accurate fire detection can improve the efficiency of firefighting. The traditional fire detection methods rely on the temperature sensors and smoke sensors, and they have the disadvantages of limited space and poor reliability [1]. With the development of computer vision technology, image-based fire detection methods have been widely researched [2]. These researches mostly focus on the color features of fire, so they cannot meet the requirements of high accuracy rate and low false rate for fire detection. Thus, researchers begin to utilize more features of fire to detect fire. Phillips et al. [3] built a gaussian smooth color histogram based on the color and motion information to detect fire. Liu et al. [4] combined the spectrum, space and time model of the fire to detect fire. Chen et al. [5] proposed a multi-feature mixed fire detection method to detect fire. Yan et al. [6] used the bright feature to locate the suspected fire regions, then the local binary pattern (LBP) and support vector machine (SVM) were utilized to detect fire.

The human visual system can search the salient region in the complex scenes. Itti et al. [7] extracted multi-scale image features to obtain the salient region. On

the basis of Ittis method, Harel et al. [8] proposed a graph-based visual saliency (GBVS) detection method and introduced the Markov chain to obtain the salient region. Moreover, some researchers have found that there is a strong contrast between the salient region and the background region. Cheng et al. [9] proposed a histogram-based contrast (HC) saliency detection method to obtain the salient region. Then Cheng et al. [10] proposed a region contrast (RC) saliency detection method, which considers both global contrast and spatial location.

In this paper, we combine the visual saliency detection method with fire features, and propose a multi-scale visual saliency fusing with spatiotemporal features for fire detection method. Firstly, the simple linear iterative clustering (SLIC) method and RC method are used to realize multi-scale visual saliency detection of fire images. Then the spatiotemporal features of fire are extracted in the time domain and space domain respectively. Finally, multi-scale visual saliency detection and the spatiotemporal features of fire are combined to realize the fire detection.

2 Multi-scale Visual Saliency Detection

The scale of fire in the image varies greatly, so the fire region cannot be accurately detected according to existing methods. Superpixel is a collection of local and coherent pixels, and it is widely used in the field of image segmentation [11]. We can combine superpixel and visual saliency detection in different scales to obtain the salient region.

2.1 Superpixel Segmentation

SLIC is a segmentation method that generates regular and compact superpixels by using the color similarity and spatial distance of pixels to cluster pixels [12]. Firstly, RGB color space image is converted to Lab color space image, and the cluster center needs to be initialized according to the number of seed points. Secondly, the seed points are moved to the position with the smallest gradient in a 3×3 neighborhood, and each pixel is assigned a class label in a $2S \times 2S$ region around the cluster center. Finally, the cluster center is constantly updated until it no longer changes. The entire process is shown in Fig. 1. The calculation and synthesis of the distance in the clustering process can be defined as formula (1), formula (2) and formula (3).

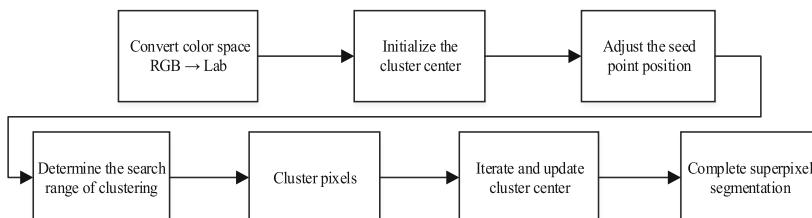


Fig. 1. Simple linear iterative clustering flowchart

$$d_c = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \quad (1)$$

$$d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (2)$$

$$D = \sqrt{\frac{d_c^2}{N_c} + \frac{d_s^2}{N_s}} \quad (3)$$

where d_c represents the color distance in Lab color space between superpixels i and j , and d_s represents the spatial distance between superpixels i and j . N_c represents the maximum color distance, and the distance generally takes a constant of 10. N_s represents the maximum space distance between seed points, and the distance generally is defined as $S = \sqrt{\frac{N}{K}}$. Where K represents the number of seed points, and N represents the number of pixels.

2.2 Region Contrast

The image is segmented into various numbers of superpixels by using the SLIC method in Sect. 2.1. The distance between superpixels also affects the obtention of the salient map, so we introduce spatial weight to balance the effect of superpixel distance on visual saliency detection. For a superpixel, its saliency value is calculated according to region contrast as formula (4).

$$C_i = \sum_{j=1}^N d_c(P_i, P_j) \omega(P_i, P_j) \quad (4)$$

where the color distance $d_c(P_i, P_j)$ can be defined as formula (5), and the spatial weight $\omega(P_i, P_j)$ can be defined as formula (6). N represents the number of superpixels in the image.

$$d_c(P_i, P_j) = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \quad (5)$$

$$\omega(P_i, P_j) = \exp\left(-\frac{d_p(P_i, P_j)^2}{\sigma}\right) \quad (6)$$

where the spatial distance $d_p(P_i, P_j)$ can be defined as formula (7), and σ plays an adjustment role.

$$d_p(P_i, P_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (7)$$

2.3 Fusion of Multi-scale Saliency Maps

The saliency maps at different scales are different. Therefore, the Normalized Euclidean distance in the Lab color space between the pixels and the center of superpixels where the pixels are located is used as the weight to fuse multi-scale saliency maps in this paper. The fused saliency map is defined as formula (8).

$$S_p = \frac{\sum_{m=1}^M W^m(p) \times S^m(p)}{\sum_{m=1}^M W^m(p)} \quad (8)$$

where $S^m(p)$ represents the single-scale saliency map obtained at scale m ($m = 1, 2, \dots, M$), and $W^m(p)$ represents the fusion weight, which is defined as formula (9).

$$W^m(p) = \frac{1}{1 + d_c(p, s(p))^\sigma} \quad (9)$$

where $d_c(p, s(p))$ represents the Normalized Euclidean distance in the Lab color space between the pixel p and the center of superpixel $s(p)$ where the pixel p is located. σ plays an adjustment role.

The fusion of multi-scale saliency maps can improve the robustness of visual saliency detection for segmentation scale and make the salient region more intuitive.

3 Spatiotemporal Features of the Fire

3.1 Spatial Features of the Fire

The spatial features of the fire include special color and outstanding brightness. If the spatial feature extraction is only realized in the RGB color space, the brightness feature will be lost. However, in Lab color space, L component represents brightness. Therefore, the features of Lab color space and RGB color space are comprehensively considered in this paper. According to the above characteristics and a large number of experimental summaries, the spatial features of the fire can be defined as formula (10).

$$L > L_N, a > a_N, b > b_N, R - B > K \quad (10)$$

where a_N and b_N represent the color threshold, and L_N represents the brightness threshold. K represents the relation threshold in RGB color space.

3.2 Temporal Features of the Fire

The fire changes dynamically with time, so the motion of the region is one of temporal features of the fire. The inter-frame difference method [13], the background difference method [14] and the optical flow method [15] are common motion detection methods. It is difficult to meet the application conditions of the optical flow method in the complex fire scene. Therefore, this paper combines the inter-frame difference method and the background difference method to improve the adaptability of temporal feature extraction for the fire scene.

When the motion target moves slowly, there are many holes in the motion target obtained by the inter-frame difference method. Aiming at this problem, the cumulative three-frame difference method first binarizes the frame difference results, then superimposes the binary images, so that the obtained motion target has fewer holes. Although the motion target will be slightly enlarged by this method, it has little effect on the detection of fire. Comparing the current frame

with the background frame to realize motion detection is the basic idea of the background difference method. Background modeling directly affects the result of detection. The adaptive background modeling method with good effect and fast speed is used in this paper. The method can be defined as formula (11).

$$B_{f_{r+1}}(x, y) = (1 - \alpha) \cdot B_{f_r}(x, y) + \alpha \cdot C_{f_r}(x, y) \quad (11)$$

where $B_{f_r}(x, y)$ represents the current background frame. $C_{f_r}(x, y)$ represents the current frame. $B_{f_{r+1}}(x, y)$ represents the next background frame. α represents the background update rate.

4 Multi-scale Visual Saliency Fusing with Spatiotemporal Features for Fire Detection

The flowchart of multi-scale visual saliency fusing with spatiotemporal features for fire detection is shown in Fig. 2.

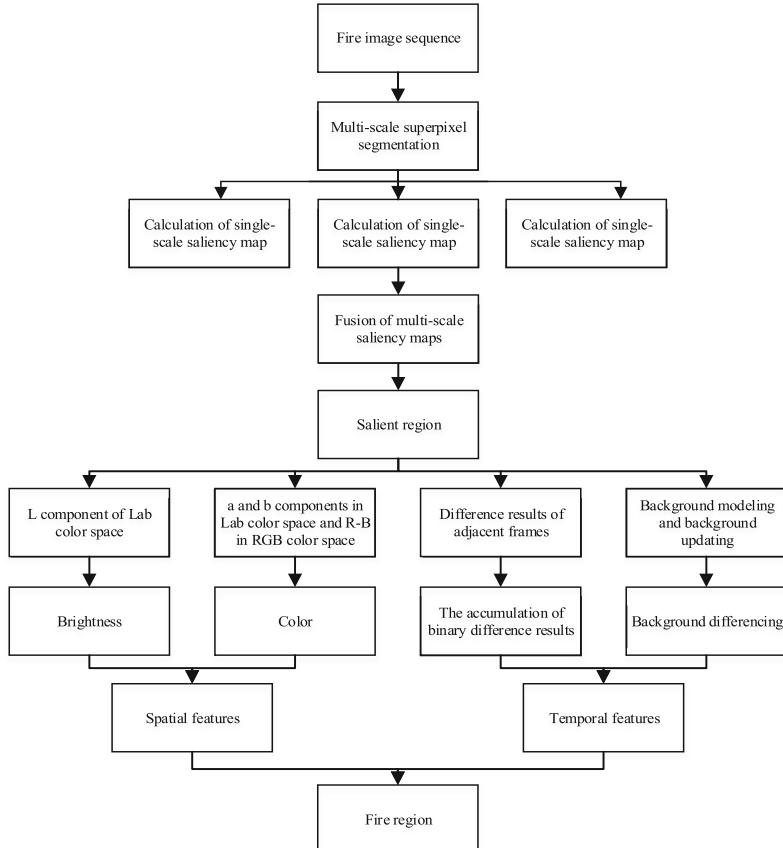


Fig. 2. Multi-scale visual saliency fusing with spatiotemporal features for fire detection flowchart

The steps of multi-scale visual saliency fusing with spatiotemporal features for fire detection are as follows:

- (1) Use the SLIC method to realize multi-scale superpixel segmentation of fire image sequence;
- (2) Use the RC method to obtain the saliency map at a single scale;
- (3) Fuse saliency maps at different scales to obtain the salient region;
- (4) Obtain the spatial features of fire according to the color and brightness of fire;
- (5) Obtain the temporal features of fire by combining the cumulative three-frame difference method and the background difference method;
- (6) Identify whether the salient region is the fire region according to spatiotemporal features.

5 Experiment

In this section, several videos from the Bilkent University fire video library and the Internet are used to experiment on Visual studio 2017. The multi-scale visual saliency detection method in this paper is used to compare with four saliency detection methods (LC, SR, HC, RC). The visual comparison results are shown in Fig. 3.

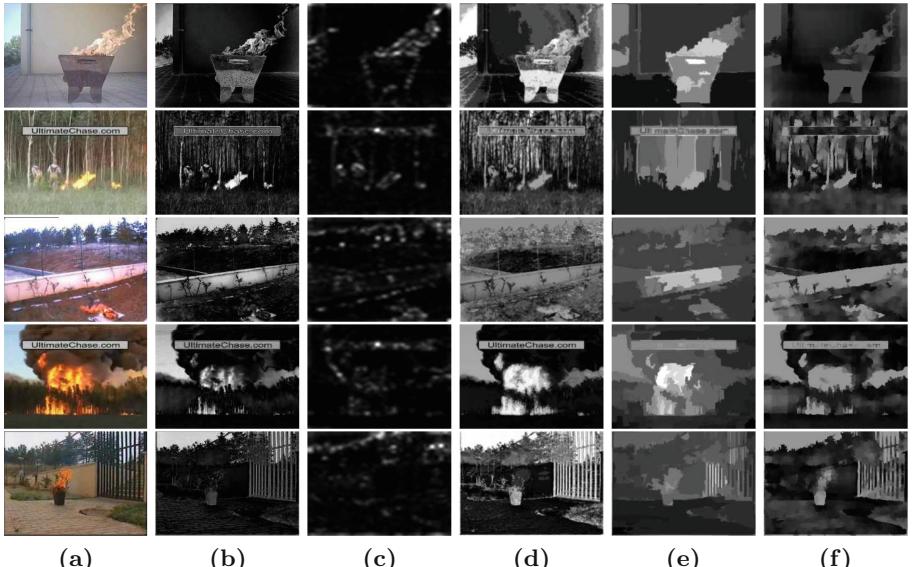


Fig. 3. The visual comparison results of different methods (a) Images (b) LC (c) SR (d) HC (e) RC (f) Ours

It can be seen from the figure that the LC method cannot accurately detect salient regions that are similar to background regions in the third and fifth fire images because of the lack of local contrast. The SR method only considers the overall information of the image in the frequency domain, so the salient regions obtained are not complete enough in the fourth and fifth fire images. The HC method has a poor detection effect on the fire images with complex texture, so the salient regions are not obvious enough in the second and third fire images. The RC method detection result is easily affected by the accuracy of region segmentation, so the regions around the fire regions are also recognized as the salient regions in the second and fifth fire images. The proposed method is superior to the other four visual saliency detection methods from a subjective visual point of view, and the salient region obtained is more intuitive and complete.

However, the salient region obtained by the multi-scale visual saliency detection method not only includes the fire region, but also includes other regions with strong color contrast in the image. Therefore, the spatiotemporal features of fire are used to further detect fire. The value of each component in formula (10) is expanded to [0, 255], and the threshold is obtained by the method of statistical histogram. The threshold range is $L_N \in [100, 150]$, $a_N \in [100, 120]$, $b_N \in [100, 120]$, $K \in [35, 80]$. The experimental effects are shown in Fig. 4.



Fig. 4. Fire detection effects. (a) Original images. (b) Fire regions

Table 1. Fire videos detection results

Fire videos		Accuracy frames		Accuracy rate/%	
Number	Fire frames	Yan et al. [6]	Ours	Yan et al. [6]	Ours
1	170	160	155	94.1	91.2
2	224	210	216	93.8	96.4
3	84	78	81	92.9	96.4
4	98	91	94	92.9	95.9
5	200	180	189	90	94.5

Table 2. Non-fire videos detection results

Non-fire videos		False frames		False rate/%	
Number	Non-fire frames	Yan et al. [6]	Ours	Yan et al. [6]	Ours
6	67	4	2	6	3.0
7	197	13	10	6.6	5.1
8	239	83	54	34.7	22.6

We compare the proposed method and the method proposed by Yan [6] to conduct experiments on 5 pieces of fire video and 3 pieces of non-fire video respectively. The experimental comparison results are shown in Table 1 and Table 2.

The fire in video 1 is erratic, and the color of the fire is close to the color of the wall. Yan's method only considers the texture and color features of the fire, so the accuracy rate is slightly higher than the proposed method. Although the proposed method in videos 2, 3, 4 and 5 can detect the fire region in each frame, non-fire regions are mistakenly detected as the fire regions in some frames. Thus, the accuracy rate has not reached 100%. Video 6 is an indoor video with a complex background change. Video 7 is a video of cars passing through a tunnel. Video 8 is a video of vehicles driving on the street at night. For the above three non-fire videos, the change of lights and the movement of nearby cars have a great interference on the detection of fire. There are regions in the image that are similar to the fire regions in the spatial and temporal features, so a small number of false detection frames exist in the two methods. However, the overall accuracy rate of the proposed fire detection method is higher, and the overall false rate is lower.

6 Conclusions

This paper proposes a multi-scale visual saliency fusing with spatiotemporal features for fire detection method. We combine the visual saliency detection method and the fire features to detect fire. The experimental results show the proposed method has a higher accuracy rate and a lower false rate. However, when there are some motion targets with similar spatial features to the fire region

in the scene, the proposed method has a higher false rate. Our future work is to improve the robustness of the proposed method for this situation.

Acknowledgements. This work was supported by the National Key Research and Development Program of China [No. 2017YFF0107303].

References

1. Li, Z.L., Mihaykova, L.S., Isupova, O., Rossi, L.: Autonomous flame detection in videos with a Dirichlet process Gaussian mixture color model. *IEEE Trans. Ind. Inform.* **14**(3), 1146–1154 (2018). <https://doi.org/10.1109/TII.2017.2768530>
2. Li, W.H., Wang, Y.: High-precision video flame detection algorithm based on multi-feature fusion. *J. Jilin Univ.* **40**(3), 69–775 (2010). <https://doi.org/10.13229/j.cnki.jxbgxb2010.03.040>
3. Phillips, W., Shah, M., Lobo, N.D.V.: Flame recognition in video. *Pattern Recogn. Lett.* **23**(1), 319–327 (2002). [https://doi.org/10.1016/S0167-8655\(01\)00135-0](https://doi.org/10.1016/S0167-8655(01)00135-0)
4. Liu C.B., Ahuja N.: Vision based fire detection. In: International Conference on Pattern Recognition, vol. 4, pp. 134–137 (2004). <https://doi.org/10.1109/ICPR.2004.979>
5. Chen, J., He, Y.P., Wang, J.: Multi-feature fusion based fast video flame detection. *Build. Environ.* **45**(5), 1113–1122 (2010). <https://doi.org/10.1016/j.buildenv.2009.10.017>
6. Yan, Y.Y., Tang, Y.Y., Liu, Y.A., Zhang, T.Y.: Flame detection based on LBP features with multi-scales and SVM. *J. Shandong Univ.* **42**(5), 47–52 (2012). <https://doi.org/10.6040/j.issn.1672-3961.2012.05.009>
7. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(11), 1254–1259 (1998). <https://doi.org/10.1109/34.730558>
8. Harel, J., Koch, C., Perona, P.: Graph-based visual saliency. In: Neural Information Processing Systems, vol. 19, pp. 545–552 (2006). <https://doi.org/10.7551/mitpress/7503.003.0073>
9. Cheng, M.M., Zhang, G.X., Mitra, N.J., Huang, X.L., Hu, S.M.: Global contrast based salient region detection. *Comput. Vision Pattern Recogn.* **37**(3), 409–416 (2011). <https://doi.org/10.1109/CVPR.2011.5995344>
10. Cheng, M.M., Mitra, N.J., Huang, X.L., Torr, P.H.S., Hu, S.M.: Global contrast based salient region detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(3), 569–582 (2015). <https://doi.org/10.1109/tpami.2014.2345401>
11. Ren, X.F., Malik, J.: Learning a classification model for segmentation. In: International Conference on Computer Vision, pp. 10–17 (2003). <https://doi.org/10.1109/ICCV.2003.1238308>
12. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(11), 2274–2282 (2012). <https://doi.org/10.1109/TPAMI.2012.120>
13. Kim, W., Jung, C., Kim, C.: Spatiotemporal saliency detection and its applications in static and dynamic scenes. *IEEE Trans. Circuit. Syst. Video Technol.* **21**(4), 446–456 (2011). <https://doi.org/10.1109/TCSVT.2011.2125450>

14. Amri, S., Barhoumi, W., Zagrouba, E.: A robust framework for joint background/foreground segmentation of complex video scenes filmed with freely moving camera. *Multimed. Tools Appl.* **46**(2), 175–205 (2010). <https://doi.org/10.1007/s11042-009-0348-y>
15. Rahtu, E., Kannala, J., Salo, M., Heikkila, J.: Segmenting salient objects from images and videos. In: European Conference on Computer Vision, pp. 366–379 (2010). https://doi.org/10.1007/978-3-642-15555-0_27



Summary and Prospect of Predictive Control for a Class of Economic System

Yiwen Yang^{1(✉)} and Haotian Peng²

¹ Business School, Hohai University, Nanjing 211100, China
120429248@qq.com

² College of Mechanical and Electrical Engineering,
Hohai University, Changzhou 213022, China
karl.raphe1@hhu.edu.cn

Abstract. This paper reviews the development of predictive control in economic systems. This paper mainly introduces the predictive control of economic system based on mathematical model, especially the predictive control of economic system based on non-cobweb model and the predictive control of economic system based on nonlinear analysis. The control strategy is derived step by step from the establishment of the model and the conditional assumptions, and the application of robust control in guaranteeing the stability of the economic system and the application of nonlinear science in the nonlinear aspect of the economic system are explained in detail.

Keywords: Economic system predictive control · Non - cobweb model · Nonlinear system

1 Introduction

Predictive control is a computer control algorithm that adopts multi-step prediction, rolling optimization and feedback correction [1] control strategy. The basic step of predictive control is to predict the future state of the system to determine the current control strategy, and then use the model output error for feedback correction. Due to its good tracking performance, anti-interference ability and robustness to model errors [2], predictive control has derived dozens of models and algorithms, which have obtained a series of gratifying research results in many fields such as petroleum, electric power, agriculture and machinery, and have been widely applied.

The economic system is a diversified and complex system, which is influenced by many factors such as politics, environment and so on. All of the economic system through continuous input and output in a variety of ways and the environment interaction [3]. Moreover, the economic system also has a unique, multiple targets, economic benefit and social benefit need to be considered at the same time. More importantly, the economic system is a directly decided by the people and the control system. Because of people's preferences and different

ways of thinking, Economic system with uncertainty and fuzziness. How to keep the economic openness and stability of the system and reduce the subjectivity and limitation of human factor management control has become a need to input energy to the problem of serious study. The advantages of predictive control thought just can provide some solutions to this problem. In order to improve the economic system is scientific and accuracy, and to combine the predictive control thoughts and economic system and related research appears very necessary [4].

The rest of this paper is organized as follows: the second part provides a detailed literature review of the predictive control of economic systems based on the non-cobweb model. In the third part, a detailed literature review is made on the economic model predictive control based on nonlinear analysis system. The fourth part gives some new development of predictive control in economic system and prospects the application prospect of predictive control in economic system. Finally, the conclusion is given in the fifth part.

2 Predictive Control of Economic System Based in Non-cobweb Model

Under normal circumstances, the market price, quantity supplied and quantity demanded of commodities will change with the change of time and show the law of alternating change. Therefore, the following disequilibrium cobweb mode with external interference is established [5]:

$$\begin{aligned} D_k &= a_1 G_k + \beta_1 p_k, \beta_1 < 0, \\ S_k &= a_2 M_k + \beta_2 p_k^*, \beta_2 > 0, \\ p_k^* &= p_{k-1} + \gamma (p_{k-1} - p_{k-2}), \\ Q_k &= \min \{D_k, S_k\}, k = 0, 1, 2 \dots, N. \end{aligned} \quad (1)$$

In the formula: D_k is the effective demand for goods in phase k . S_k is the effective supply quantity of goods in phase k . p_k is the price of goods in phase k . p_k^* is the expected price of goods in phase k . Q_k is the trading volume of the goods in issue k . G_k is the external variables that affect demand. M_k is the external variable that affects the supply. $a_1, a_2, \beta_1, \beta_2, \gamma$ is the coefficient of each variable. The third formula represents the price adjustment equation, indicating the economic characteristics of prices rising when demand exceeds supply and falling when supply exceeds demand in the market. The third formula says “short side rule”.

Based on the above model, the following assumptions are made [6]:

$$\begin{aligned} x_1(k) &= D_k - S_k, x_2(k) = D_{k-1} - S_{k-1}, \\ x_3(k) &= P_k - P_{k-1}, w_1(k) = G_{k+1} - G_k, \\ w_2(k) &= M_{k+1} - M_k. \end{aligned} \quad (2)$$

Therefore, the unbalanced cobweb model is transformed into the state space form as:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + Dw(k), \\ y(k) = D_k - S_k = x_1(k) = Cx(k), \end{cases} \quad (3)$$

where $x(k) = [x_1(k), x_2(k), x_3(k)]^T$ is a state vector, $u(k) = [(P_{k+1} - P_k) - (P_k - P_{k-1})]$ is the control input vector, $w(k) = [w_1(k), w_2(k)]^T$ is the interference vector,

$$\text{and } A = \begin{bmatrix} 1 - \beta_2\gamma & \beta_2\gamma & \beta_1 - \beta_2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} \beta_1 \\ 0 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^T, D = \begin{bmatrix} a_1 & a_2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Assume that the external disturbance input $w(k)$ is the step function and in a step, $\mathbf{w}^* = (w_1^*, w_2^*)^T$ is a constant vector of unknown value. At this point, the following robust control strategy can be established [7]:

$$\begin{aligned} x(k+1) &= x(k) + y(k) = x(k) + Cx(k), \\ u(k) &= (k_1 k_2 k_3) x(k) + k_4 x(k), \end{aligned} \quad (4)$$

where k_1, k_2, k_3, k_4 are the undetermined parameter.

Then the state matrix and the corresponding characteristic equation of the system are derived, and the value of k_1, k_2, k_3, k_4 are expressed under the condition that the closed-loop pole assignment is 0. Put the value of k_1, k_2, k_3, k_4 into the control system and get the following feedback control strategy:

$$u_k = p_{k+1} - p_k = k_1 (D_k - S_k) + k_2 (D_{k-1} - S_{k-1}) + k_3 (p_k - p_{k-1}) + k_4 x(k). \quad (5)$$

Due to the uncertain disturbance of demand and supply in the unbalanced economic system, to achieve the goal of price control in the economic system, a robust control feedback mechanism is established, that is, to control the external uncertain disturbance by adjusting the price difference between two adjacent periods. γ represents the degree of inhibition of disturbance, and formula $\|\mathbf{y}\|_2 \leq \gamma \|w\|_2$ represents the gain of two external inputs interfering with the difference between supply and demand of commodities. The minimum difference between supply and demand of commodities can be obtained by robust control based on the non-cobweb model, so that the supply and demand of commodities can reach the ideal value [8]. For an open dynamic economic system that is subject to constant external interference, robust control can make the system run stably and achieve the desired adjustment purpose. Many simulation experiments have also verified its feasibility [9, 10].

3 Predictive Control of Economic Systems Based on Nonlinear Analysis

The economic system is complicated and involves many aspects. The analysis and research of many models on the economic system is chaotic, and more and more economic paradoxes appear with the development of time. In order to improve the scientificity and accuracy of economic system and strengthen the guidance of practice, nonlinear economics emerged. Nonlinear economics uses nonlinear science to remove the fog of chaos in the economic system, and provides a basis for the law behind the economic system and its predictive control [11, 12].

Firstly, a discrete - time nonlinear system is established [13]:

$$x_{k+1} = f(x_k, u_k), k \in I_{\geq 0}, \quad (6)$$

where $x_k \in R^n, u_k \in R^m$ are state variables and control variables at time k , respectively, $f : R^n \times R^m \rightarrow R^n$ is a continuous function of (x, u) . Let's say there's an equilibrium point in the system where (x, u) satisfies $\mathbf{x}_s = f(x_s, u_s)$. Then there are states and control constraints on the system as follows:

$$\mathbf{x}_k \in X, u_k \in U, k \in I_{\geq 0}, \quad (7)$$

where constraint set $X \subset R^n, U \subset R^m$ is the convex compact set, and equilibrium point $(\mathbf{x}_s, \mathbf{u}_s)$ is its interior point. Assume that the system state is fully measurable.

Let the prediction time domain be $N \in I_{\geq 1}$ and define the prediction control sequence at moment k as $u_k = \{u_0|_k, u_1|_k, \dots, u_{N-1}|_k\}$ and the corresponding prediction state sequence $\mathbf{x}_k = \{x_1|_k, x_2|_k, \dots, x_N|_k\}$.

Given a compact set $X_T \subset X$, if condition $x_{i+1|k} \in X, u_{i|k} \in U, \forall i \in I_{0:N-1}, x_{N|K} \in X_T$ is satisfied at \mathbf{u}_k , then the above control sequence is a feasible predictive control sequence of the system.

Then, based on the feasible predictive control u_k of the system and the corresponding state prediction sequence \mathbf{x}_k , the following objective function is defined:

$$V_a(x_k, u_k) = E_a(x_{N|k}) + \sum_{i=0}^{N-1} (x_{i|k}, u_{i|k}), \quad (8)$$

where $\mathbf{x}_{0|k} \in X$ is the state at time k .

Combining the continuous bounded functions $L_a : X \times U \rightarrow R_{\geq 0}$ and $E_a : X \in R_{\geq 0}$, the economic objective function is further defined as follows:

$$J(x_k, u_k) = \sum_{i=0}^{N-1} (x_{i|k}, u_{i|k}). \quad (9)$$

Then, the following time domain optimal control problem is solved at each time:

$$\begin{aligned} u_k^* &= \arg \min_{u_k} J(x_k, u_k), \\ s.t. x_{i+1|k} &= f(x_{i|k}, u_{i|k}), i \in I_{0:N-1}, \\ (x_{i|k}, u_{i|k}) &\in X \times U, i \in I_{0:N-1}, \\ \mathbf{x}_{0|k} &= x_k, \mathbf{x}_{N|k} \in X_T, \\ V_a(x_k, u_k) &\leq \eta(x_k, \alpha). \end{aligned} \quad (10)$$

In the formula: u_k^* is the optimal solution of the optimization problem, $\mathbf{x}_{0|k} = x_k, \mathbf{x}_{N|k} \in X_T$ are the initial condition and the terminal constraint condition respectively, Terminal constraint set $X_T \subset X$ and $x_s \in X_T$.

Let the constraint condition of function η be $X \times R_{\geq 0} \rightarrow R_{\geq 0}$, and constructor X define the auxiliary optimization problem:

$$u_k^o = \arg \min_{u_k} \{V_a(x_k, u_k)\} \mid s.t. x_{i+1|k} = f(x_{i|k}, u_{i|k}) - (x_{i|k}, u_{i|k}) \in X \times U \mid . \quad (11)$$

In the formula: u_k^o is the optimal solution to the optimization problem, put the values of u_k^o and u_k^* into formula (8) to get the value function at time k :

$$\begin{aligned} V_a^*(x_k) &:= V_a(x_k, u_k^*), \\ V_a^o(x_k) &:= V_a(x_k, u_k^o). \end{aligned} \quad (12)$$

From the above discussion, function η can be defined as:

$$\eta(x_k, a) = V_a^o(x_k) + \alpha [V_a^*(x_{k-1}) - V_a^o(x_k)], \quad (13)$$

where coefficient $\alpha \geq 0$, it can be proved by derivation that the inequality $\eta(x, \alpha) \geq 0$ is true for any $x \in X$ and $\alpha \geq 0$.

As complex working conditions often change in the economic production process, and it is difficult for a single model to obtain an accurate global model for the actual process, MPC control of multi-model strategy based on mixed logic is designed to control such complex nonlinear systems [14, 15].

The prediction time domain and parameter time domain are represented by the parameter T .

The model prediction process is as follows:

$$\left\{ \begin{array}{l} x(k+1/k) = Ax(k/k) + B_1u(k/k) + B_2\delta(k/k) + B_3z(k/k) + B_4, \\ x(k+2/k) = Ax(k+1/k) + B_1u(k+1/k) + B_2\delta(k+1/k) + B_3z(k+1/k) \\ \quad + B_4, \\ \vdots \\ x(k+T-1/k) = Ax(k+T-2/k) + B_1u(k+T-2/k) + B_2\delta(k+T-2/k) \\ \quad + B_3z(k+T-2/k) + B_4, \\ x(k+T/k) = x_f. \end{array} \right. \quad (14)$$

Making the state of the system after the future time equal to the expected terminal state is a commonly used constraint method to guarantee the stability of the closed-loop system in MPC system. When T is correlated with the prediction time domain, the value of T is small, which will make the system unfeasible, while increasing the value of T can make the optimization problem reach a better solution. With the increase of the number of models, the optimization process will be close to the global optimal solution. However, more is not always better, because the more models there are, the higher the number of calculations will be. Therefore, the most important thing is that we should choose the number of required models according to the actual situation, flexibly balance the control effect and calculation amount, and set the appropriate prediction time domain and control time domain, so as to facilitate the acquisition and better control of the economic system [16].

It is worth noting that although the nonlinear system plays a role in the predictive control of the economic system and makes up for the lack of the singleness of the linear system, its disadvantages are still obvious, such as the lack of a sound basic theory and many available mature methods. Therefore, in the application of nonlinear systems, the application of nonlinear science should be more in-depth understanding and research. And to be clear, the nonlinear system cannot completely replace the previous linear system. Therefore, in the study of specific economic systems, how to create a new nonlinear method and give full play to its advantages is an important subject that still needs to spend energy and be studied well in the future [17].

4 Future Prospects

In essence, the application of predictive control in economic systems is to create new methods and strategies based on relevant models. Due to the advantages of model predictive control such as convenience in modeling, system robustness, good stability, rolling optimization strategy, ability to effectively deal with multiple variables and constrained problems [18], economic systems will still show a trend of combining with model predictive control in the future. In recent years, model predictive control has been widely used in economic systems. By establishing a stochastic model, Weiguo Zhang et al. studied the dynamic hedging problem of European options and conducted simulation experiments to verify the effectiveness of the stochastic model [19]. Ronggui Luo et al. 's adaptive neuro-fuzzy inference system (ANFIS) combined with BP neural network to simulate the functions of eco-economic systems for predictive control [20]. Zixian Li et al. introduced the idea of predictive control into cost control, and established a cost predictive control model with wavelet neural network as the model prediction algorithm and CUSUM control diagram as the cost predictive control law [21]. In general, predictive control in economic system applications will continue to be based on models and algorithms, and on the basis of which innovations and adjustments will be made flexibly, so as to obtain predictive control strategies or mechanisms that are suitable for various practical economic systems.

5 Conclusion

Economic system is an open system with complex structure and interlaced relationship, which is restricted by many factors. Its operation is a complex process of multivariable coupling. The stability and shock of the actual economic system have a direct and important influence on the realization of the intended target. If the stability of the economic system can be judged in advance and the corresponding system can be controlled before, during or after the event, then the economic system can be closer to the desired goal and obtain more economic benefits. In this paper, the predictive control of economic system is summarized, and the establishment and implementation of predictive control model based on cobweb model and nonlinear system are discussed. However, as predictive control

is firstly applied in the industrial field and industrial systems are different from economic systems, there are still many urgent problems in how to establish more accurate or optimized models that adapt to the characteristics of economic systems while ensuring stability. In addition, since the predictive control uses the previous information, there is sometimes a large error between the predictive control and the actual value. What's more, the study of predictive control in economic systems is still more linear, but in the actual economic process, there is often a nonlinear situation. Therefore, how to select optimization strategy, design a wide range of control effects, strong adaptability, strong robustness of predictive control system and establish a practical feedback correction mechanism and strengthen the study of nonlinear economic system is still the direction of in-depth study of predictive control.

References

1. Yang, Q.: Predictive control. In: Automation Panorama, vol. 1, pp. 45–47 (1999)
2. Hu, Y.H., Jia, X.L.: Summarization of generalized predictive control. Inf. Control **3**, 248–256 (2000)
3. Zhou, S.: Lecture on economic cybernetics - the fourth lecture on the characteristics, stability and oscillation of economic systems. Econ. Theory Econ. Manage. **2**, 56–61 (1986)
4. Ding, L., Tong, M., Li, R., Leng, X., Yuan, Y., Yang, L.: Mathematical models were used to predict the wreckage. SME Manage. Technol. **14**, 1–2 (2016)
5. Liu, X., Liu, J.: Predictive control theory is applied to the analysis of unbalanced economic systems. Control Decis. Making **2**, 1018–1019 (2006)
6. Qikun C.: Robust control policies on price adjustment for non equilibrium micro-market. J. Xiamen Univ. (Nat. Sci.) **z1**, 215–218 (2001)
7. Zhang, J.S.: Economic cybernetics of deterministic dynamic systems. Econ. Cybern., 263–264 (1989)
8. Li, C., Yu, M., Chen, X.: Robust control of a class of the disequilibrium economic systems and its application. J. Northeast. Univ. (Nat. Sci.) **3**, 323–326 (2012)
9. Yong, W.: The rude stick control strategy of macroeconomic system. J. Chongqing Inst. Technol. **15**(4), 22–25 (2001)
10. Liu, M., Liu, X., Han, C.: Robust control of macroeconomic system. Yantai Teach. Univ. J. (Nat. Sci. Edn.) **22**(2), 96–98 (2006)
11. Ellis, M., Durand, H., Christofides, P.D.: A tutorial review of economic model predictive control methods. J. Process Control **24**(8), 1156–1178 (2014)
12. Huang, D., Li, H.: The growth and development on the nonlinear economics. Science **4**, 37–39 (1992)
13. He, D.: Predictive control of a stable economic model for constrained nonlinear systems. Acta Automatica Sinica **042**(011), 1680–1690 (2016)
14. Zou, T., Wang, X., Li, S.: Multi-model predictive control for nonlinear systems based on mixed logic. Acta Automatica Sinica **033**(002), 188–192 (2007)
15. Xi, Y.G., Wang, F.: Multi-model method for predictive control of nonlinear systems. Acta Automatica **22**(4), 456–461 (1996)
16. Williams, H.P.: Model Building in Mathematical Program, 3rd edn. Wiley, New York (1993). ming
17. Xu, Y., Zhan, L., Wang, Y.: A universal economic model - a nonlinear system of prediction and control. Syst. Eng. **12**(2), 12–17 (1994)

18. Kao, Y., Li, D., Ling, S.: Model predictive control: status quo and challenges. *Act Automatica* **03**, 28–42 (2013)
19. Zhang, W., Du, Q.: Dynamic hedging European option based on stochastic model predictive control. *J. South China Univ. Technol. (Soc. Sci. Edn.)* **018**(004), 1–9 (2016)
20. Luo, R., Huang, M.: On sustainable development prediction and adjustment model based on adaptive network-based fuzzy inference system. *Training Res.-J. Hubei Coll. Educ.* **22**(1), 55–58 (2005)
21. Liu, Z., Wu, L., Ji, S.: Study on method for cost predictive control. *J. Syst. Eng.* **21**(6), 659–662 (2006)



The Conjugate Gradient Algorithm for Control Systems with a Sine Excitation

Ling Xu^(✉)

School of Internet of Things Technology, Wuxi Vocational Institute of Commerce,
Wuxi 214153, People's Republic of China
lingxu0848@163.com

Abstract. This paper studies the problem of the parameter estimation for control systems by means of the discrete observations under a sine excitation signal. In order to seize the dynamical characteristics of systems and obtain higher estimation accuracy, an objective function by using the dynamical data is constructed and optimized by the negative gradient search. For the purpose of obtaining fast convergence speed, a conjugate gradient algorithm is developed to estimate the system parameters, in which the search direction, i.e., the step-size can be determined in accordance with the variation of the objective function. Finally, a numerical example is provided to test the performance to the proposed method and the simulation results show that the presented algorithm based on the conjugate gradient is effective for process systems.

1 Introduction

Mathematical models are very important in industrial processes, which can determine the performance of control systems [1–6]. Because many control methods depend on system models, system identification has been studied in many areas [7–9]. In terms of most industrial processes, it is difficult to construct their process models by their mechanism. Therefore, the system modeling based on the statistic principle is more effective and can obtain satisfactory models for controller design [10]. In system identification and optimization, the iterative techniques play an important role and used widely in solving system identification and optimization problems. The system modeling based on the statistic principle uses the discrete measured data to construct a criterion function with respect to the system parameters. Generally, the discrete measured data can be collected from the response experiments by applying an excitation signal and the criterion function is built by using these observations. Because the system dynamical information is contained in the observations, the system models can be obtained by optimizing the criterion functions. In this work, we study the control system which can be described by a linear first-order transfer function.

For linear systems, the system responses can be impulse responses, step responses and sine response by employing the impulse signal, step signal and sine

signal. Even though the systems are linear, these system responses are nonlinear. Therefore, the criterion functions are highly nonlinear function with respect to the system parameters. Thus, the off-line optimization strategies of an iterative nature is useful [11]. Moreover, optimizing the nonlinear criterion functions requires the nonlinear optimization such as the gradient optimization, Newton optimization [12–16].

Among the nonlinear optimization, the conjugate gradient method is a method between steepest descent method and Newton method, which only needs the first derivative information. The conjugate gradient method is an effective algorithm to large-scale nonlinear optimization. It has many advantages such as small storage, step convergence, high stability etc. In this work, we introduce the conjugate gradient optimization to develop the identification algorithm under the sine input for the first-order control systems.

2 Problem Description

Consider the following first-order system:

$$G(s) = \frac{K}{Ts + 1},$$

where K is the system gain, T is the time constant.

Suppose that the input signal is $r(t) = A \sin \omega t$, where A is the amplitude of the input, ω is the angular frequency of the input sine signal. For a linear system, when the input is a sine signal the output response is also a sine signal. In the frequency response test, we apply a sine signal with known frequencies and amplitudes to the system to be identified. The output is given by

$$\begin{aligned} y(t) &= B \sin(\omega t + \varphi) \\ &= B \sin \omega t \cos \varphi + B \cos \omega t \sin \varphi \\ &= B \cos \varphi \sin \omega t + B \sin \varphi \cos \omega t, \end{aligned} \tag{1}$$

where B is the amplitude of the output, φ is the phase of the output. According to the definition the frequency characteristic, we have

$$B = |G(j\omega)|A = \frac{KA}{\sqrt{(T\omega)^2 + 1}}, \tag{2}$$

$$\varphi = \arg G(j\omega) = -\arctan T\omega.$$

Let $b_1 := B \cos \varphi$, $b_2 := B \sin \varphi$. Then Eq. (1) becomes

$$y(t) = b_1 \sin \omega t + b_2 \cos \omega t.$$

Suppose that the sampling period is Δt , the corresponding outputs are described by $y(t_i)$, $i = 1, \dots, n$. $t_i = t_1 + n\Delta t$. t_i is sample time.

Define the parameter vector $\boldsymbol{\theta} := [b_1, b_2]^\top$. Define the residual:

$$\varepsilon_i = y(t_i) - b_1 \sin \omega t_i - b_2 \cos \omega t_i.$$

Define a quadratic cost function:

$$J(\boldsymbol{\theta}) := \frac{1}{2} \sum_{i=1}^L \varepsilon_i^2 = \frac{1}{2} \sum_{i=1}^L [y(t_i) - b_1 \sin \omega t_i - b_2 \cos \omega t_i]^2. \quad (3)$$

The parameter estimates can be obtained by minimizing the cost function $J(\boldsymbol{\theta})$. Because the cost function is a nonlinear function, we use the nonlinear optimization to derive the parameter estimation algorithm. In this paper, the nonlinear conjugate gradient optimization is introduced to minimize the cost function.

3 The Conjugate Gradient Parameter Estimation Algorithm

Taking the partial derivation of $J(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ gives the gradient of $J(\boldsymbol{\theta})$

$$\begin{aligned} \mathbf{g}(\boldsymbol{\theta}) &:= \frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} -\sum_{i=1}^L [y(t_i) - (b_1 \sin \omega t_i + b_2 \cos \omega t_i)] \sin \omega t_i \\ -\sum_{i=1}^L [y(t_i) - (b_1 \sin \omega t_i + b_2 \cos \omega t_i)] \cos \omega t_i \end{bmatrix} \\ &= \begin{bmatrix} -\sum_{i=1}^L \varepsilon_i \sin \omega t_i \\ -\sum_{i=1}^L \varepsilon_i \cos \omega t_i \end{bmatrix} \in \mathbb{R}^2. \end{aligned}$$

Let $k = 1, 2, 3, \dots$ be an iteration variable and $\hat{\boldsymbol{\theta}}_k := [b_{1k}, b_{2k}]^\top$ be the iterative estimate of $\boldsymbol{\theta}$. The conjugate gradient iterative estimation formula is given by

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k + \lambda_k \mathbf{d}(\hat{\boldsymbol{\theta}}_k), \quad (4)$$

where λ_k is the step-size and $\mathbf{d}(\hat{\boldsymbol{\theta}}_k)$ is the search direction of the negative gradient at each iteration. Let $\hat{\boldsymbol{\theta}}_0 := [b_{10}, b_{20}]^\top$ be the initial value. Then the search direction is determined by

$$\begin{aligned} \mathbf{d}(\hat{\boldsymbol{\theta}}_0) &= -\mathbf{g}(\hat{\boldsymbol{\theta}}_0) \\ &= \begin{bmatrix} \sum_{i=1}^L [y(t_i) - (b_{10} \sin \omega t_i + b_{20} \cos \omega t_i)] \sin \omega t_i \\ \sum_{i=1}^L [y(t_i) - (b_{10} \sin \omega t_i + b_{20} \cos \omega t_i)] \cos \omega t_i \end{bmatrix}, \quad k = 0, \end{aligned} \quad (5)$$

$$\mathbf{d}(\hat{\boldsymbol{\theta}}_k) = -\mathbf{g}(\hat{\boldsymbol{\theta}}_k) + \beta_k \mathbf{d}(\hat{\boldsymbol{\theta}}_{k-1}), \quad k = 1, 2, \dots, \quad (6)$$

where

$$\mathbf{g}(\hat{\boldsymbol{\theta}}_k) = \begin{bmatrix} \sum_{i=1}^L [y(t_i) - (\hat{b}_{1k} \sin \omega t_i + \hat{b}_{2k} \cos \omega t_i)] \sin \omega t_i \\ \sum_{i=1}^L [y(t_i) - (\hat{b}_{1k} \sin \omega t_i + \hat{b}_{2k} \cos \omega t_i)] \cos \omega t_i \end{bmatrix}. \quad (7)$$

The choice of parameter β_k in (6) is determined by the Fletcher-Reeves method, i.e.,

$$\begin{aligned}\beta_k &= \frac{\mathbf{g}^T(\hat{\boldsymbol{\theta}}_k)\mathbf{g}(\hat{\boldsymbol{\theta}}_k)}{\mathbf{g}^T(\hat{\boldsymbol{\theta}}_{k-1})\mathbf{g}(\hat{\boldsymbol{\theta}}_{k-1})} \\ &= \frac{\left[\sum_{i=1}^L \varepsilon_i(\hat{\boldsymbol{\theta}}_k) \sin \omega t_i\right]^2 + \left[\sum_{i=1}^L \varepsilon_i(\hat{\boldsymbol{\theta}}_k) \cos \omega t_i\right]^2}{\left[\sum_{i=1}^L \varepsilon_i(\hat{\boldsymbol{\theta}}_{k-1}) \sin \omega t_i\right]^2 + \left[\sum_{i=1}^L \varepsilon_i(\hat{\boldsymbol{\theta}}_{k-1}) \cos \omega t_i\right]^2}, \\ \varepsilon_i(\hat{\boldsymbol{\theta}}_k) &= y(t_i) - \hat{b}_{1k} \sin \omega t_i - \hat{b}_{2k} \cos \omega t_i.\end{aligned}\quad (8)$$

The best step-size is determined by solving $\lambda_k = \underset{\lambda \geq 0}{\operatorname{argmin}} J[\hat{\boldsymbol{\theta}}_k + \lambda \mathbf{d}(\hat{\boldsymbol{\theta}}_k)]$. For the purposes of convenient analysis, define $\mathbf{d}(\hat{\boldsymbol{\theta}}_k) := [d_{1k}, d_{2k}]^T$. Then we have

$$J[\hat{\boldsymbol{\theta}}_k + \lambda \mathbf{d}(\hat{\boldsymbol{\theta}}_k)] = \frac{1}{2} \sum_{i=1}^L [y(t_i) - (\hat{b}_{1k} + \lambda \hat{d}_{1k}) \sin \omega t_i - (\hat{b}_{2k} + \lambda \hat{d}_{2k}) \cos \omega t_i]^2.$$

Taking the first-order partial derivation of $J[\hat{\boldsymbol{\theta}}_k + \lambda \mathbf{d}(\hat{\boldsymbol{\theta}}_k)]$ with respect to λ yields

$$\begin{aligned}\frac{\partial J[\hat{\boldsymbol{\theta}}_k + \lambda \mathbf{d}(\hat{\boldsymbol{\theta}}_k)]}{\partial \lambda_k} &= \frac{\frac{1}{2} \partial \sum_{i=1}^L [y(t_i) - (\hat{b}_{1k} + \lambda \hat{d}_{1k}) \sin \omega t_i - (\hat{b}_{2k} + \lambda \hat{d}_{2k}) \cos \omega t_i]^2}{\partial \lambda_k} \\ &= - \sum_{i=1}^L [y(t_i) - (\hat{b}_{1k} + \lambda \hat{d}_{1k}) \sin \omega t_i - (\hat{b}_{2k} + \lambda \hat{d}_{2k}) \cos \omega t_i] \times \\ &\quad (\hat{d}_{1k} \sin \omega t_i + \hat{d}_{2k} \cos \omega t_i).\end{aligned}$$

Let

$$\frac{\partial J[\hat{\boldsymbol{\theta}}_k + \lambda \mathbf{d}(\hat{\boldsymbol{\theta}}_k)]}{\partial \lambda_k} = 0.$$

Then the optimal step-size λ_k is given by

$$\lambda_k = \frac{\sum_{i=1}^L [y(t_i) - \hat{b}_{1k} \sin \omega t_i - \hat{b}_{2k} \cos \omega t_i]}{\sum_{i=1}^L (\hat{d}_{1k} \sin \omega t_i + \hat{d}_{2k} \cos \omega t_i)}. \quad (9)$$

For the conjugate gradient parameter estimation algorithm, the parameters are not the system parameters. In fact, the system parameters to be estimated are the gain K and time constant T .

From $b_1 := B \cos \varphi$, $b_2 := B \sin \varphi$, we have

$$\frac{b_2}{b_1} = \frac{B \sin \varphi}{B \cos \varphi} = \tan \varphi = -\tan \arctan T\omega = -T\omega. \quad (10)$$

$$b_1^2 + b_2^2 = B^2 \sin^2 \varphi + B^2 \cos^2 \varphi = B^2 (\sin^2 \varphi + \cos^2 \varphi) = B^2. \quad (11)$$

Substituting (2) into (11) gives

$$b_1^2 + b_2^2 = \frac{K^2 A^2}{T^2 \omega^2 + 1}. \quad (12)$$

Equations (10)–(12) show the following relations, i.e.,

$$K = \frac{b_1^2 + b_2^2}{Ab_1}, \quad T = -\frac{b_2}{b_1\omega}.$$

As a result, the system parameter estimates are given by

$$\hat{K}_k = \frac{\hat{b}_{1k}^2 + \hat{b}_{2k}^2}{A\hat{b}_{1k}}, \quad \hat{T}_k = -\frac{\hat{b}_{2k}}{\hat{b}_{1k}\omega}. \quad (13)$$

The steps of implementing the conjugate gradient iterative algorithm are listed as follows:

1. To initialize: let $k = 0$, $\hat{\theta}_0 = [b_{10}, b_{20}]^T$ be a random vector, and give a small number $\varepsilon > 0$. Set the data length L .
2. Collect the measured data $\{t_i, y(t_i)\}$, $i = 1, 2, 3, \dots, L$.
3. Compute $\mathbf{g}(\hat{\theta}_0)$ using (5). Compute $\mathbf{d}(\hat{\theta}_0) = -\mathbf{g}(\hat{\theta}_0)$ using (6). Compute λ_0 using (9).
4. Compute $\mathbf{g}(\hat{\theta}_k)$ using (7). Compute β_k using (8). Compute $\mathbf{d}(\hat{\theta}_k)$ using (5).
5. Update λ_k using (8). Update the parameter estimate $\hat{\theta}_{k+1}$ by (4). If $\|\hat{\theta}_{k+1} - \hat{\theta}_k\| < \varepsilon$, terminate the procedure, obtain the parameter estimate $\hat{\theta}_{k+1}$ and compute the system parameter estimates using (13); otherwise $k := k + 1$ and go to step 4.

4 Numerical Example

Consider a dynamical system:

$$G(s) = \frac{4}{5s + 1},$$

where the true parameters are $K = 4$ and $T = 5$.

In order to set the input signal, we compute the cutoff frequency of the system to be identified. As we all know, the cutoff frequency of systems ω_b refers to the frequency when the logarithmic amplitude frequency characteristic of the closed-loop system drops to 3 dB frequency below its zero amplitude frequency, i.e.,

$$20 \lg M(\omega_b) = 20 \lg M(0) - 3.$$

In the frequency test, we apply a sine signal to the system in the bandwidth range. The system bandwidth refers to the frequency range from 0 to the cutoff frequency. For the system in example 1, the cutoff frequency is computed as $\omega_b = 1/T = 1/5 = 0.2$ rad/s. Therefore the input is set as $r(t) = 2 \sin 0.1t$. In the simulation experiment, the following cases are considered to test the performance of the proposed method.

Case 1. Different noise variance

In this case, the sample period is set as 1s and the response time is $t = 1000$ s. In the real control condition, many disturbance signals are contained

Table 1. The parameter estimates and their estimation errors $t = 1000$

σ^2	k	K	T	δ
0.10^2	1	3.99276	4.96075	0.00623
	2	3.98798	4.93554	0.01024
	3	3.98796	4.93555	0.01024
	4	3.98796	4.93555	0.01024
	5	3.98796	4.93555	0.01024
	6	3.98796	4.93555	0.01024
0.50^2	1	3.99643	4.92302	0.01204
	2	3.99178	4.89799	0.01598
	3	3.99176	4.89799	0.01598
	4	3.99176	4.89799	0.01598
	5	3.99176	4.89799	0.01598
	6	3.99176	4.89799	0.01598
4.00^2	1	4.08614	4.89932	0.02069
	2	4.07775	4.89934	0.01986
	3	4.08523	4.89758	0.02081
	4	4.08538	4.89754	0.02083
	5	4.08531	4.89756	0.02082
	6	4.08531	4.89756	0.02082
True values		4.00000	5.00000	

in the measured data. In order to simulate practical conditions, we take the unrelated white noise with noise variance $\sigma = 0.10^2$, $\sigma = 0.50^2$ and $\sigma = 4.00^2$ as the disturbance signal respectively. Using the presented algorithm to estimate the system parameters, the parameter estimates and their estimation errors are shown in Table 1. The estimation errors $\delta := \|\hat{\theta}_k - \theta\|/\|\theta\|$ versus k are shown in Fig. 1.

Case 2. Different response time

In order to study the precision of the parameter estimates in different response time during the frequency response test, we take a disturbance with the noise variance $\sigma^2 = 0.40^2$ to apply to the system. The response time is set as $t = 1000$ s and $t = 2000$ s respectively. Using the presented algorithm to estimate the system parameters, the parameter estimates and their estimation errors are shown in Table 2. The parameter estimation errors versus k are shown in Fig. 2.

Case 3 Model validation

Selecting the estimated parameters with the noise variance $\sigma^2 = 0.40^2$ and $t = 1000$ s obtains the estimated system model

$$\hat{G}(s) = \frac{3.99080}{4.90736s + 1}.$$

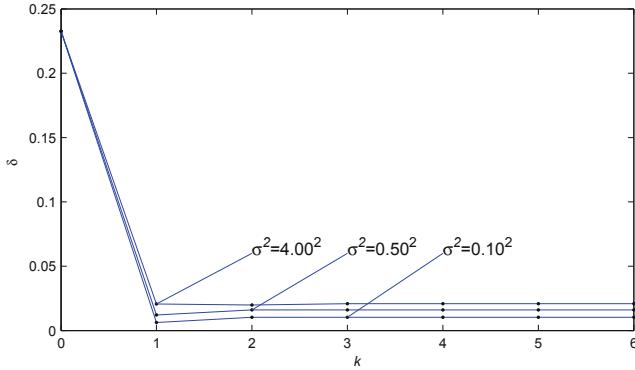


Fig. 1. The parameter estimation errors δ versus k for the first-order system

Table 2. The parameter estimates and their estimation errors ($\sigma^2 = 0.40^2$)

t	k	K	T	δ
1000	1	3.99546	4.93244	0.01058
	2	3.99082	4.90735	0.01454
	3	3.99080	4.90736	0.01454
	4	3.99080	4.90736	0.01454
	5	3.99080	4.90736	0.01454
	6	3.99080	4.90736	0.01454
2000	1	4.00365	4.99088	0.00153
	2	3.99883	4.96562	0.00537
	3	3.99881	4.96563	0.00537
	4	3.99881	4.96563	0.00537
	5	3.99881	4.96563	0.00537
	6	3.99881	4.96563	0.00537
True values		4.00000	5.00000	

For the sake of validating the accuracy of the estimated model, we draw the Bode diagrams of the estimated model and true model. The Bode diagrams are shown in Fig. 3, where the dot-line is the Bode diagram of estimated model and the solid-line is the Bode diagram of the true model.

From the simulation results, we can summarize the following conclusions.

- The parameter estimation errors become smaller very fast with the increasing of iteration k under different noise variances, which means that the proposed method is effective.
- The estimation accuracy is related to the datum length t . The observations with large datum length are taken in the computation will lead to more accurate parameter estimates.

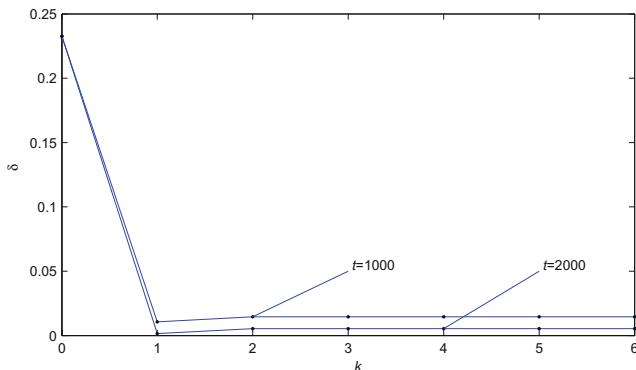


Fig. 2. The estimation errors δ versus k for the first-order system

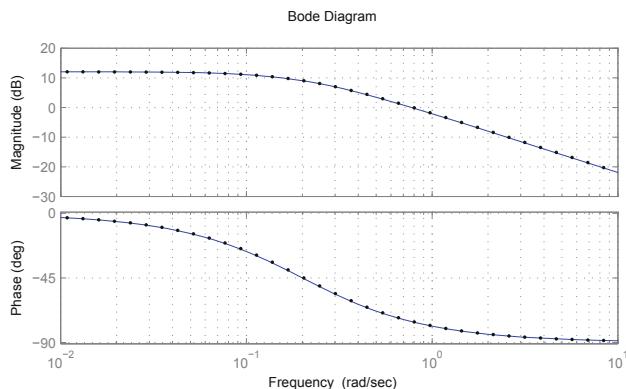


Fig. 3. The Bode diagrams of the estimated model and real model

- From the Bode diagrams of the estimated model and real model, we can see that they are very close, which means that the estimated model can be used to capture the dynamical characteristic of the system.

Remark 1: It is noted that multiple frequency points should be adopted in the range of the system bandwidth during the frequency characteristic test process in order to assure the accuracy of the parameter estimate. The parameter estimates are obtained by calculating the average values at each test.

Remark 2: For the sine signal with single frequency used as the input, the measured data should be collected when the output responses are stable, otherwise the measure data will be imprecise. For slow processes, the single frequency sine signal will take very long time to obtain the measured data using the frequency characteristic test.

5 Conclusions

This paper presents a conjugate gradient based on iterative parameter estimation algorithm by constructing an objective function based on the off-line observations in terms of the first-order system. The simulation results show that the proposed method is effective for solving the identification problem for control systems.

Acknowledgments. This work was supported by Qing Lan Project of Jiangsu Province, by the “333” Project of Jiangsu Province (No. BRA2018328), by Jiangsu Overseas Visiting Scholar Program for University Prominent Young and Middle-aged Teachers and Presidents.

References

1. Ding, F.: System Identification - New Theory and Methods. Science Press, Beijing (2013)
2. Ding, F.: System Identification - Performance Analysis for Identification Methods. Science Press, Beijing (2014)
3. Ding, F.: System Identification - Auxiliary Model Identification Idea and Methods. Science Press, Beijing (2017)
4. Ding, F.: System Identification - Iterative Search Principle and Identification Methods. Science Press, Beijing (2018)
5. Ding, F.: System Identification - Multi-Innovation Identification Theory and Methods. Science Press, Beijing (2016)
6. Ding, F.: Modern Control Theory. Tsinghua University Press, Beijing (2018)
7. Wang, Y.J., Ding, F.: Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model. *Automatica* **71**, 308–313 (2016)
8. Ding, F., Liu, G., Liu, X.P.: Partially coupled stochastic gradient identification methods for non-uniformly sampled systems. *IEEE Trans. Autom. Control* **55**(8), 1976–1981 (2010)
9. Ding, J., Ding, F., Liu, X.P., Liu, G.: Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data. *IEEE Trans. Autom. Control* **56**(11), 2677–2683 (2011)
10. Ding, F.: Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling. *Appl. Math. Model.* **37**(4), 1694–1704 (2013)
11. Ding, F.: Several multi-innovation identification methods. *Digital Signal Process.* **20**(4), 1027–1039 (2020)
12. Xu, L., Ding, F., Zhu, Q.M.: Hierarchical Newton and least squares iterative estimation algorithm for dynamic systems by transfer functions based on the impulse responses. *Int. J. Syst. Sci.* **50**(1), 141–151 (2019)
13. Ding, F.: Combined state and least squares parameter estimation algorithms for dynamic systems. *Appl. Math. Model.* **38**(1), 403–412 (2014)
14. Ding, F.: State filtering and parameter estimation for state space systems with scarce measurements. *Signal Process.* **104**, 369–380 (2014)
15. Ding, J., Chen, J.Z., Lin, J.X., Jiang, G.P.: Particle filtering-based recursive identification for controlled auto-regressive systems with quantised output. *IET Control Theory Appl.* **13**(14), 2181–2187 (2019)
16. Chen, J., Shen, Q.Y., Ma, J.X., Liu, Y.J.: Stochastic average gradient algorithm for multirate FIR models with varying time delays using self-organizing maps. *Int. J. Adap. Control Signal Process.* **34**(7), 955–970 (2020)



Privacy-Preserving Consensus of Continuous Multi-agent Systems with Communication Delay

Yize Yang^{1,2}, Hongyong Yang^{1(✉)}, Meiyuan Yu¹, and Yujiao Sun¹

¹ School of Information and Electrical Engineering,
Ludong University, Yantai 264025, People's Republic of China
hyyang@ldu.edu.cn

² School of Electrical Engineering and Telecommunications,
University of New South Wales, Sydney, Australia

Abstract. The privacy-preserving consensus problem for the continuous multi-agent systems with communication delay is investigated in this paper, and a novel privacy-preserving strategy is proposed to protect all initial state of agents from being acquired by other agents. Then, a consensus algorithm based on this privacy-preserving method is designed. Different from other protection of private information methods, the proposed privacy-preserving method can not only ensure that movement state of agents accurately converge to the stable position but cover their true motion path. According to the designed control algorithm, we conduct a rigorous mathematical analysis for the consensus problem of MAS. Finally, data simulations verifies the effectiveness of the proposed privacy-preserving consensus algorithm.

Keywords: Privacy-preserving · Communication delay · Multi-agent systems · Consensus control

1 Introduction

In this age of data explosion, it is especially important to cover the privacy of individual information. The privacy protection of data information and the method of privacy protection has become a hot topic of research. In recent years, researchers in the control field of multi-agent systems (MAS) have gradually applied the concept of privacy protection to the consensus problem of MAS [1,2]. The privacy-preserving is that the MAS can make the initial state of multi-agents indiscernible by others. If a agent in MAS knows the movement information for other agents, it can infer the state of all agents under certain observable conditions. This also means that the true initial state of the agent can be calculated by other agents, which leads to the exposure of the state information of all agents. Therefore, it is necessary to consider the privacy-preserving problem in the consensus algorithm of MAS, and it is challenging to design a new consensus control algorithm to solve the privacy-preserving problem in MAS.

The theme of privacy protection originated in the field of database security and attracted wide attention from the academic circles. At present, many researchers have begun to study the privacy-preserving consensus of MAS, which means that all agents in MAS will achieve consensus state, but also ensure the privacy of the state of agents. In [3–5], researchers studied the average consensus of privacy protection in discrete MAS. In [3], an optimal privacy-preserving protocol is proposed to achieve average consensus and the high-quality data protection results. To ensure the privacy of the initial state and the asymptotic consensus of precise increment of the initial value. A consensus protocol is proposed in [4] to protect personal privacy by considering and randomizing noise during the movement process. To expand its research, an sufficient differentially private protocol is proposed in [5], and the maximum consensus of multi-nodes systems with Laplacian noises is studied.

Differential privacy [6–8], as an existing method of privacy-preserving, has been widely concerned by researchers because of its rigorous formulaic form and proven security characteristics, and also independence, flexibility for auxiliary information and variability for post-processing from the adversary dynamics. For example, the control problem of the average consensus for the discrete-time dynamic MAS with differentially private strategy is investigated in [2]. In [6], the analytical expression for dynamic MAS with a novel maximum consensus algorithm is given and the differential privacy is considered.

Then, the privacy-preserving average consensus problem of MAS is investigated in [7]. In [8], the multi-agent average consensus problem is investigated. In this paper, the adversary of agent has access to all the messages under the conditions of differential privacy in the initial states of agent.

The differential privacy method is only applicable to discrete MAS, and cannot achieve accurate consensus of MAS. On the other hand, the traditional privacy-preserving methods are implemented by considering the consensus algorithm with random noise. Thus, the research results obtained in [7] are all proved by statistics in practical engineering with random noise, which leads to a huge amount of calculation. So we are driven to propose a novel privacy-preserving method instead of adding random noise.

In order to solve some problems that cannot be solved by existing privacy methods, a novel privacy-preserving method is proposed for the consensus problem of continuous MAS. And the privacy-preserving consensus control algorithm ensure that the agents accurately converge to a common value instead of converging in a mean square.

2 Preliminaries

Considered the network $g = (V, \omega, A)$ be a weighted undirected network topology with n nodes. $V = \{1, 2, \dots, n\}$ is a collection of communication vertices. $\omega \subseteq V \times V$ is a collection of communication edges. $A = [a_{ij}] \in R^{n \times n}$ is the connection adjacency matrix of the nodes with $a_{ij} \geq 0$, where $a_{ij} > 0$ if $(i, j) \in \omega$, otherwise $a_{ij} = 0$; there is $a_{ii} = 0$ for $\forall i \in V$. The Neighbor collection of node i is defined

as $N_i = \{j \in V | (i, j) \in \omega\}$. Let the matrix $D = \text{diag}\{d_1, d_2, \dots, d_n\} \in R^{n \times n}$ be degree matrix of the graph g , where $d_i = \sum_{j \in N_i} a_{ij}$, for $i = 1, 2, \dots, n$. The Laplacian matrix of Topology diagram g is defined as $L = D - A \in R^{n \times n}$, where D is the degree matrix corresponding to the topological structure between the intelligent agents.

For an undirected connected network topology, the eigenvalue of the connection matrix A satisfying $\lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_n(A)$, where the eigenvalue of Laplacian matrix L satisfying $\lambda_1(L) \geq \lambda_2(L) \geq \dots \geq \lambda_n(L) = 0$.

Assumption 1. Considered the communication topology composed of multi-agents in this paper is undirected and connected. And the weighted connection matrix A satisfies $A \cdot 1 = 1$.

Remark 1. According to Assumption 1, the connection matrix A is symmetric matrix with $\lambda_n(A) = 1$ and L is a positive semi-definite matrix with $\lambda_n(L) = 0$.

Definition 1. Function $h_i(t, x_i, \pi_i)$ is a function with the ability to cover the internal state of the agent, if the following conditions are satisfying: C1: $h_i(0, x_i, \pi_i) \neq x_i, \forall x_i \in R^n, i = 1, 2, \dots, n$.

C2: $h_i(0, x_i, \pi_i)$ guarantees unrecognizable initial conditions

C3: $h_i(0, x_i, \pi_i)$ cannot protect the nearby value of agent state $\forall x_i \in R^n$.

C4: If given t and π_i , $h_i(0, x_i, \pi_i)$ increases monotonically with x_i .

C5: If given t and π_i , $|h_i(0, x_i, \pi_i) - x_i|$ is monotonically decreasing, and $\lim_{t \rightarrow \infty} h_i(0, x_i, \pi_i) = x_i, i = 1, 2, \dots, n$.

Then, we design the privacy-preserving function:

$$y_i(t) = h_i(t, x_i, \pi_i)$$

$$h_i(t, x_i, \pi_i) = \left(1 + \frac{\phi_i}{1 + at}\right) \left(x_i(t) + \frac{\gamma_i}{1 + at}\right)$$

where $y_i(t)$ is the output state of function $h_i(t, x_i, \pi_i)$. And $\pi_i = \{\phi_i, \gamma_i\}$, $\phi_i, \gamma_i > 0$, $a > 0$.

Assumption 2. The input $u_i(t)$ and output $y_i(t)$ are public and known in the MAS. The real state $x_i(t)$ and function $h_i(t, x_i, \pi_i)$ are private and unknowable to each agent.

3 Consensus Control and Privacy-Preserving of the Multi-agent Systems

Based on the proposed privacy-preserving function and Definition 1, we propose the privacy-preserving consensus control algorithm as follow:

$$\dot{x}_i(t) = x_i(t) + u_i(t) \quad (1)$$

$$u_i(t) = c \sum_{j \in N_i} a_{ij} (y_j(t - \tau) - y_j(t)) \quad (2)$$

$$y_i(t) = h_i(t, x_i, \pi_i) \quad (3)$$

$$h_i(t, x_i, \pi_i) = \left(1 + \frac{\phi_i}{1+at}\right) \left(x_i(t) + \frac{\gamma_i}{1+at}\right) \quad (4)$$

where $x_i(t)$ denotes the internal state of agent i . $h_i(t, x_i, \pi_i)$ is the privacy-preserving function. $y_i(t)$ denotes the output state of function $h_i(t, x_i, \pi_i)$. $u_i(t)$ is the input of the MAS. τ is communication time-delay between agents. $c > 0$ is control gain. Next, we analyze the consensus control and privacy-preserving of the MAS.

Lemma 1. Function $h_i(t, x_i, \pi_i) = \left(1 + \frac{\phi_i}{1+at}\right) \left(x_i(t) + \frac{\gamma_i}{1+at}\right)$ is a privacy-preserving function that masks the internal real state $x_i(t)$ of agent i .

Proof. Since $h_i(0, x_i, \pi_i) = (1 + \phi_i)(x_i(0) + \gamma_i) \neq x_i(0)$, So C1 in Definition 1 is satisfied.

In the privacy-preserving function, the information of $\dot{x}_i(t)$ and $y_i(t)$ cannot determine the parameters ϕ_i and γ_i . Therefore, we cannot draw the image of $h_i(t, x_i, \pi_i)$ unilaterally so that we cannot recognize the initial conditions of the agent. That is, C2 is satisfied.

If x^+ is any point in R^n and satisfies $\|x(0) - x^+\| < \xi$, then $\|h(0, x, \pi) - x^+\| = \|(I + \Phi)(x + \Upsilon)\| \leq \|(I + \Phi)x - x^+\| + \|(I + \Phi)\Upsilon\|$, where $\Phi = \text{diag}(\phi_1, \dots, \phi_n), \Upsilon = [\gamma_1, \dots, \gamma_n]^T$. The inequality does not belong to the ξ neighborhood of x^+ . So C3 is satisfied.

Since $h_i(t, x_i, \pi_i) = \left(1 + \frac{\phi_i}{1+at}\right) \left(x_i(t) + \frac{\gamma_i}{1+at}\right)$, then $\frac{\partial h_i(t, x_i, \pi_i)}{\partial x_i} = \left(1 + \frac{\phi_i}{1+at}\right) > 0$. This means that $h_i(t, x_i, \pi_i)$ is monotonically increasing with respect to x_i . That is, C4 is satisfied.

Since

$$|h_i(t, x_i, \pi_i) - x_i| = \frac{\phi_i}{1+at} x_i(t) + \frac{\phi_i \gamma_i}{(1+at)^2},$$

this function is an increasing function. In addition, $\lim_{t \rightarrow \infty} h_i(t, x_i, \pi_i) = x_i$. So C5 is satisfied.

In summary, according to Definition 1, $h_i(t, x_i, \pi_i) = \left(1 + \frac{\phi_i}{1+at}\right) (x_i(t) + \frac{\gamma_i}{1+at})$ is a function with privacy-preserving performance.

Theorem 1. Under Assumption 1 and Assumption 2, if the communication time-delay satisfy $\tau = (0, \tau^0)$, $\tau = \frac{\ln \rho}{\rho}$. The privacy-preserving control algorithm (2) can ensure the asymptotic consensus of MAS (1).

Proof. According to system dynamics (1), its vector form is as follows

$$\begin{aligned}\dot{x}(t) = & \left[1 - c\left(1 + \frac{\phi}{1+at}\right)\right]x(t) + c\left(1 + \frac{\phi}{1+a(t-\tau)}\right)x(t-\tau) \\ & + c\left(1 + \frac{\phi}{1+a(t-\tau)}\right)\frac{\gamma}{1+a(t-\tau)} - c\left(1 + \frac{\phi}{1+at}\right)\frac{\gamma}{1+at}\end{aligned}\quad (5)$$

where $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$, $\phi(t) = [\phi_1(t), \phi_2(t), \dots, \phi_n(t)]^T$, $\gamma(t) = [\gamma_1(t), \gamma_2(t), \dots, \gamma_n(t)]^T$.

Integrate (5), then we can get

$$x(t) = \int_{t_0}^t \left[1 - c\left(1 + \frac{\phi}{1+as}\right)\right]dsx(t_0) + c \int_{t_0}^t F(\theta)d\theta \quad (6)$$

where $F(\theta) = c\left(1 + \frac{\phi}{1+a(\theta-\tau)}\right)x(\theta-\tau) + c\left(1 + \frac{\phi}{1+a(\theta-\tau)}\right)\frac{\gamma}{1+a(\theta-\tau)} - c\left(1 + \frac{\phi}{1+a\theta}\right)\frac{\gamma}{1+a\theta}$.

Thus,

$$\|x(t)\| = \left\| \int_{t_0}^t \left[1 - c\left(1 + \frac{\phi}{1+as}\right)\right]dsx(t_0) \right\| + \left\| c \int_{t_0}^t F(\theta)d\theta \right\|. \quad (7)$$

Note that the following equation holds,

$$\left\| \int_{t_0}^t \left[1 - c\left(1 + \frac{\phi}{1+as}\right)\right]dsx(t_0) \right\| \leq (1-c)(t-t_0) - \frac{\phi c}{a} \ln\left(\frac{1+at}{1+at_0}\right) \quad (8)$$

$$\left\| c \int_{t_0}^t \left(1 + \frac{\phi}{1+a(\theta-\tau)}\right) \frac{\gamma}{1+a(\theta-\tau)} d\theta \right\| \leq \frac{\gamma c}{a} \ln\left(\frac{1+at}{1+at_0}\right) + \phi\gamma c \left(\frac{1}{1+at_0} - \frac{1}{1+at}\right) \quad (9)$$

Substituting (8) and (9) into (7), it gives

$$\|x(t)\| \leq (1-c)(t-t_0) - \frac{\phi c}{a} \ln\left(\frac{1+at}{1+at_0}\right) + \frac{\gamma c}{a} \ln\left(\frac{1+at}{1+at_0}\right) + \phi\gamma c \left(\frac{1}{1+at_0} - \frac{1}{1+at}\right) \quad (10)$$

Therefore, The MAS can converge to a common value and cover the true motion path with the proposed privacy-preserving method.

4 Simulation Examples

In this part, it mainly verifies the effectiveness of the privacy-preserving consensus algorithm (2) and privacy-preserving function (4). Consider an undirected topology network composed of five agents as shown in Fig. 1. If the weights of lines in the topology diagram are equal to 1, thus the system's analogous Laplacian matrix is

$$L = \begin{pmatrix} 2 & 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 1 & 0 \\ 1 & 1 & 4 & 1 & 1 \\ 0 & 1 & 1 & 3 & 1 \\ 0 & 0 & 1 & 1 & 2 \end{pmatrix}$$

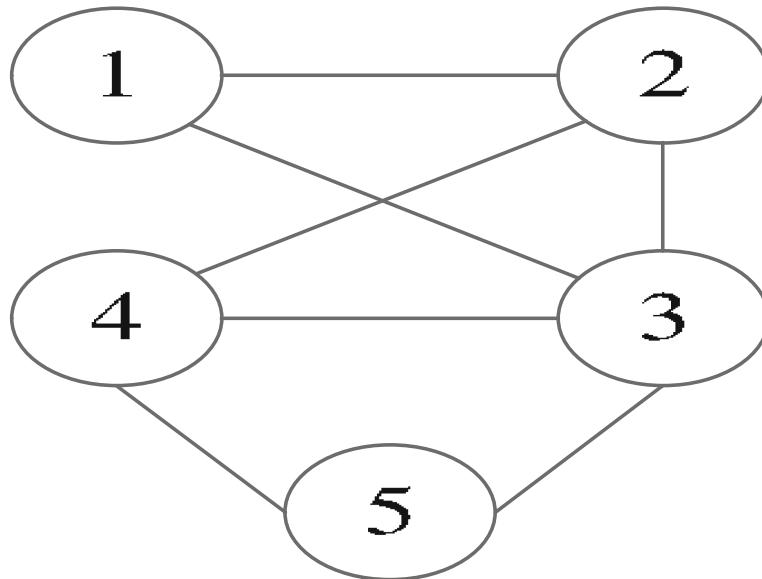


Fig. 1. Communication topology composed of 5 agents

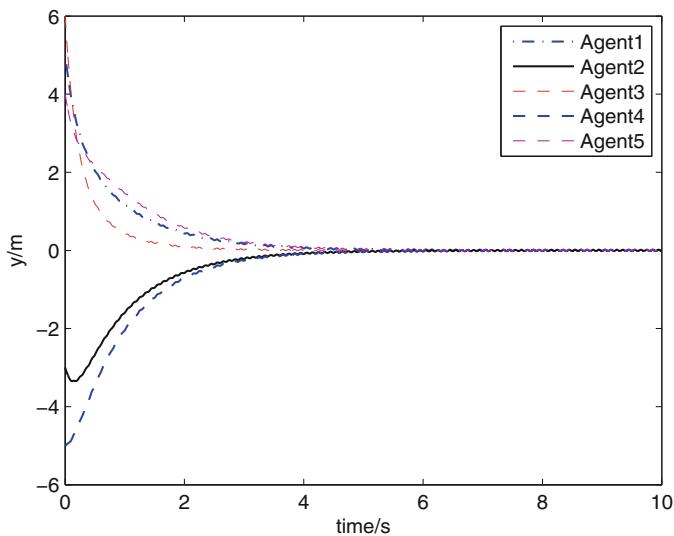


Fig. 2. The trajectory of the agent's movement state

The parameters in the control algorithm are set to $a = 1$, $c = 0.1$, $\phi_i = 0.3$ and $\gamma_i = 0.7$. Thus we got Fig. 2 and Fig. 3 are movement states and privacy-preserving outputs of the agents respectively.

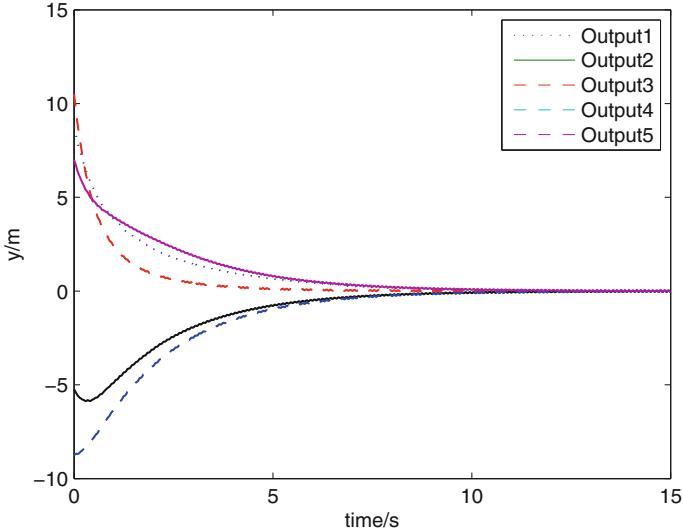


Fig. 3. The trajectory of the agent's privacy-preserving output

Figure 2 is the trajectory of the internal state $x_i(t)$ of agents. It can be seen that all agents finally converges to a same value. Figure 3 is the trajectory of $y_i(t)$. Therefore, we can conclude that $x_i(t)$ and $y_i(t)$ tend to the same value from Fig. 2 and Fig. 3. The proposed privacy-preserving method can not only ensure that all agents accurately converge to a common value but cover their true motion path is proven.

5 Conclusion

The privacy-preserving consensus problem for the continuous multi-agent systems with communication delay is studied in this paper. A new privacy-preserving method is proposed, which is a function constructed to protect the real state of all agents. Based on the proposed privacy-preserving consensus control algorithm, the consensus conditions of multi-agent systems with communication delay are obtained.

Acknowledgments. The research is supported by the National Natural Science Foundation of China (61673200), the Major Basic Research Project of Natural Science Foundation of Shandong Province of China (ZR2018ZC0438) and the Key Research and Development Program of Yantai City of China (2019XDHZ085).

References

1. Le Ny, J., Pappas, G.: Differentially private filtering. *IEEE Trans. Autom. Control* **59**(2), 341–354 (2014). <https://doi.org/10.1109/tac.2013.2283096>
2. Wang, A., Liao, X., He, H.: Event-triggered differentially private average consensus for multi-agent network. *IEEE/CAA J. Automatica Sinica* **6**(1), 75–83 (2019). <https://doi.org/10.1109/jas.2019.1911327>
3. He, J., Cai, L., Zhao, C., Cheng, P., Guan, X.: Privacy-preserving average consensus: privacy analysis and algorithm design. *IEEE Trans. Signal Inf. Process. Netw.* **5**(1), 127–138 (2019). <https://doi.org/10.1109/tsipn.2018.2866342>
4. Mo, Y., Murray, R.: Privacy preserving average consensus. *IEEE Trans. Autom. Control* **62**(2), 753–765 (2017). <https://doi.org/10.1109/tac.2016.2564339>
5. Wang, X., He, J., Cheng, P., Chen, J.: Differentially private maximum consensus. *IFAC-Papers OnLine* **50**(1), 9509–9514 (2017). <https://doi.org/10.1016/j.ifacol.2017.08.1597>
6. Wang, X., He, J., Cheng, P., Chen, J.: Differentially private maximum consensus: design, analysis and impossibility result. *IEEE Trans. Netw. Sci. Eng.* **6**(4), 928–939 (2019). <https://doi.org/10.1109/tnse.2018.2879795>
7. Nozari, E., Tallapragada, P., Corts, J.: Differentially private average consensus with optimal noise selection. *IFAC-Papers OnLine* **48**(22), 203–208 (2015). <https://doi.org/10.1016/j.ifacol.2015.10.331>
8. Nozari, E., Tallapragada, P., Corts, J.: Differentially private average consensus: obstructions, trade-offs, and optimal algorithm design. *IEEE Trans. Autom. Control* **81**, 221–231 (2017). <https://doi.org/10.1109/j.automatica.2017.03.016>



Output Feedback Stabilization for a Class of Nonlinear Systems with Unknown Output Function

Jiao Dai and Junyong Zhai^(✉)

School of Automation, Southeast University, Nanjing 210096, Jiangsu, China
jyzhai@seu.edu.cn

Abstract. This article proposes an output feedback control method for a class of nonlinear systems in non-triangular form with unknown output function. A state observer and observer-based controller are designed to globally stabilize the nominal system. Then, the coordinate transformation is introduced into the original nonlinear system to make its form consistent with the nominal system. On this basis, a homogeneous observer and controller are designed to regulate the nonlinear system states to zero. Finally, the correctness of the theoretical derivation is verified by an example.

Keywords: Unknown growth rate · Output feedback · Unknown output function

1 Introduction

In this article, the nonlinear system is described by

$$\begin{aligned}\dot{x}_i &= x_{i+1} + \phi_i(t, x, u), i = 1, \dots, n-1 \\ \dot{x}_n &= u + \phi_n(t, x, u) \\ y &= h(x_1)\end{aligned}\tag{1}$$

where $x = (x_1, \dots, x_n)^T \in R^n$, $y \in R$ and $u \in R$ are system states, output and control input, respectively. The continuous nonlinear term $\phi_i(\cdot) : R \times R^n \times R^n \times R \rightarrow R$ is a nonlinear function and the uncertain output function $h(\cdot)$ satisfies $h(0) = 0$.

Because of disturbances and errors, many industrial processes appear as nonlinear systems. Researchers usually adopt state-feedback method or output feedback method to deal with the nonlinear system, such as [1–3] and the references therein. Some restrictions should be imposed to the nonlinear term $\phi(\cdot)$ because the separation principle is not suitable in nonlinear systems such as [4–9]. To accommodate more general nonlinear systems, homogeneous domination approach is developed to cope with the nonlinear terms with homogeneous growth

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 257–265, 2021.

https://doi.org/10.1007/978-981-15-8458-9_28

condition in [10] and [11] with known growth rates. When the growth rate is unknown, [12–14] propose a universal control scheme to solve the problem. When dealing with the nonlinear systems, we need to obtain the exact information of the state x_1 , such as [15]. However, the relationship between x_1 of the system and the sensors output is always nonlinear, changeable and uncertain in practical control systems. In [16], the output function is assumed to satisfy Lipschitz condition. To handle the unknown output function, the work [17] supposed the differentiation of the output function was bounded.

In this article, we discuss the problem of global adaptive output feedback stabilization for nonlinear systems in non-triangular form. Different from [17], the growth rate of the nonlinear terms here is unknown. First, an observer and output feedback controller are constructed for the nominal system. Then, the homogeneous observer and controller are introduced after the transformation of coordinates. Finally, we prove the system state can converge to the origin asymptotically.

2 Preliminaries and Problem Statement

Lemma 1 [20]. *Suppose that $V(x) : R^n \rightarrow R$ is a homogeneous function of degree τ with respect to Δ . Then $\partial V / \partial x_i$ is homogeneous of degree $\tau - r_i$.*

Lemma 2. *Suppose c, d and γ are positive numbers and $x \in R$, $y \in R$. Then, one has*

$$|x|^c |y|^d \leq \frac{c}{c+d} \gamma |x|^{c+d} + \frac{d}{c+d} \gamma^{-\frac{c}{d}} |y|^{c+d} \quad (2)$$

Lemma 3 [21]. *In terms of the following system*

$$\begin{aligned} \dot{z}_i &= z_{i+1}, i = 1, \dots, n-1 \\ \dot{z}_n &= \nu \\ y &= h(z_1) \end{aligned} \quad (3)$$

there exists an output feedback controller of the form

$$\begin{aligned} \dot{\hat{z}}_i &= \hat{z}_{i+1} - a_i \hat{z}_1^{r_{i+1}}, i = 1, \dots, n-1 \\ \dot{\hat{z}}_n &= \nu - a_n \hat{z}_1^{r_{n+1}} \\ \nu &= -\beta_n (\hat{z}_n^{\frac{r_{n+1}}{r_n}} + \beta_{n-1}^{\frac{r_{n+1}}{r_n}} (\hat{z}_{n-1}^{\frac{r_{n+1}}{r_{n-1}}} + \dots + \beta_2^{\frac{r_{n+1}}{r_3}} (\hat{z}_2^{\frac{r_{n+1}}{r_2}} + \beta_1^{\frac{r_{n+1}}{r_2}} y^{r_{n+1}}) \dots)) \end{aligned} \quad (4)$$

where $\beta_i > 0$ are constants and r_i is defined as

$$r_1 = 1, r_{i+1} = r_i + \tau, i = 1, \dots, n \quad (5)$$

such that the states in system (3) and (4) can converge to zero.

In this paper, we suppose $\tau > 0$. The specific definition of τ can be found in [21]. Define the following vector:

$$Z := [e_1, \dots, e_n, y, \hat{z}_2, \dots, \hat{z}_n]^T \quad (6)$$

As a matter of fact, choosing the dilation weight

$$\Delta = (\Delta_e, \Delta_{\hat{z}}), \Delta_e = (r_1, r_2, \dots, r_n) = \Delta_{\hat{z}} \quad (7)$$

Then, the Lyapunov function of this system satisfies

$$\dot{V}_n|_{(3)} \leq -b \|Z\|_{\Delta}^{\mu+\tau} \quad (8)$$

where $b > 0$, $\mu = 2r_n$.

3 Obverser and Controller Design

In this part, we will construct observer and controller for nonlinear system (1). At the beginning, we adopt the following coordinate transformation for system (1)

$$z_i = \frac{x_i}{L^{i-1}}, \hat{z}_i = \frac{\hat{x}_i}{L^{i-1}}, \nu = \frac{u}{L^n}, i = 1, \dots, n \quad (9)$$

where L is a dynamic gain and will be assigned later.

Assumption 1 . *The derivative of output function is bounded by two positive constants:*

$$\underline{\theta} \leq \frac{\partial h(s)}{\partial s} \leq \bar{\theta}, \forall s \in R. \quad (10)$$

Assumption 2 . *There exists an unknown constant $\theta > 0$ and $0 \leq \alpha_i < 1$, such that under the change of coordinates (9)*

$$\left| \frac{\phi_i(\cdot)}{L^{i-1}} \right| \leq \theta L^{\alpha_i} \left(\sum_{j=1}^n |z_j|^{\frac{r_i+\tau}{r_j}} + |\nu|^{\frac{r_i+\tau}{r_n+\tau}} \right), i = 1, \dots, n \quad (11)$$

Theorem 1. *Under Assumption 1 and Assumption 2, there exists a homogeneous output feedback controller*

$$\begin{aligned} \dot{\hat{z}}_i &= L(\hat{z}_{i+1} - a_i \hat{z}_1^{r_i+1}) - (i-1) \frac{\dot{L}}{L} \hat{z}_i, \quad i = 1, \dots, n-1 \\ \dot{\hat{z}}_n &= L(\nu - a_n \hat{z}_1^{r_n+1}) - (n-1) \frac{\dot{L}}{L} \hat{z}_n \\ \nu &= -\beta_n (\hat{z}_n^{\frac{r_n+1}{r_n}} + \beta_{n-1}^{\frac{r_n+1}{r_n}} (\hat{z}_{n-1}^{\frac{r_n+1}{r_{n-1}}} + \dots + \beta_2^{\frac{r_n+1}{r_3}} (\hat{z}_2^{\frac{r_n+1}{r_2}} + \beta_1^{\frac{r_n+1}{r_2}} y^{r_n+1}) \dots)) \\ \dot{L} &= L \min\{y^\tau, y^{\mu+\tau}\}, \quad L(0) = 1 \end{aligned} \quad (12)$$

to guarantee system (1) globally asymptotically stable.

Proof. Using (9), system (1) can be changed as:

$$\begin{aligned}\dot{z}_i &= Lz_{i+1} + \frac{\phi_i(\cdot)}{L^{i-1}} - (i-1)\frac{\dot{L}}{L}z_i, \quad i = 1, \dots, n-1 \\ \dot{z}_n &= L\nu + \frac{\phi_n(\cdot)}{L^{n-1}} - (n-1)\frac{\dot{L}}{L}z_n\end{aligned}\tag{13}$$

Define the estimation errors as $e_i = z_i - \hat{z}_i$, $i = 1, \dots, n$. Then, one has

$$\begin{aligned}\dot{e}_i &= L(e_{i+1} - a_i e_1^{r_{i+1}}) + \frac{\phi_i(\cdot)}{L^{i-1}} - (i-1)\frac{\dot{L}}{L}e_i \\ &\quad + La_i(\hat{z}_1^{r_{i+1}} + (z_1 - \hat{z}_1)^{r_{i+1}}), \quad i = 1, \dots, n-1 \\ \dot{e}_n &= -La_n e_1^{r_{n+1}} + \frac{\phi_n(\cdot)}{L^{n-1}} - (n-1)\frac{\dot{L}}{L}e_n \\ &\quad + La_n(\hat{z}_1^{r_{n+1}} + (z_1 - \hat{z}_1)^{r_{n+1}})\end{aligned}\tag{14}$$

Using the notations in (6), using the same Lyapunov function V_n adopted in Lemma 3, it's easy to obtain

$$\begin{aligned}\dot{V}_n &\leq -Lb\|Z\|_{\Delta}^{\mu+\tau} + \frac{\partial V_n}{\partial Z}[\phi_1(\cdot), \dots, \frac{\phi_n(\cdot)}{L^{n-1}}, \frac{\partial h}{\partial z_1}\phi_1(\cdot), 0, \dots, 0]^T \\ &\quad + \frac{\partial V_n}{\partial Z}[0, -\frac{\dot{L}}{L}e_2, \dots, -(n-1)\frac{\dot{L}}{L}e_n, 0, \dots, 0]^T \\ &\quad + \frac{\partial V_n}{\partial Z}[0, \dots, 0, 0, -\frac{\dot{L}}{L}\hat{z}_2, \dots, -(n-1)\frac{\dot{L}}{L}\hat{z}_n]^T\end{aligned}\tag{15}$$

Next, we estimate some items in (15).

By the definition of homogeneous systems in [19] and Lemma 1, one has

$$\left|\frac{\partial V_n}{\partial Z_i}\right| \leq d_i\|Z\|_{\Delta}^{\mu-r_i}, \quad \left|\frac{\partial V_n}{\partial Z_{n+i}}\right| \leq d_i\|Z\|_{\Delta}^{\mu-r_i}, \quad i = 1, \dots, n\tag{16}$$

where $d_i > 0$ is a constant.

From Assumption 2 and (16), it yields

$$\sum_{i=1}^n \frac{\partial V_n}{\partial Z_i} \frac{\phi_i(\cdot)}{L^{i-1}} \leq \theta \sum_{i=1}^n d_i L^{\alpha_i} \|Z\|_{\Delta}^{\mu+\tau}\tag{17}$$

By Assumption 1, Assumption 2 and (16), it can be concluded that

$$\begin{aligned}\left|\frac{\partial V_n}{\partial Z_{n+1}}\right| \cdot \left|\frac{\partial h}{\partial z_1}\phi_1(\cdot)\right| &\leq d_1\|Z\|_{\Delta}^{\mu-r_1} \bar{\theta} L^{\alpha_1} \|Z\|_{\Delta}^{r_1+\tau} \\ &\leq \theta d_1 \bar{\theta} L^{\alpha_1} \|Z\|_{\Delta}^{\mu+\tau}\end{aligned}\tag{18}$$

From the definition of a homogeneous p -norm, we have

$$|e_i| \leq (|e_i|^{\frac{2}{r_i}})^{\frac{r_i}{2}} \leq \|Z\|_{\Delta}^{r_i}, \quad i = 1, \dots, n\tag{19}$$

Using the same approach used in (19), one has

$$|y|^\tau \leq |y^{\frac{2}{r_1}}|^{\frac{r_1}{2} \cdot \tau} \leq \|Z\|_\Delta^\tau, |\hat{z}_i| \leq \left[\sum_{j=2}^n |\hat{z}_j|^{\frac{2}{r_j}} \right]^{\frac{r_i}{2}} \leq \|Z\|_\Delta^{r_i}, i = 2, \dots, n \quad (20)$$

From (12), (16), (19) and (20), one has

$$\sum_{i=1}^n \frac{\partial V_n}{\partial Z_i} \left(-\frac{\dot{L}}{L} \right) (i-1) e_i \leq \sum_{i=1}^n d_i (i-1) \|Z\|_\Delta^\mu \|Z\|_\Delta^\tau \leq \bar{d}_1 \|Z\|_\Delta^{\mu+\tau} \quad (21)$$

where $\bar{d}_1 > 0$ is a constant.

From (12), (16) and (20), one has

$$\sum_{i=2}^n \frac{\partial V_n}{\partial Z_{n+i}} \left(-\frac{\dot{L}}{L} \right) (i-1) \hat{z}_i \leq \sum_{i=2}^n d_i (i-1) \|Z\|_\Delta^{\mu-r_i} |y|^\tau \|Z\|_\Delta^{r_i} \leq \bar{d}_2 \|Z\|_\Delta^{\mu+\tau} \quad (22)$$

where $\bar{d}_2 > 0$ is a constant.

Substituting (17), (18), (21) and (22) into (15), one has

$$\begin{aligned} \dot{V}_n|_{(15)} &\leq -bL\|Z\|_\Delta^{\mu+\tau} + \bar{d}\|Z\|_\Delta^{\mu+\tau} + \theta \sum_{i=1}^n d_i L^{\alpha_i} \|Z\|_\Delta^{\mu+\tau} + \theta d_1 \bar{\theta} L^{\alpha_1} \|Z\|_\Delta^{\mu+\tau} \\ &\leq -bL\|Z\|_\Delta^{\mu+\tau} + \bar{d}\|Z\|_\Delta^{\mu+\tau} + \theta(1 + \bar{\theta}) L^{\alpha_{max}} \sum_{i=1}^n d_i \|Z\|_\Delta^{\mu+\tau} \\ &\leq -L(b - \frac{\bar{d}}{L} - \theta(1 + \bar{\theta}) L^{\alpha_{max}-1} \sum_{i=1}^n d_i) \|Z\|_\Delta^{\mu+\tau} \end{aligned} \quad (23)$$

where $\bar{d} = \bar{d}_1 + \bar{d}_2$ is a constant and $\alpha_{max} = \max\{\alpha_1, \dots, \alpha_n\}$.

4 Stability Analysis

In this part, we will show that $L(t)$, $e(t)$ and $\hat{z}(t)$ are bounded on $[0, t_f]$. At first, we prove $L(t)$ cannot escape at $t = t_f$. To show this, we assume that $\lim_{t \rightarrow t_f} L(t) = +\infty$. From (12), one has $\dot{L} \geq 0$, which indicates $L(t)$ is a monotonic nondecreasing function. As a result, there exists a finite time $t_0 \in (0, t_f)$ and a constant $c > 0$, such that

$$b - \frac{\bar{d}}{L} - \theta(1 + \bar{\theta}) L^{\alpha_{max}-1} \sum_{i=1}^n d_i > c, \quad \forall t \in [t_0, t_f] \quad (24)$$

under which (23) becomes

$$\dot{V}_n \leq -cL\|Z\|_\Delta^{\mu+\tau}, \quad \forall t \in [t_0, t_f] \quad (25)$$

It can be deduced from (12) that

$$\begin{aligned}
+\infty &= L(t_f) - L(t_0) = \int_{t_0}^{t_f} L \min\{y^\tau, y^{\mu+\tau}\} ds \leq \int_{t_0}^{t_f} L |y|^{\mu+\tau} ds \\
&\leq \int_{t_0}^{t_f} L \|Z\|_\Delta^{\mu+\tau} ds \leq - \int_{t_0}^{t_f} \frac{\dot{V}_n}{c} ds \leq - \frac{1}{c} [V_n(t_f) - V_n(t_0)] \\
&\leq \frac{1}{c} V_n(t_0) = \text{constant}
\end{aligned} \tag{26}$$

which indicates that $L(t)$ and $\int_0^{t_f} L \|Z\|_\Delta^{\mu+\tau} ds$ are bounded on $[0, t_f]$. Then, we can conclude $\int_0^{t_f} \|Z\|_\Delta^{\mu+\tau} ds$ is also bounded.

In (6), Z has been defined as $Z := [e_1, \dots, e_n, y, \hat{z}_2, \dots, \hat{z}_n]^T$. Thus, it can be proved that $e(t)$ and $\hat{z}(t)$ is bounded on $[0, t_f]$. What proved above contradicts with $\lim_{t \rightarrow t_f} \sup \|Z(t), L(t)\| = \infty$. From (12) and (14), it can be deduced that $\dot{e}(t)$ and $\dot{\hat{z}}(t)$ are bounded on $[0, \infty)$. By Barbalat Lemma, we finally prove $\lim_{t \rightarrow +\infty} (e(t), \hat{z}(t)) = 0$.

5 An Example

Consider

$$\begin{aligned}
\dot{x}_1 &= x_2 + \theta x_1^{\frac{6}{7}} x_3^{\frac{3}{11}}, \quad \dot{x}_2 = x_3, \\
\dot{x}_3 &= u, \quad y = d_1 x_1 + d_2 \sin(x_1)
\end{aligned} \tag{27}$$

where θ , d_1 and d_2 are unknown constants. By Lemma 2,

$$|x_1^{\frac{6}{7}} x_3^{\frac{3}{11}}| = |z_1^{\frac{6}{7}} (L^2 z_3)^{\frac{3}{11}}| \leq L^{\frac{6}{11}} (|z_1|^{\frac{9}{7}} + |z_3|^{\frac{9}{11}}) \tag{28}$$

which implies that Assumption 2 holds with $\alpha_1 = \frac{6}{11}$ and $\tau = \frac{2}{7}$. For system (27), we design the following output controller

$$\begin{aligned}
\dot{\hat{z}}_1 &= L(\hat{z}_2 - a_1 \hat{z}_1^{r_2}) \\
\dot{\hat{z}}_2 &= L(\hat{z}_3 - a_2 \hat{z}_1^{r_3}) - \frac{\dot{L}}{L} \hat{z}_2 \\
\dot{\hat{z}}_3 &= L(\nu - a_3 \hat{z}_1^{r_4}) - 2 \frac{\dot{L}}{L} \hat{z}_3 \\
u &= L^3 \nu = -L^3 \beta_3 (\hat{z}_3^{\frac{r_4}{r_3}} + \beta_2^{\frac{r_4}{r_3}} (\hat{z}_2^{\frac{r_4}{r_2}} + \beta_1^{\frac{r_4}{r_2}} y^{r_4}))
\end{aligned} \tag{29}$$

In the computer simulation, the parameters are selected as follows: $\theta = 1.4$, $d_1 = 1.2$, $d_2 = 1.5$, $\beta_1 = 1$, $\beta_2 = 5$, $\beta_3 = 8$, $a_1 = 6$, $a_2 = 11$, $a_3 = 6$. The simulation results are shown in Fig. 1 and Fig. 2 with the initial value $[x_1(0), x_2(0), x_3(0), \hat{z}_1(0), \hat{z}_2(0), \hat{z}_3(0)] = [0.5, -1, 1, 0, 0, 0]$. Obviously, the states and their estimations are globally asymptotically stable. The dynamic gain L is also bounded.

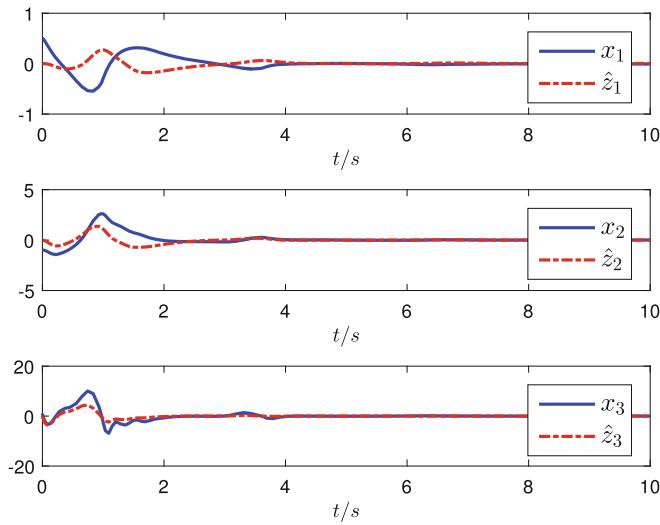


Fig. 1. The curves of $x_1, \hat{z}_1, x_2, \hat{z}_2, x_3, \hat{z}_3$

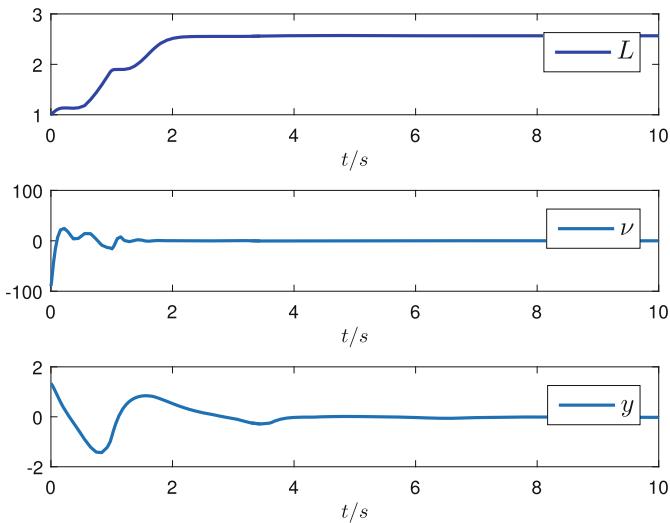


Fig. 2. The curves of L, ν, y

6 Conclusion

In conclusion, the global stabilization of output feedback for a class of nonlinear systems with unknown growth rate and unknown output function is studied above. This paper introduces a dynamic gain into coordinate transformation. Then, the adjustable scaling gain is adopted in the design of the output feedback controller. By utilizing the homogeneous domination, it suffices to show that the

controller can render the system states globally asymptotically converge to the origin.

Acknowledgments. This work is supported in part by National Natural Science Foundation of China [grant number 61873061].

References

1. Moog, C.H., Xia, X.: Output feedback control design for nonlinear systems. IFAC Proc. Vol. **30**(21), 69–73 (1997)
2. Man, Y.C., Liu, Y.G.: Global output-feedback stabilization for high-order nonlinear systems with unknown control coefficients. IFAC-PapersOnLine **48**(28), 1214–1219 (2015)
3. Li, E., Long, L., Zhao, J.: Global output-feedback stabilization for a class of switched uncertain nonlinear systems. Appl. Math. Comput. **256**, 551–564 (2015)
4. Tsinias, J.: A theorem on global stabilization of nonlinear systems by linear feedback. Syst. Control Lett. **17**(5), 357–362 (1991)
5. Gildas, B.: State-affine systems and observer-based control. IFAC Proc. Vol. **31**(17), 391–396 (1998)
6. Yan, X.H., Liu, Y.G., Zheng, W.X.: Global adaptive output-feedback stabilization for a class of uncertain nonlinear systems with unknown growth rate and unknown output function. Automatica **104**, 173–181 (2019)
7. Qian, C.J., Lin, W.: Output feedback control of a class of nonlinear systems: a nonseparation principle paradigm. IEEE Trans. Autom. Control **47**(10), 1710–1715 (2002)
8. Praly, L., Jiang, Z.P.: Linear output feedback with dynamic high gain for nonlinear systems. Syst. Control Lett. **53**(2), 107–116 (2004)
9. Guo, T.T., Zhang, K.M., Xie, X.J.: Output feedback stabilization of high-order nonlinear systems with polynomial nonlinearity. J. Frank. Inst. **355**(14), 6579–6596 (2018)
10. Qian, C.J.: A homogeneous domination approach for global output feedback stabilization of a class of nonlinear systems. In: Proceedings of the 2005 American Control Conference, pp. 4708–4715 . IEEE (2005)
11. Yang, S.H., Sun, Z.Y., Wang, Z., Li, T.: A new approach to global stabilization of high-order time-delay uncertain nonlinear systems via time-varying feedback and homogeneous domination. J. Frank. Inst. **355**(14), 6469–6492 (2018)
12. Lei, H., Lin, W.: Universal output feedback control of nonlinear systems with unknown growth rate. IFAC Proc. Vol. **38**(1), 1073–1078 (2005)
13. Li, F.Z., Liu, Y.G.: Global stabilization via time-varying output-feedback for stochastic nonlinear systems with unknown growth rate. Syst. Control Lett. **77**, 69–79 (2015)
14. Zhao, Z.L., Guo, B.Z.: On active disturbance rejection control for nonlinear systems using time-varying gain. Eur. J. Control **23**, 62–70 (2015)
15. Choi, H.L., Lim, J.T.: Stabilization of a class of nonlinear systems by adaptive output feedback. Automatica **41**(6), 1091–1097 (2005)
16. Wang, P., Yu, C.P.: Output feedback control for nonlinear systems with uncertainties on output functions and growth rates. Eur. J. Control (2020)
17. Zhai, J.Y.: Global finite-time output feedback stabilisation for a class of uncertain nontriangular nonlinear systems. Int. J. Syst. Sci. **45**(3), 637–646 (2014)

18. Yan, X.H., Song, X.M., Wang, Z.H., Zhang, Y.: Global output-feedback adaptive stabilization for planar nonlinear systems with unknown growth rate and output function. *Appl. Math. Comput.* **314**, 299–309 (2017)
19. Bacciotti, A., Rosier, L.: Liapunov Functions and Stability in Control Theory, 2nd edn. Springer, Berlin (2005)
20. Zhai, J.Y.: Dynamic output-feedback control for nonlinear time-delay systems and applications to chemical reactor systems. *IEEE Trans. Circ. Syst. II Express Briefs* **66**(11), 1845–1849 (2019)
21. Zhai, J.Y., Qian, C.J.: Global control of nonlinear systems with uncertain output function using homogeneous domination approach. *Int. J. Robust Nonlinear Control* **22**(14), 1543–1561 (2012)



Adaptive Control for Planar Nonlinear Systems with Input Quantization and Unknown Control Directions

Ce Liu and Junyong Zhai^(✉)

School of Automation, Southeast University, Nanjing 210096, Jiangsu, China
jyzhai@seu.edu.cn

Abstract. This paper studies the problem of adaptive control for planar nonlinear systems with input quantization. The Nussbaum-type function is applied to handle the unknown control directions. By virtue of dynamic gain technology, an adaptive controller is designed to ensure that all signals of the system are globally bounded.

Keywords: Nonlinear systems · Unknown control directions · Quantization

1 Introduction

Non-linearity is the most common phenomenon in nature and engineering technology. Recently, research on nonlinear systems has made gratifying progress, such as [1, 2, 4, 6, 13, 14, 18, 19] and the references therein. The work [19] investigated the state feedback control problem for uncertain nonlinear systems with input quantization by a hysteretic logarithmic quantizer. Then, a hysteresis uniform quantizer was applied to decrease chattering with bounded quantization error in [18]. The work [6] considered state feedback stabilization problem for high-order nonlinear systems with input quantization.

On the other hand, the problem of nonlinear systems with unknown control directions has also aroused great attention [5, 9, 15] and the references therein. By virtue of Nussbaum-type function and adaptive control technique, an adaptive controller was designed to stabilize planar nonlinear systems with unknown control directions. The work [8] considered adaptive control problem for quantized nonlinearly parameterized systems with unknown control directions. Then, an adaptive state feedback controller was proposed for a class of feedforward nonlinear systems in [12].

In this note, we consider the following system described by

$$\dot{\eta}_1 = g_1 \eta_2 + \phi_1(\eta_1), \quad \dot{\eta}_2 = g_2 q(u) + \phi_2(\eta_1, \eta_2), \quad y = \eta_1 \quad (1)$$

where $\eta = (\eta_1, \eta_2)^T$ is the system state, u is the control input, y is measured output. $\phi_1(\cdot), \phi_2(\cdot)$ are uncertain nonlinear functions. g_1, g_2 are unknown control

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 266–274, 2021.

https://doi.org/10.1007/978-981-15-8458-9_29

coefficients. $q(\cdot)$ is the logarithmic quantizer described as follows:

$$q(u) = \begin{cases} u_i, & \frac{u_i}{1+\delta} \leq u < \frac{u_i}{1-\delta} \\ 0, & 0 \leq u < \frac{u_{\min}}{1+\delta} \\ -q(-u), & u < 0 \end{cases} \quad (2)$$

with $u_i = \rho^{1-i} u_{\min}$ ($i = 1, 2, \dots$), $u_{\min} > 0$, $0 < \rho < 1$ and $\delta = \frac{1-\rho}{1+\rho}$. ρ denotes the logarithmic quantizer's density and u_{\min} determines the dead zone of quantizer. Most of works on quantization only consider the problem of state feedback controller design. In practice applications, it is difficult to obtain all states information and output feedback control is a better way [13, 14]. The work [17] discussed the stabilization problem for nonlinear systems with unknown output function by sampled-data control.

2 Preliminaries

We introduce a definition of Nussbaum function and some useful lemmas.

Definition 1 [9]. A continuous function $N : [0, \infty) \rightarrow \mathbb{R}$, is called Nussbaum function:

$$\lim_{s \rightarrow \infty} \sup\left(\frac{1}{s} \int_0^s N(\xi) d\xi\right) = +\infty, \quad \lim_{s \rightarrow \infty} \inf\left(\frac{1}{s} \int_0^s N(\xi) d\xi\right) = -\infty.$$

Remark 1. From Definition 1, the Nussbaum-type functions can be chosen as $\xi^2 \cos(\xi)$, $e^{\xi^2} \cos(\xi)$, $\xi^2 \sin(\xi)$ and $e^{\xi^2} \cos(\frac{\pi}{2}\xi)$, which have been proved in [10, 16].

Lemma 1 [3]. Suppose that $V(t)$ and $\xi(t)$ are smooth functions with $V(t) \geq 0, \forall t \in [0, t_f]$, and $N(\xi)$ is a Nussbaum-type function. If

$$V(t) \leq \alpha_0 + e^{-\alpha_1 t} \int_0^t (g(v)N(\xi(v)) + 1)\dot{\xi}(v)e^{\alpha_1 v} dv, \quad \forall t \in [0, t_f)$$

where α_0 is a suitable constant, $\alpha_1 > 0$ is a constant, and $g(\cdot)$ is a time-varying parameter taken values in the closed intervals $I := [l^-, l^+]$ with $0 \notin I$, then $\xi(t), V(t)$ and $\int_0^t g(v)N(\xi(v))\dot{\xi}(v)dv$ are bounded on $[0, t_f]$.

Lemma 2 [11]. For any positive real numbers ρ, ϱ and γ , one has

$$|x|^{\rho}|y|^{\varrho} \leq \frac{\rho}{\rho + \varrho} \gamma |x|^{\rho+\varrho} + \frac{\varrho}{\rho + \varrho} \gamma^{-\frac{\rho}{\varrho}} |y|^{\rho+\varrho}.$$

3 Main Result

To solve this problem, we first make the following assumptions.

Assumption 1. There exists a constant $c \geq 0$, such that

$$\phi_1(\eta_1) \leq c|\eta_1|, \quad \phi_2(\eta_1, \eta_2) \leq c(|\eta_1| + |\eta_2|).$$

Assumption 2. Each control coefficient $g_i (i = 1, 2)$ equals +1 or -1, but unknown.

Before moving on, we introduce the following change of coordinates:

$$x_i = \frac{1}{\prod_{j=i}^2 g_j} \eta_j, \quad i = 1, 2 \quad (3)$$

under which system (1) becomes

$$\dot{x}_1 = x_2 + \frac{1}{g} \phi_1(\cdot), \quad \dot{x}_2 = q(u) + \frac{1}{g_2} \phi_2(\cdot), \quad y = g x_1 \quad (4)$$

where $g = g_1 g_2$. Under Assumption 1, one has

$$|\frac{1}{g} \phi_1(\cdot)| \leq c|x_1|, \quad |\frac{1}{g_2} \phi_2(\cdot)| \leq c(|x_1| + |x_2|). \quad (5)$$

Motivated by [17], we design the state observer for system (4) as:

$$\dot{\hat{x}}_1 = \hat{x}_2 - L a_1 \hat{x}_1, \quad \dot{\hat{x}}_2 = q(u) - L^2 a_2 \hat{x}_1 \quad (6)$$

where $a_1 > 0$ and $a_2 > 0$. $L > 1$ is a dynamic gain.

The quantizer is decomposed into two parts as [7]:

$$q(u) = q_1 u + q_2$$

where

$$q_1 = \begin{cases} \frac{q(u)}{u}, & |u(t)| \geq \frac{u_{\min}}{1+\delta} \\ 1, & |u(t)| < \frac{u_{\min}}{1+\delta} \end{cases}, \quad q_2 = \begin{cases} 0, & |u(t)| \geq \frac{u_{\min}}{1+\delta} \\ q(u) - u, & |u(t)| < \frac{u_{\min}}{1+\delta} \end{cases}$$

with $1 - \delta \leq q_1 \leq 1 + \delta$, $|q_2| \leq \frac{u_{\min}}{1+\delta}$ and the quantized parameter $\delta \in (0, 1)$.

Next, one introduces the following transformation:

$$z_i = \frac{\hat{x}_i}{L^i}, \quad i = 1, 2, \quad \epsilon_1 = \frac{y}{L}, \quad \epsilon_2 = z_2 - \alpha(\cdot) \quad (7)$$

where $\alpha(\cdot)$ is a smooth function.

Define $\varepsilon_i = \frac{x_i - \hat{x}_i}{L^i}, i = 1, 2$. Then, one has

$$\dot{\varepsilon} = L A \varepsilon + \Phi + a x_1 - \frac{\dot{L}}{L} D \varepsilon \quad (8)$$

where $\varepsilon = (\varepsilon_1, \varepsilon_2)^T$, $A = \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix}$, $\Phi = (\frac{\phi_1}{gL}, \frac{\phi_2}{g_2 L^2})^T$, $a = (a_1, a_2)^T$, $D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.

Since A is a Hurwitz matrix, there exists a positive definite matrix P , such that

$$A^T P + P A \leq -I, \quad P D + D^T P \geq 0. \quad (9)$$

Design the adaptive law of L as

$$\dot{L} = \max\left\{\frac{\beta(\cdot)}{L} - L^2, 0\right\}, \quad L(0) = 6c^2\|P\|^2 + 4c\|P\| + 1 \quad (10)$$

where $\beta(\cdot) > 0$ is a continuous function.

Choosing $V_0 = \varepsilon^T P \varepsilon$, one has

$$\begin{aligned} \dot{V}_0 &= L\varepsilon^T (A^T P + PA)\varepsilon + 2\varepsilon^T P(\Phi + ax_1) - \frac{\dot{L}}{L}\varepsilon^T (PD + D^T P)\varepsilon \\ &\leq -L\|\varepsilon\|^2 + 2\varepsilon^T P\Phi + 2\varepsilon^T Pax_1 \end{aligned} \quad (11)$$

Due to $L(t) > 1$ for $\forall t \geq 0$, by Assumption 1 and (7), one has

$$\begin{aligned} \|\Phi\|_2 &\leq \|\Phi\|_1 \leq \frac{2c|y|}{L} + \frac{c|L^2\varepsilon_2 + L^2z_2|}{L^2} \\ &\leq 2c|\epsilon_1| + c\|\varepsilon\| + c|z_2| \end{aligned}$$

under which one has

$$2\varepsilon^T P\Phi \leq (3c^2\|P\|^2 + 2c\|P\|)\|\varepsilon\|^2 + z_2^2 + 2\epsilon_1^2. \quad (12)$$

By Lemma 2, one has

$$|2\varepsilon^T Pax_1| = |2\varepsilon^T PaL\epsilon_1| \leq \frac{1}{4}L\|\varepsilon\|^2 + 4L\|Pa\|^2\epsilon_1^2. \quad (13)$$

Substituting (12) and (13) into (11), we have

$$\dot{V}_0 \leq -\left(\frac{3}{4}L - \Theta_1\right)\|\varepsilon\|^2 + (4L\|Pa\|^2 + 2)\epsilon_1^2 + z_2^2$$

where $\Theta_1 = 3c^2\|P\|^2 + 2c\|P\|$.

Step 1 : Choose the Lyapunov function $V_1 = V_0 + \frac{1}{2}\epsilon_1^2$. Then, one has

$$\begin{aligned} \dot{V}_1 &= \dot{V}_0 + \left(gL\varepsilon_2 + gLz_2 + \frac{1}{L}\phi_1 - \frac{\dot{L}}{L}\epsilon_1\right)\epsilon_1 \\ &\leq -\left(\frac{3}{4}L - \Theta_1\right)\|\varepsilon\|^2 + (4L\|Pa\|^2 + 2)\epsilon_1^2 + z_2^2 \\ &\quad + gL\epsilon_1\varepsilon_2 + gL\epsilon_1z_2 + \frac{1}{L}\epsilon_1\phi_1 - \frac{\dot{L}}{L}\epsilon_1^2. \end{aligned} \quad (14)$$

Now, we estimate some cross terms in (14). By Lemma 2, one has

$$\begin{aligned} gL\epsilon_1\varepsilon_2 &\leq \frac{L}{8}\|\varepsilon\|^2 + 2L\epsilon_1^2, \\ gL\epsilon_1z_2 &= gL\epsilon_1(\epsilon_2 + \alpha(\cdot)), \\ \frac{1}{L}\epsilon_1\phi_1 &\leq c\epsilon_1^2. \end{aligned} \quad (15)$$

From (10), one has $\dot{L} \geq \frac{\beta(\cdot)}{L} - L^2$. Then

$$-\frac{\dot{L}}{L}\epsilon_1^2 \leq -\frac{1}{L}\left(\frac{\beta(\cdot)}{L} - L^2\right)\epsilon_1^2 \leq -\frac{\beta(\cdot)}{L^2}\epsilon_1^2 + L\epsilon_1^2. \quad (16)$$

Taking (15) and (16) into (14), it yields

$$\begin{aligned} \dot{V}_1 &\leq -\left(\frac{5}{8}L - \Theta_1\right)\|\varepsilon\|^2 + (4L\|Pa\|^2 + 3L + c + 2)\epsilon_1^2 \\ &\quad + z_2^2 + gL\epsilon_1\epsilon_2 + gL\epsilon_1\alpha(\cdot) - \frac{\beta(\cdot)}{L^2}\epsilon_1^2 \\ &\leq -\left(\frac{5}{8}L - \Theta_1\right)\|\varepsilon\|^2 - 2(L + c)\epsilon_1^2 + (4L\|Pa\|^2 + 5L + 3c + 2)\epsilon_1^2 \\ &\quad + 2(\epsilon_2^2 + \alpha^2(\cdot)) + gL\epsilon_1\epsilon_2 + gL\epsilon_1\alpha(\cdot) - \frac{\beta(\cdot)}{L^2}\epsilon_1^2. \end{aligned} \quad (17)$$

Design the following virtual controller

$$\begin{aligned} \alpha(\xi, L, \epsilon_1) &= \frac{1}{L}N(\xi)(4L\|Pa\|^2 + 5L + 3c + 2)\epsilon_1, \\ \dot{\xi} &= (4L\|Pa\|^2 + 5L + 3c + 2)\epsilon_1^2 \end{aligned} \quad (18)$$

where $N(\cdot) = e^\xi \cos(\pi\xi/2)$. Under which (17) becomes

$$\begin{aligned} \dot{V}_1 &\leq -\left(\frac{5}{8}L - \Theta_1\right)\|\varepsilon\|^2 - 2(L + c)\epsilon_1^2 + (gN(\xi) + 1)\dot{\xi} \\ &\quad + 2\epsilon_2^2 + 2\alpha^2(\cdot) + gL\epsilon_1\epsilon_2 - \frac{\beta(\cdot)}{L^2}\epsilon_1^2. \end{aligned} \quad (19)$$

Step 2 : Construct the Lyapunov function $V_2 = V_1 + \frac{1}{2}\epsilon_2^2$. From the definition of ϵ_2 , one has

$$\begin{aligned} \dot{\epsilon}_2 &= \frac{1}{L^2}q(u) - a_2\hat{x}_1 - 2\frac{\dot{L}}{L}z_2 - \frac{\partial\alpha}{\partial L}\dot{L} - \frac{\partial\alpha}{\partial\xi}\dot{\xi} \\ &\quad - \frac{\partial\alpha}{\partial\epsilon_1}(gL\varepsilon_2 + gLz_2 + \frac{1}{L}\phi_1 - \frac{\dot{L}}{L}\epsilon_1). \end{aligned}$$

Then, one has

$$\begin{aligned} \dot{V}_2 &\leq -\left(\frac{5}{8}L - \Theta_1\right)\|\varepsilon\|^2 - 2(L + c)\epsilon_1^2 + (gN(\xi) + 1)\dot{\xi} + 2\epsilon_2^2 + 2\alpha^2(\cdot) \\ &\quad + gL\epsilon_1\epsilon_2 - \frac{\beta(\cdot)}{L^2}\epsilon_1^2 + \frac{1}{L^2}\epsilon_2q(u) - a_2\epsilon_2\hat{x}_1 - 2\frac{\dot{L}}{L}\epsilon_2z_2 - \frac{\partial\alpha}{\partial L}\epsilon_2\dot{L} - \frac{\partial\alpha}{\partial\xi}\epsilon_2\dot{\xi} \\ &\quad - gL\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_2\varepsilon_2 - gL\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_2z_2 - \frac{1}{L}\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_2\phi_1 + \frac{\dot{L}}{L}\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_2\epsilon_1. \end{aligned} \quad (20)$$

Similar to (15), one has

$$\begin{aligned} gL\epsilon_1\epsilon_2 &\leq L\epsilon_1^2 + \frac{L}{4}\epsilon_2^2, \\ -gL\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_2\varepsilon_2 &\leq \frac{L}{8}\|\varepsilon\|^2 + 2L(\frac{\partial\alpha}{\partial\epsilon_1})^2\epsilon_2^2, \\ -gL\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_2z_2 &\leq \frac{1}{2}z_2^2 + \frac{1}{2}L^2(\frac{\partial\alpha}{\partial\epsilon_1})^2\epsilon_2^2 \leq \epsilon_2^2 + a^2(\cdot) + \frac{1}{2}L^2(\frac{\partial\alpha}{\partial\epsilon_1})^2\epsilon_2^2, \\ -\frac{1}{L}\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_2\phi_1 &\leq c|\frac{\partial\alpha}{\partial\epsilon_1}||\epsilon_2||\epsilon_1| \leq c\epsilon_1^2 + \frac{c}{4}(\frac{\partial\alpha}{\partial\epsilon_1})^2\epsilon_2^2. \end{aligned}$$

Substituting the above inequalities into (20), one has

$$\begin{aligned} \dot{V}_2 &\leq -\left(\frac{1}{2}L - \Theta_1\right)\|\varepsilon\|^2 - (L + c)\epsilon_1^2 + (gN(\xi) + 1)\dot{\xi} \\ &\quad + 3\epsilon_2^2 + 3\frac{1}{L^2}N^2(\xi)(4L\|Pa\|^2 + 5L + 3c + 2)^2\epsilon_1^2 \\ &\quad + \frac{\epsilon_2}{L^2}(q_1u + q_2) + \epsilon_2\left(-a_2\hat{x}_1 - 2\frac{\dot{L}}{L}z_2 - \frac{\partial\alpha}{\partial L}\dot{L} - \frac{\partial\alpha}{\partial\xi}\dot{\xi}\right. \\ &\quad \left.+ (\frac{1}{2}L^2 + 2L + \frac{c}{4})(\frac{\partial\alpha}{\partial\epsilon_1})^2\epsilon_2 + \frac{\dot{L}}{L}\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_1 + \frac{L}{4}\epsilon_2\right) - \frac{\beta(\cdot)}{L^2}\epsilon_1^2. \end{aligned} \quad (21)$$

Using Lemma 2, one has

$$\frac{\epsilon_2}{L^2}q_2 \leq \frac{1}{L^2}\epsilon_2^2 + \frac{1}{4L^2}q_2^2 \leq \frac{1}{L^2}\epsilon_2^2 + \frac{d^2}{4L^2(0)}. \quad (22)$$

Choose the continuous function $\beta(\cdot)$ as

$$\beta(\xi, L) = 3N^2(\xi)(4L\|Pa\|^2 + 5L + 3c + 2)^2. \quad (23)$$

Note that $L \geq L(0)$ and substitute (22) and (23) into (21), one has

$$\begin{aligned} \dot{V}_2 &\leq -\frac{1}{2}\|\varepsilon\|^2 - (L + c)\epsilon_1^2 + (gN(\xi) + 1)\dot{\xi} + \frac{\epsilon_2}{L^2}q_1u \\ &\quad + \epsilon_2\left(-a_2\hat{x}_1 - 2\frac{\dot{L}}{L}z_2 - \frac{\partial\alpha}{\partial L}\dot{L} - \frac{\partial\alpha}{\partial\xi}\dot{\xi} + (\frac{1}{2}L^2 + 2L + \frac{c}{4})(\frac{\partial\alpha}{\partial\epsilon_1})^2\epsilon_2\right. \\ &\quad \left.+ \frac{\dot{L}}{L}\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_1 + (\frac{1}{L^2} + \frac{L}{4} + 3)\epsilon_2\right) + \frac{d^2}{4L^2(0)}. \end{aligned}$$

Design the controller u as

$$u = -\frac{L^2}{1 - \text{sign}(f(\cdot)\epsilon_2)\delta}f(\cdot) \quad (24)$$

where $f(\cdot) = -a_2\hat{x}_1 - 2\frac{\dot{L}}{L}z_2 - \frac{\partial\alpha}{\partial L}\dot{L} - \frac{\partial\alpha}{\partial\xi}\dot{\xi} + (\frac{1}{2}L^2 + 2L + \frac{c}{4})(\frac{\partial\alpha}{\partial\epsilon_1})^2\epsilon_2 + \frac{\dot{L}}{L}\frac{\partial\alpha}{\partial\epsilon_1}\epsilon_1 + (\frac{1}{L^2} + \frac{L}{4} + 3)\epsilon_2 + \epsilon_2$. Under which one has

$$\begin{aligned}\dot{V}_2 &\leq -\frac{1}{2}\|\varepsilon\|^2 - c\epsilon_1^2 - \epsilon_2^2 + (gN(\xi) + 1)\dot{\xi} + \frac{d^2}{4L^2(0)} \\ &\leq -\gamma V_2 + (gN(\xi) + 1)\dot{\xi} + \frac{d^2}{4L^2(0)}\end{aligned}\quad (25)$$

where $\gamma > 0$ is a constant.

Theorem 1. *Under Assumptions 1 and 2, there exists an adaptive controller (24) with dynamic gain (10), such that the system states are globally bounded.*

Proof. We will prove that for any initial value (x_0, \hat{x}_0, ξ_0) and $L(0)$, the corresponding solution $(x(t), \hat{x}(t), \xi(t), L(t))$ of the closed-loop system (4)–(6)–(10)–(24) exists on $[0, \infty)$, and is unique and globally bounded.

Since the system is locally Lipschitz in $(x, \hat{x}, \xi) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}$. Then, for any initial value $(x_0, \hat{x}_0, \xi_0) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}$ and $L(0)$, then the system has a unique solution on $[0, t_f]$.

Firstly, we prove the boundedness of $\xi(t)$ on $[0, t_f]$. From (25), one has

$$\begin{aligned}V_2(t) &\leq V_2(0)e^{-\gamma t} + \frac{d^2}{4\gamma L^2(0)}(1 - e^{-\gamma t}) + e^{-\gamma t} \int_0^t (gN(\xi(v)) + 1)\dot{\xi}(v)e^{\gamma v} dv \\ &\leq V_2(0) + \frac{d^2}{4\gamma L^2(0)} + e^{-\gamma t} \int_0^t (gN(\xi(v)) + 1)\dot{\xi}(v)e^{\gamma v} dv\end{aligned}\quad (26)$$

By virtue of Lemma 1, one can obtain that $V_2(t)$, $\xi(t)$ and $\int_0^t (gN(\xi(v)) + 1)\dot{\xi}(v)dv$ are bounded on $[0, t_f]$. Since $V_2 = \varepsilon^T P \varepsilon + \frac{1}{2}\epsilon_1^2 + \frac{1}{2}\epsilon_2^2$, one can obtain $\varepsilon(t)$, $\epsilon_1(t)$ and $\epsilon_2(t)$ are bounded on $[0, t_f]$.

Next, we will prove that $L(t)$ is bounded. Suppose that $\lim_{t \rightarrow t_f} L(t) = +\infty$. Note that $\xi(t)$ is bounded on $[0, t_f]$ and $N(\cdot)$ is continuous with respect to $\xi(t)$, which implies $N(\cdot)$ is bounded. It can be obtained from (10) that there is a time $t^* \in [0, t_f]$ such that $\frac{\beta(\cdot)}{L(t)} - L^2(t) < 0$ and $\dot{L}(t) = 0$ for $\forall t \in [t^*, t_f]$. This contradicts to $\lim_{t \rightarrow t_f} L(t) = +\infty$. Thus, L is bounded on $[0, t_f]$.

Finally, we should prove the boundedness of $x(t), \hat{x}(t)$. From (7) and the definition of ϵ_2 , one has

$$\begin{aligned}\hat{x}_1 &= x_1 - L\varepsilon_1 = L\left(\frac{\epsilon_1}{g} - \varepsilon_1\right), \\ \hat{x}_2 &= L^2 z_2 = L^2(\epsilon_1 + \alpha(\xi, L, \epsilon_1)).\end{aligned}\quad (27)$$

Since $\alpha(\cdot)$ is a continuous function and $\xi(t), L(t), \epsilon_1(t)$ are bounded, one has $\alpha(\cdot)$ and $z_2(t)$ are bounded. Consequently, (27) indicates that $\hat{x}_i(t), i = 1, 2$ are bounded. In the same way, one has $x_1 = \frac{L\epsilon_1}{g}$ and $x_2 = L^2\varepsilon_2 + \hat{x}_2$. From the boundedness of L , ϵ_1 and ε , we can obtain that $x_i(t), i = 1, 2$ are bounded on $[0, t_f]$.

Based on the above analysis, it can be verified that $(x(t), \hat{x}(t), \xi(t), L(t))$ is bounded on $[0, t_f]$, which contradicts with $\lim_{t \rightarrow t_f} \sup \| (x(t), \hat{x}(t), \xi(t), L(t)) \| = +\infty$.

From the boundedness of $\xi, \varepsilon, \epsilon_1$ on $[0, +\infty)$, we can obtain from (18) that V_2 and $(gN(\xi) + 1)\dot{\xi}$ in (26) are bounded on $[0, +\infty)$. Thus, there is a time instant t^* and a small constant M_1 , such that $\dot{V}_2 \leq -M_1 V_2 + \frac{d^2}{4L^2(0)}$. That is, system (1) is globally bounded under the controller (24). \square

4 Conclusion

In this note, an adaptive controller has been designed for planar nonlinear systems with input quantization such that the system states are globally bounded. A Nussbaum-type function and a new dynamic gain are combined to handle unknown control directions. In the future, we will construct a feasible strategy to handle the global stabilization for n -dimensional nonlinear systems.

References

1. Coutinho, D.F., Fu, M., de Souza, C.E.: Input and output quantized feedback linear systems. *IEEE Trans. Autom. Control* **55**(3), 761–766 (2010)
2. Fu, M., Xie, L.: The sector bound approach to quantized feedback control. *IEEE Trans. Autom. Control* **50**(11), 1698–1711 (2005)
3. Ge, S.S., Wang, J.: Robust adaptive tracking for time-varying uncertain nonlinear systems with unknown control coefficients. *IEEE Trans. Autom. Control* **48**(8), 1463–1469 (2003)
4. Hayakawa, T., Ishii, H., Tsumura, K.: Adaptive quantized control for nonlinear uncertain systems. *Syst. Control Lett.* **58**(9), 625–632 (2009)
5. Jiang, Z., Mareels, I., Hill, D., Huang, J.: Universal output feedback controllers for nonlinear systems with unknown control direction. In: Proceedings of the 2003 American Control Conference, pp. 573–578 (2003)
6. Li, G., Lin, Y., Zhang, X.: Global state feedback stabilisation for a class of high-order nonlinear systems with quantised input and state. *Int. J. Syst. Sci.* **48**(5), 921–929 (2017)
7. Liu, W., Ho, D.W., Xu, S., Zhang, B.: Adaptive finite-time stabilization of a class of quantized nonlinearly parameterized systems. *Int. J. Robust Nonlinear Control* **27**(18), 4554–4573 (2017)
8. Liu, W., Lim, C.C., Xu, S.: Adaptive control of a class of quantised nonlinearly parameterised systems with unknown control directions. *Int. J. Syst. Sci.* **48**(5), 941–951 (2017)
9. Liu, Y.: Output-feedback adaptive control for a class of nonlinear systems with unknown control directions. *Acta Autom. Sin.* **33**(12), 1306–1312 (2007)
10. Nussbaum, R.D.: Some remarks on a conjecture in parameter adaptive control. *Syst. Control Lett.* **3**(5), 243–246 (1983)
11. Qian, C., Lin, W.: Non-Lipschitz continuous stabilizers for nonlinear systems with uncontrollable unstable linearization. *Syst. Control Lett.* **42**(3), 185–200 (2001)
12. Shang, F., Liu, Y., Li, C.: Global stabilization for feedforward nonlinear systems with unknown control direction and unknown growth rate. *Syst. Control Lett.* **111**, 58–63 (2018)

13. Xing, L., Wen, C., Wang, L., Liu, Z., Su, H.: Adaptive output feedback regulation for a class of nonlinear systems subject to input and output quantization. *J. Frank. Inst.* **354**(15), 6536–6549 (2017)
14. Xing, L., Wen, C., Zhu, Y., Su, H., Liu, Z.: Output feedback control for uncertain nonlinear systems with input quantization. *Automatica* **65**, 191–202 (2016)
15. Ye, X.: Asymptotic regulation of time-varying uncertain nonlinear systems with unknown control directions. *Automatica* **35**(5), 929–935 (1999)
16. Ye, X., Jiang, J.: Adaptive nonlinear design without a priori knowledge of control directions. *IEEE Trans. Autom. Control* **43**(11), 1617–1621 (1998)
17. Zhai, J., Du, H., Fei, S.: Global sampled-data output feedback stabilisation for a class of nonlinear systems with unknown output function. *Int. J. Control.* **89**(3), 469–480 (2016)
18. Zhou, J.: Decentralized adaptive control for interconnected nonlinear systems with input quantization. *IFAC-PapersOnLine* **50**(1), 10419–10424 (2017)
19. Zhou, J., Wen, C., Yang, G.: Adaptive backstepping stabilization of nonlinear uncertain systems with quantized input signal. *IEEE Trans. Autom. Control* **59**(2), 460–464 (2014)



Numerical Modelling and Velocity Tracking Control for Autonomous Heavy-Duty Trucks

Tongtong Lei and Mingxing Li^(✉)

The Seventh Research Division and the Center for Information and Control,
School of Automation Science and Electrical Engineering,
Beihang University (BUAA), Beijing 100191, People's Republic of China
lmx196@126.com

Abstract. In this paper, a numerical modelling method is presented and a sliding mode controller is designed for the heavy-duty truck. Based on the off-line data, a numerical model is established which is a simpler third-order system compared to the traditional longitudinal dynamic model in which the engine model is needed to be considered. And then a sliding mode controller with excellent velocity tracking performance is designed. To verify tracking performances, simulations based on a typical heavy-duty truck in Trucksim are implemented and results conclude that the new model and control strategy are very effective.

Keywords: Heavy-duty truck · Velocity tracking · Numerical modelling method

1 Introduction

Automatic driving technology especially for heavy-duty trucks in mine, goods transport systems and other applications has attracted a growing attention in the last years. As a basic and important aspect, velocity tracking problem must be solved to make the vehicle moving with the desired velocity which is given by the planning and decision module.

There are many control methods have been established to solved this problems for small passenger cars such as [1–7]. In [1], the PID control was used. To improve the controlled performance, fuzzy control [3], sliding mode control [4] and other nonlinear control methods such as parameter adaptive control and model predictive control were considered in [5–7]. However, the controlled system of the velocity control problem has a strong nonlinear nature [8] which makes it difficult to obtain satisfy tracking results [9].

Different from above results, in some other work such as [10] and our previous work [11], numerical modeling methods were established to control the velocity of the passenger car. Based on the speed-throttle-acceleration mapping and a

first-order linear time-invariant approach was presented in [10] and a third-order time varying approach was proposed in [11]. By using these methods, the higher accurately tracking performance is obtained while the complex parameters of the engine system are not needed to be known.

Compared with the small passenger car, the velocity tracking problem of the heavy-duty truck is quite different and is more difficult [12]. The main differences and difficulties are as following: firstly, the throttle control input is easy to saturate and the acceleration capability is very limited. Secondly, there are larger actuator time delays in both throttle control and brake control. Last but not least, mass has greater uncertainty. Thus, the nonlinear characteristics are stronger, the model is more complexity, and the velocity control problem is more challenging for the heavy-duty truck.

In this paper, a numerical model is established based on the off-line data which is good match with the vehicle dynamics. Then, based on this model, sliding mode controller is designed which has good tracking precision.

This paper is organized as follows: modeling of the heavy-duty truck is introduced in Sect. 2. A sliding mode controller is designed in Sect. 3. Simulating results are described in Sect. 4 and conclusion is made at last in Sect. 5.

2 Numerical Modeling

By using Newton's second law, the longitudinal dynamics is [5]:

$$\dot{v} = (F_v - F_a - F_r - F_g - F_e)/m \quad (1)$$

where m , v , F_v , F_a , F_r , F_g and F_e represent the mass, velocity, sum accelerating /braking force, aerodynamic drag force, rolling resistance force, gravitational force and the other external forces, respectively.

According to works [9] and [12], it is a complex nonlinear system while the driving force is taken as output variable and the throttle angle ψ_t and the braking pressure ψ_b are taken as input variables, i.e., $F_v = F_v(v, \psi_t, \psi_b)$. This relationship is almost not changing with the external environmental, the load and other external factors. Thus, if a heavy-duty truck is moving on horizontal ground without wind, then $F_g = F_e = 0$ and let $F_{vn} \triangleq F_v - F_a - F_r$, then we have:

$$\dot{v} = F_{vn}/m. \quad (2)$$

This system describes a common moving situation and it is easy to obtain that $F_{vs} = F_{vs}(v, \psi_t, \psi_b)$, then system (2) can be rewritten as following:

$$\begin{cases} \dot{v} = F_{vn}/m \\ \dot{F}_{vn} = \frac{\partial F_{vs}}{\partial v} F_{vs} + \frac{\partial F_{vs}}{\partial \psi_t} \dot{\psi}_t + \frac{\partial F_{vs}}{\partial \psi_b} \dot{\psi}_b \end{cases} \quad (3)$$

Notice that, operations of accelerating and braking cannot be implemented simultaneously. Hence we have $\dot{\psi}_t \dot{\psi}_b \equiv 0$. Let $u = \dot{\psi}_t$, $\psi = \psi_t$ while the vehicle is

accelerating, and $u = \dot{\psi}_b, \psi = \psi_b$ while the vehicle is braking. According to the results of [11], system (3) can be rewritten as follows:

$$\begin{cases} \dot{\psi} = u \\ \dot{v}_n = F_{vn} \\ \dot{F}_{vn} = \frac{\partial F_{vn}}{\partial v_n} F_{vn} + \frac{\partial F_{vn}}{\partial \psi} u \end{cases} \quad (4)$$

Assume that Lipschitz conditions of F_{vn} to v and ψ both hold, which means there are constants $l_v > 0$ and $l_\psi > 0$ such that:

$$\left| \frac{\partial F_{vn}}{\partial v} \right| \leq l_v, \left| \frac{\partial F_{vn}}{\partial \psi} \right| \leq l_\psi \quad (5)$$

Hence, $|\dot{F}_{vn}|$ is bounded, i.e., there is a real number $l_{vn} \in (0, \infty)$ such that $|\dot{F}_{vn}| \leq l_{vn}$. According to above results, and though driving by a experienced driver, we can get a data mapping $\psi_k = MP(\dot{v}_j, v_i)$ where k, i, j are sample points, and the following approaching system:

$$\dot{\psi} = u \quad (6)$$

$$\dot{v} = F_{vma}/m \quad (7)$$

$$\dot{F}_{vma} = f_{ij} F_{vma} + g_{ij} u \quad (8)$$

$$f_{ij} = \frac{F_{vn}(v_i, \psi_j) - F_{vn}(v_{i-1}, \psi_j)}{v_i - v_{i-1}} \quad (9)$$

$$g_{ij} = \frac{F_{vn}(v_i, \psi_j) - F_{vn}(v_i, \psi_{j-1})}{\psi_j - \psi_{j-1}} \quad (10)$$

To the above system, we can easy obtain that both f_{ij} and g_{ij} are bounded. From now on, we get dynamics model of F_{vma} and u . This dynamics is used to approximate to the dynamics of F_{vn} in (4).

3 Design of the Velocity Tracking Controller

The aim of this section is to design a controller such that the heavy-duty truck can track the desired velocity v_d . From Eq. (1)–(2), the controlled system can be rewritten into:

$$\dot{v} = F_{vn}/m - F_g/m - F_e/m \quad (11)$$

To this system, v and \dot{v} are both measurable. Let:

$$e = v - v_d \quad (12)$$

then $\dot{e} = \dot{v} - \dot{v}_d$. Furthermore, let $\omega = (F_{vn} - F_{vma} - F_g - F_e)/m$ then ω is bounded and we have the following equation:

$$\begin{cases} \dot{v} = F_{vma}/m + \omega \\ \dot{F}_{vma} = f_{ij} F_{vma} + g_{ij} u \end{cases} \quad (13)$$

To the above system, we according to the data mapping can get $\psi_d \approx MP(\dot{v}_d, v_d)$ and the following error system:

$$\begin{cases} \dot{e} = F_e - \omega \\ \dot{F}_e = f_{ij}F_e + g_{ij}u. \end{cases} \quad (14)$$

where $e(t_0) = 0$, $F_e(t_0) = 0$ and $\psi = \psi_d + \int_{t_0}^t u d\tau$. Since only e and \dot{e} are measurable, the following sliding mode surface is taken:

$$s = \int_{t_0}^t K_p e d\tau + e \quad (15)$$

where $K_p > 0$. It is easy to obtain that:

$$\begin{aligned} \dot{s} &= K_p e + \dot{e} \\ &= K_p e + F_e - \omega \\ &= K_p e + (1 - f_{ij})F_e + f_{ij}F_e - \omega \\ &= K_p e + (1 - f_{ij})F_e - g_{ij}u + \dot{F}_e - \omega \\ &= K_p e + (1 - f_{ij})(\dot{e} + \omega) - g_{ij}u + \dot{F}_e - \omega \\ &= K_p e + (1 - f_{ij})\dot{e} - g_{ij}u + \dot{F}_e - f_{ij}\omega \\ &= K_p e + \dot{e} - f_{ij}\dot{e} - g_{ij}u + \dot{F}_e - f_{ij}\omega \end{aligned}$$

Thus, if u is taken as following:

$$u = -g_{ij}^{-1}(-K_p e - (1 - f_{ij})\dot{e} - K_I s - D \text{sign}(s)) \quad (16)$$

where $D > \sup |\dot{F}_e - f_{ij}\omega|$, then we have:

Theorem 1. *If the control input u is taken as Eq. (16), then the controlled error system (14) is exponentially stable.*

Proof. Take the Liapunov function as following form:

$$V = \frac{1}{2}s^2$$

then take the derivative of V along system (14) while u is taken as (16), we have:

$$\begin{aligned} \dot{V} &= s\dot{s} \\ &= s(K_p e + \dot{e} - f_{ij}\dot{e} - g_{ij}u + \dot{F}_e - f_{ij}\omega) \\ &= s(-K_I s - D \text{sign}(s) + \dot{F}_e - f_{ij}\omega) \\ &= -K_I s^2 - D|s| + (\dot{F}_e - f_{ij}\omega)s \end{aligned}$$

Because of $D > \sup |\dot{F}_e - f_{ij}\omega|$, the following inequality is established:

$$\dot{V} = s\dot{s} < -K_I s^2 = -2K_I V < 0$$

Thus, system (14) is exponentially stable while u is taken as (16).

4 Simulations

To verify the control effect of our control scheme, the model of three axle conventional van in Trucksim2016 is used. As shown in Fig. 1, the nonlinear characteristics are very strong between throttle input and output. And the nonlinear characteristics are stronger while value of the throttle angle is taken small such as 0.1 and big such as 0.9 and 1.0. As shown in Fig. 2, to brake input, there are also strong nonlinear characteristics. Points marked by red triangle are the ABS starting points. After these points, the ABS is working and the nonlinear becomes more serious. Hence, there will be conflicts between the ABS and the velocity tracking controller.

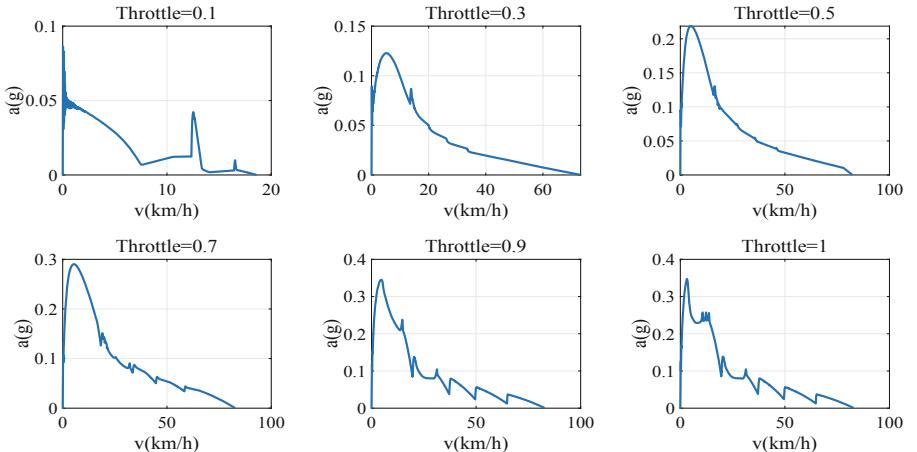


Fig. 1. Relationship of v , \dot{v}_d and throttle input

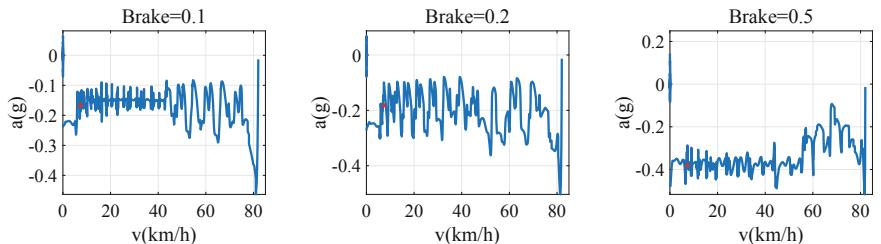


Fig. 2. Three special results of outputs v , \dot{v}_d and brake input

To solve above problems, the feed forward compensation is established by a online data mapping among \dot{v} , v and ψ . To set up this data mapping, we set the

throttle input in a fixed value to collect the data of v and v_d . And there are ten fixed values of throttle input, which are 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0. To the brake situation shown by Fig. 3, outputs v and \dot{v} are changing almost the same with each other while the brake input value are bigger than 0.3. Thus, there are only three fixed values which are 0.1, 0.2 and 0.5 in the data mapping. By doing this, the data mapping of \dot{v} , v and ψ is established. According to the desired value of v_d , \dot{v}_d and the data mapping of v , \dot{v} and ψ , the moving situation such as braking or accelerating and the feed forward input value of ψ_d can be obtained to lookup the data mapping.

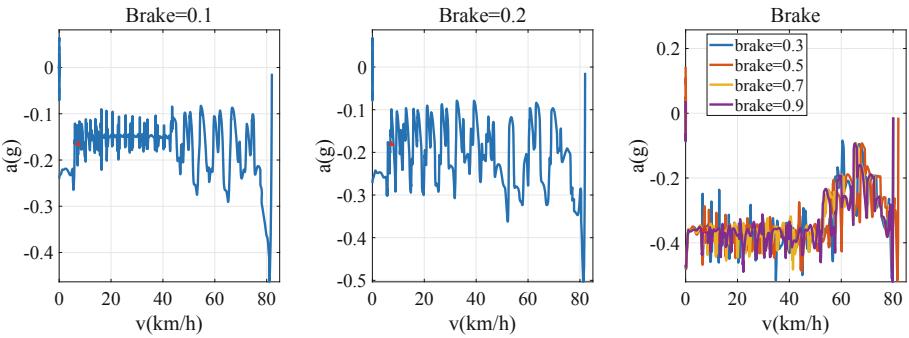


Fig. 3. Relationship of v , \dot{v}_d and brake input

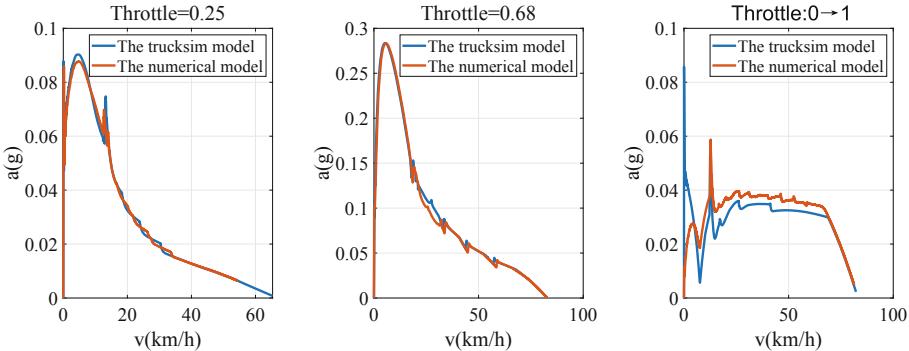


Fig. 4. Verification of the numerical model in accelerating situation

In order to test and verify the effectiveness of the above feed forward compensation method, two kinds of fixed value inputs and one kind varying inputs of both accelerating and braking situations are simulated. Simulation results are shown by Fig. 4 and Fig. 5. As shown in Fig. 4, the feed forward compensation

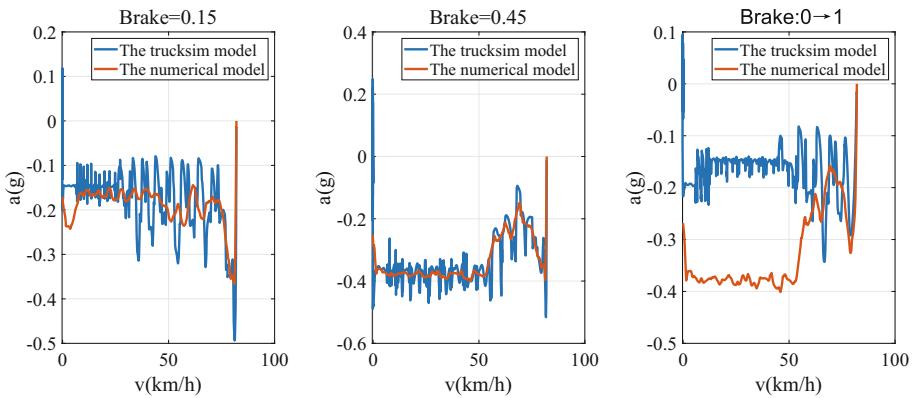


Fig. 5. Verification of the numerical model in braking situation

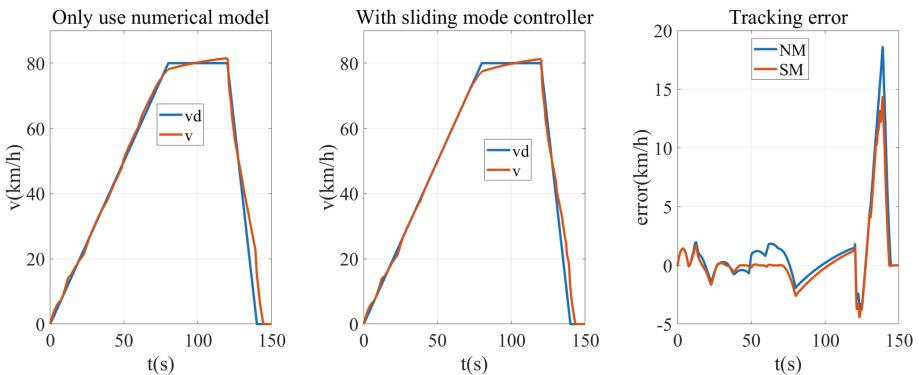


Fig. 6. Velocity tracking results with the sliding mode controller (16)

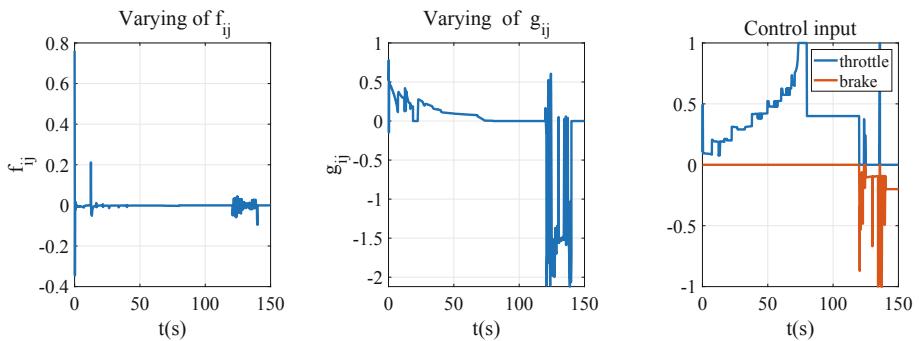


Fig. 7. Inputs and f_{ij} and g_{ij} with the sliding mode controller (16)

by the data mapping is almost can be tracking the desired value of the velocity and acceleration in the accelerating situation. And compared to the varying input, the tracking accuracy is higher while inputs are fixed values. As shown in Fig. 5, the feed forward compensation results are not so good in the braking situation. And the tracking accuracy is even worse while inputs are varying values. However, tracking errors are bounded and acceptable.

To test and verify the effectiveness of the sliding mode controller, simulations shown by Fig. 6 and Fig. 7 are done. The tracking results with sliding mode controller are shown by Fig. 6. These results are shown that the tracking accuracy is improved significantly while the sliding mode controller (16) is used. Control inputs and varying of f_{ij} and g_{ij} are shown by Fig. 7. In Fig. 6 and Fig. 7, tracking errors are big and there is a high gain fluttering of the input signal and parameter f_{ij} and g_{ij} while the truck is braking. Because of ABS is conflict with velocity tracking control meanwhile big actuator time delays exist. Even though, from Fig. 6 we can conclude that the track results are improved while the sliding controller is used. Thus, based on the above results, we conclude that the effectiveness of our control scheme is verified. And with the sliding mode controller, the tracking accuracy is improved.

5 Conclusions

In this paper, the velocity tracking problem of the heavy-duty truck is studied. A numerical model is established and a sliding mode controller is designed. From the simulating results, the high accurately tracking result is obtained and the effectiveness of our control scheme is verified. Furthermore, the coupling performance of the lateral dynamics is not considered in this paper. It should be studied further in the future.

References

1. De Santis, R.M.: A novel PID configuration for speed and position control. *J. Dyn. Syst. Meas. Control* **116**(3), 542–549 (1994)
2. Tian, L., Hao, W., Zhu, X.: Research on cruise control system based on fuzzy adaptive method. *Autom. Instrum.* **9**, 11–14 (2014)
3. Naranjo, J.E., Gonzalez, C.: ACC+Stop&Go maneuvers with throttle and brake fuzzy control. *IEEE Trans. Intell. Transp. Syst.* **7**(2), 213–225 (2006)
4. Li, Y., He, L.: Counterbalancing speed control for hydrostatic drive heavy vehicle under long down-slope. *IEEE/ASME Trans. Mechatron.* **20**(4), 1533–1542 (2015)
5. Kim, H., Kim, D., Shu, I., Yi, K.: Time-varying parameter adaptive vehicle speed control. *IEEE Trans. Veh. Technol.* **65**(2), 581–588 (2016)
6. Milanés, V., Villagrá, J., Pérez, J., González, C.: Low-speed longitudinal controllers for mass-produced cars: a comparative study. *IEEE Trans. Ind. Electron.* **59**(1), 620–628 (2012)
7. Li, S., Li, K., Rajamani, R., Wang, J.: Model predictive multi-objective vehicular adaptive cruise control. *IEEE Trans. Control Syst. Technol.* **19**(3), 556–566 (2011)

8. Kuutti, S., Bowden, R., Jin, Y., et al.: A survey of deep learning applications to autonomous vehicle control. *IEEE Trans. Intell. Transp. Syst.* 1–22 (2020)
9. Li, S.E., Gao, F., Cao, D., Li, K.: Multiple-model switching control of vehicle longitudinal dynamics for platoon-level automation. *IEEE Trans. Veh. Technol.* **65**(6), 4480–4492 (2016)
10. Xu, S., Peng, H., Song, Z., Chen, K., Tang, Y.: Accurate and smooth speed control for an autonomous vehicle. In: *IEEE Intelligent Vehicles Symposium (IV)*, pp. 1976–1982, June 2018
11. Li, M., Jia, Y.: Velocity tracking control based on throttle-pedal-moving data mapping for the autonomous vehicle. In: *Proceedings of 2019 Chinese Intelligent Systems Conference, CISC 2019*, Hainan, China, 25–27 October 2019
12. Lu, X., Hedrick, J.K.: Longitudinal control design and experiment for heavy-duty trucks. In: *Proceedings of the American Control Conference*, Denver, Colorado, USA, 4–6 June 2003



A Scheduling Problem of Joint Mobile Charging Sequence and Charging Start Time Control in Wireless Rechargeable Sensor Networks

Tao Cao^{1,2}, Chengpeng Jiang^{1,2}, and Wendong Xiao^{1,2(✉)}

¹ School of Automation and Electrical Engineering,
University of Science and Technology Beijing, Beijing 100083, China
 [{caotao3019,jcp09868}@163.com,](mailto:{caotao3019,jcp09868}@163.com)
wdxiao@ustb.edu.cn

² Beijing Engineering Research Center of Industrial Spectrum Imaging,
Beijing 100083, China

Abstract. The development of wireless power transmission technology provides a new method to the energy constraint problem of sensor networks. Extensive of research has been done to optimize the charging sequence of mobile charger to improve the charging efficiency and to prolong the life cycle of wireless rechargeable sensor networks (WRSN). Compared with the traditional research, this paper allows the nodes in the network to stop working, that is, the node energy is exhausted, and then it can resume working after getting the energy replenishment. At the same time, based on the consideration of charging efficiency, the concept of charging waiting time is introduced, and finally a novel mobile energy replenishment problem is proposed, that is the joint mobile charging sequence and charging start time control (JMCSCSTC) problem. We first build the system models and formulate the problem, then design three algorithms to address this problem. Finally a large number of simulation experiments are carried out to learn the impact of waiting time on the charging performance. At the same time, we change the parameters to compare the impact on the performance index, and verify the feasibility of the algorithm.

Keywords: Wireless rechargeable sensor network · Wireless energy transfer · Mobile charger · Scheduling

1 Introduction

Wireless sensor network (WSN) is composed of a large number of sensor nodes deployed in the monitored area. These nodes form a network system, and the information of the monitored objects in the network coverage area can be

acquired, collected and processed [1]. Nowadays, WSN has an important application in many fields, even has a unique and irreplaceable role, specific application fields include marine monitoring, indoor positioning, and human health monitoring [2].

However, the limited energy capacity of nodes becomes an important factor to restrict the WSN's continuous operation. To solve this problem, the first thing we think about is replacing the battery of the node, but frequent battery replacement is impractical in complex environments. Another way is to replenish energy through natural energy, such as solar energy, wind energy [3], but this method is uncontrollable and greatly influenced by the environment. Above methods cannot fundamentally solve the problem of node energy replenishment.

The continuous development of wireless power transmission (WPT) technology [4] brings a new way to replenish the node, that is, the car equipped with WPT technology charge each node by moving in WSN, many researchers have paid attention to this field. We use this method in this paper and call WSN which uses WPT technology to replenish energy as wireless rechargeable sensor network (WRSN). The car equipped with WPT technology is called as mobile charger (MC).

The focus of mobile energy replenishment problem is designing a charging scheduling scheme for MC to achieve the optimal charging performance. There have been a lot of studies on this issue, but in the traditional research, once the node energy is exhausted in WRSN, it is considered that WRSN stops working and the corresponding MC scheduling scheme fails. In previous studies, MC charges the node as soon as it reaches its vicinity, but there may be some residual energy in the node at this time. Considering the above two aspects, we propose a new mobile energy replenishment problem, that is, the joint mobile charging sequence and charging start time control (JMCSCSTC) problem.

The structure of this paper is as follows: in Sect. 2, the related works are introduced. In the Sect. 3, the system models are established, and the new problem are formulated. In Sect. 4, three algorithms are introduced. In Sect. 5, based on the proposed algorithm, extensive of simulation experiments are carried out. Finally, in Sect. 6, the full text is summarized and the future research is prospected.

2 Related Work

In this article, we consider using a mobile charger (MC) with WPT technology to charge each sensor in the WRSN. Researchers have done a lot of research on mobile charging scheduling, next, we divide the mobile charging scheduling problem into single MC and multiple MCs cases according to the number of MC used in the charging process, and the following is a brief description of the relevant research:

2.1 Single MC

At the beginning of the study, the researchers considered using only one MC to replenish the energy of sensors in WRSN, and MC can carry infinite energy.

For example, in [5], the author introduced the concept of renewable energy cycle. In this charging scheme, MC periodically accesses all nodes in WRSN. After that, researchers began to consider the limited energy of MC, Liang et al. [6] considered the problem that if all nodes are need charged in one cycle, which the node should be charged, to solve this problem, the author designed an approximate algorithm. In [7], Zhao et al. considered that MC can not only replenish the energy of nodes, but also collect sensor data from nodes, the energy replenishment and data collection are jointly designed by MC mobility.

In the above article, the MC has determined the scheme to charge the nodes, it's an offline charging mode, but this scheme cannot meet the dynamic characteristics of WRSN. So He et al. [8] first raised the problem of on-demand charging, and gave an online scheduling method Nearest-Job-Next with Pre-emption (NJNP). Feng et al. [9] realized that the online charging mode also has disadvantages, so the author proposed a mobile charging algorithm based on hybrid mode, namely MERSH.

2.2 Multiple MCs

With the expansion of network scale, single MC cannot guarantee the charging demand of nodes in WRSN, therefore, researchers consider using multiple MCs to replenish energy for nodes in WRSN. In [10], Xu et al. designed two algorithms for fixed and variable cycle scenarios, the optimization goal is to minimize the total moving distance of MCs. Mo et al. [11] presented an optimal multiple MCs coordination (OMC) algorithm, the algorithm divides the problem into two parts: multiple MCs scheduling based on integer linear programming and MCs charging time and moving time based on linear programming. Dai et al. [12] considered the minimum MCs problem (MinMCP), which is to find the minimum number of MCs needed under the premise of ensuring the normal operation of WRSN.

Then the researchers considered the energy transfer between multiple MCs. Zhang et al. [13] first proposed a cooperative mobile charging scheme, which is to allow energy to be intentionally transferred between MCs. Han et al. [14] considered that the nodes had the different energy consumption rate, and the WRSN is divided into three regions, then proposed multi-charger cooperation recharging algorithm.

3 Problem Formulation

In this section, we present system models, and describe the problem studied in this paper:

3.1 WRSN Model

We use a weighted undirected graph to describe the network. WRSN can be described as $G = (SN, L, P_c, E)$, $SN = \{s_1, s_2, \dots, s_i, \dots s_N\}$ represents a set of

sensor nodes, s_i means the i -th sensor node, $1 \leq i \leq N$; $L = \{l_1, l_2, \dots, l_i, \dots, l_N\}$, l_i represents the i -th node corresponding positions. This paper assumes that the energy consumption rate of the node is constant and different, using $P = \{p_1, p_2, \dots, p_i, \dots, p_N\}$ represents the energy consumption rate set of nodes, p_i means the energy consumption rate of the i -th sensor node. The residual energy of each node is different when MC starts charging, use $E = \{e_1, e_2, \dots, e_i, \dots, e_N\}$ to represent the initial energy set of nodes.

3.2 Energy Consumption and Replenishment

On the basis of the above network model, we use the way of mobile energy replenishment to charge the nodes in the network. We assume that MC can obtain all kinds of information in WRSN from the base station and MC has sufficient energy. Each node has a corresponding energy exhausted time t_e , energy exhausted time t_{ei} of the i -th node is calculated by the following formula:

$$t_{ei} = \frac{e_i}{p_i} \quad (1)$$

MC's energy has two directions, which are energy replenishment for nodes and energy consumption in the process of moving. We assume that MC's moving speed V_c and the charging power U of MC is constant. MC charges the nodes in turn through the designed charging sequence of nodes, we set the charging sequence as $R_k = \{\pi_1, \pi_2, \dots, \pi_i, \dots, \pi_N\}$, $1 \leq k \leq N$, and the number in the charging sequence corresponds to the node number in SN .

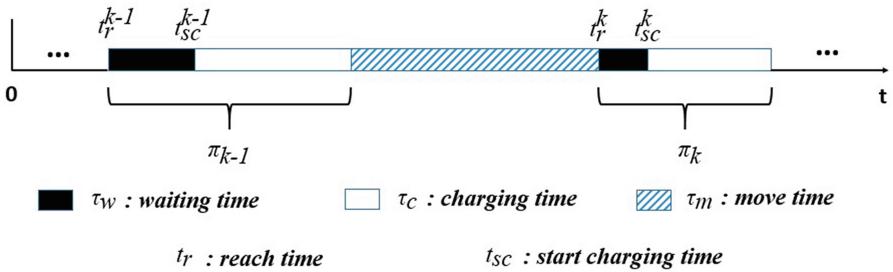


Fig. 1. The time consumption of MC charging for two adjacent nodes π_{k-1}, π_k in the charging set.

In Fig. 1, we describe the time consumption of MC charging for two adjacent nodes π_{k-1}, π_k in the charging set. In order to compare with the exhaustion time of the nodes mentioned above, we introduce the starting charging time t_{sc} of the nodes, the concept of t_{sc} is the time when a node starts charging. This time is related to two parts of time, namely MC arrival time t_r and MC waiting time τ_w , τ_w means that when MC arrives at a node, it will wait a certain time to

charge the node based on the consideration of maximizing charging efficiency, we use the following formula to calculate the starting charging time:

$$t_{sc}^k = t_r^k + \tau_w^k \quad (2)$$

From Fig. 1, we can see that the t_r^k is related to the time consumption of the k -1th node and the distance between nodes π_{k-1}, π_k . Firstly, the distance set $D = \{d_{i,j} = d(l_i, l_j) \mid l_i, l_j \in L\}$ between nodes is defined, so the moving time $\tau_m^{k-1,k}$ of MC between nodes π_{k-1}, π_k is expressed:

$$\tau_m^{k-1,k} = \frac{d_{\pi_{k-1}, \pi_k}}{V_c} \quad (3)$$

For each node, after knowing its initial energy information, the charging time τ_c can be calculated by (4):

$$\tau_c^k = \frac{E_{max} - e_k}{U} \quad (4)$$

With the above definition, we use the following formula to express t_r^k :

$$t_r^k = t_{sc}^{k-1} + \tau_w^{k-1} + \tau_c^{k-1} + \tau_m^{k-1,k} \quad (5)$$

we will combine (2) and (5):

$$t_{sc}^k = t_{sc}^{k-1} + \tau_w^{k-1} + \tau_c^{k-1} + \tau_m^{k-1,k} + \tau_w^k \quad (6)$$

Table 1. Notation definition

Symbol	Quantity
d_{π_{k-1}, π_k}	The distance between sensor node s_{k-1} and s_k
E_{max}	The maximum energy capacity of the sensor node
e_i	The residual energy of sensor node s_i
p_i	The energy consumption rate of sensor node s_i
U	MC charging speed
V_c	MC moving speed
N	Number of nodes in WRSN
t_e	Energy exhaustion time
t_r^k	The moment when the MC reaches the k -th node
τ_c^k	The energy replenishment duration for the k -th sensor node
$\tau_m^{k-1,k}$	The MC move time between sensor node s_{k-1} and s_k
τ_w^k	The waiting duration about the k -th node when the MC arrive it
t_{sc}^k	The start energy replenishment time of the k -th sensor node
t_{ed}	Energy exhausted duration
T_{ed}	Total energy exhausted duration of network

In this paper, we allow the node to stop working, that is, when the node's energy is exhausted. At the same time, we assume that the node can resume working after replenish energy. In order to evaluate this situation, the concept of energy exhausted duration t_{ed} is defined. The sum of all nodes t_{ed} is called the total energy exhausted duration T_{ed} of the system, we hope that the value is as small as possible, and the relevant calculation formula is as follows:

$$t_{ed}^k = \begin{cases} t_{sc}^k - t_e^k & t_{sc}^k \geq t_e^k \\ 0 & t_{sc}^k < t_e^k \end{cases} \quad (7)$$

$$T_{ed} = \sum_{k=1}^N t_{ed}^k \quad (8)$$

The symbols and definitions used in this paper are shown in Table 1.

3.3 Problem Statement and Formulation

In this paper, we define a problem called: the joint mobile charging sequence and charging start time control (JMCSCSTC) problem. The problem is described as follows: The nodes in WRSN are randomly distributed in two-dimensional closed area, and the nodes use MC with enough energy to replenish the energy, in which the energy consumption rate of each node is different, and there are different initial energy. In the case of allowing nodes to exhaust energy, we hope to find a charging scheme, which has higher charging efficiency and guaranteed the performance of the network. Figure 2 shows an example of JMCSCSTC problem.

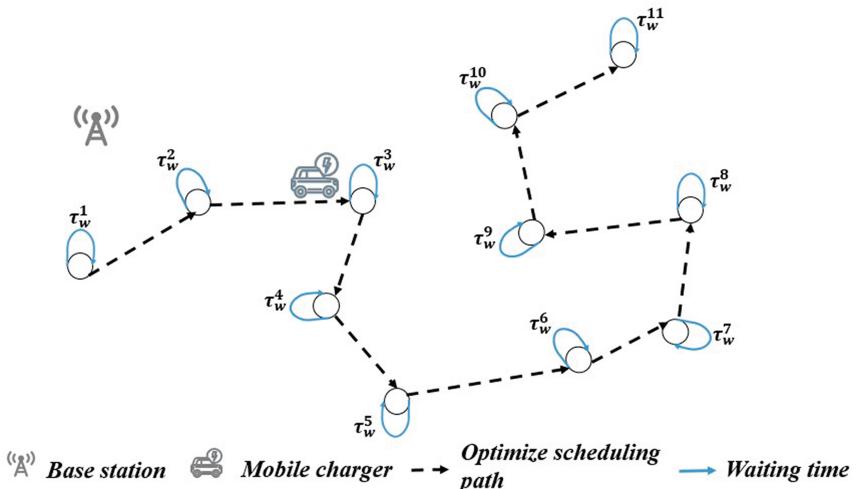


Fig. 2. The time consumption of MC charging for two adjacent nodes π_{k-1}, π_k in the charging set.

In the traditional mobile charging scheduling problem, node stop working due to energy exhaustion is not allowed, it represents the charging scheme fails. However, in this paper, this situation is permissible, and we assume that exhausted nodes will resume working when they get energy replenishment later. In order to describe this process, we build an index to evaluate the charging performance: charging quality of network (CQN):

$$CQN(k) = (t_{sc}^k - t_e^k)^2 \quad (9)$$

where $CQN(k)$ represents the charging mass of the k -th node, and the formula (10) represents the total charging quality of network ($TCQN$):

$$TCQN = \sum_{k=1}^N CQN(k) \quad (10)$$

Node stop working will inevitably affect the performance of WRSN, so we hope that the duration of the exhausted nodes is as short as possible. On the other hand, this paper considers the situation that the node has a waiting time. In short, our optimization goal is to minimize $TCQN$. Each charging scheme has a node charging sequence set R_k . All possible charging order of nodes is defined as S , so JMCSCSTC problem can be described by the following formula:

$$\begin{aligned} OPT : \min TCQN(R_k) &= \min \sum_{k=1}^N CQN(k) \\ s.t. : R_k &\in S \subseteq R^N \end{aligned} \quad (11)$$

4 Algorithme Design

This section describes the algorithm of charge scheduling:

4.1 Particle Swarm Optimization Approach for JMCSCSTC

When PSO is used to address the JMCSCSTC problem in this paper, firstly, assume that there are m particles in the N -dimensional search space, and each particle represents an MC charging sequence R_k . Each particle has two attributes: position and speed, the position of the i -th particle is expressed as $x_i = (x_{i1}, x_{i2}, \dots, x_{iN})$, $i = 1, 2, \dots, m$, and the velocity is expressed as $v_i = (v_{i1}, v_{i2}, \dots, v_{iN})$, $i = 1, 2, \dots, m$, each particle has a corresponding fitness. In this paper, the fitness is $TCQN(x_i)$, so we convert (11) to the following formula:

$$\begin{aligned} OPT : \min TCQN(x_i) &= \min \sum_{j=1}^N CQN(x_{ij}) \\ s.t. : x_i &\in S \subseteq R^N \end{aligned} \quad (12)$$

In the process of PSO iterative calculation, we need to record the historical optimal position of each particle and the historical optimal position of the whole population. We use t to express the number of iterations, and record the optimal position of the i -th particle is $P_{iN} = (P_{i1}, P_{i2}, \dots, P_{iN})$, and the optimal position of the population is $P_{gN} = (P_{g1}, P_{g2}, \dots, P_{gN})$, and the fitness is $TCQN(P_{gN})$.

P_{iN} needs to be updated after each iteration, so we use the following formula to update:

$$P_{iN}(t+1) = \begin{cases} P_{iN}(t) & TCQN(x_i(t+1)) \geq TCQN(P_{iN}(t)) \\ x_i(t+1) & TCQN(x_i(t+1)) < TCQN(P_{iN}(t)) \end{cases} \quad (13)$$

The position and velocity of particles will be updated after each iteration. The updating formula is as follows:

$$v_i(t+1) = v_i(t) + c_1 * r_1 * [P_{iN}(t) - x_i(t)] + c_2 * r_2 * [P_{gN}(t) - x_i(t)] \quad (14)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (15)$$

where c_1 and c_2 is the learning factor, which is a non negative constant, r_1 and r_2 is a random number uniformly distributed on $[0,1]$, and the operation process of the algorithm is described in Algorithm 1.

Algorithm 1: Particle swarm optimization approach for JMCSCSTC problem

Input: Known parameters in the system;

Step 1: Determine the number of particles in the population, initialize the position x_i and speed v_i of each particle;

Step 2: Use formulas (9) and (10) to calculate the fitness $TCQN(x_i)$ of each particle;

Step 3: Calculate the optimal position P_{iN} of each particle and the optimal position P_{gN} of the population;

Step 4: The optimal position of each particle is updated by formula (13), and then the optimal position of the population is updated;

Step 5: The particle velocity is updated by formula (14), the particle position is updated by formula (15);

Step 6: Determine whether the iteration termination condition is met. If it is, the optimal position of population P_{gN} will be output. Otherwise, return to Step 2.

4.2 Genetic Algorithm for JMCSCSTC

The second algorithm selected in this paper is genetic algorithm (GA). We code the N features in this study to obtain chromosome X_i by integer coding. In this problem, a charging sequence set R_k is a chromosome, and the fitness is used to evaluate the chromosome quality. In this paper, $TCQN(X_i)$ is used, so formula (11) can be rewritten as:

$$\begin{aligned} OPT : \min & \quad TCQN(X_i) \\ s.t. : & \quad X_i \in S \subseteq R^N \end{aligned} \quad (16)$$

Algorithm 2: Genetic algorithm for JMCSCSTC problem

-
- Input: Known parameters in the system;
- Step 1: Integer coding of N features in JMCSSCT problem, producing m chromosomes as initial population;
- Step 2: The fitness $TCQN(X_i)$ of each chromosome X_i was calculated by formulas (9) and (10);
- Step 3: Selection of parent chromosome by roulette selection method based on formula (17);
- Step 4: Perform cross operation on the selected chromosome;
- Step 5: Perform mutation operation on the selected chromosome;
- Step 6: Determine whether the iteration termination condition is met. If it is, Output convergence result. Otherwise, return to Step 2.
-

The genetic operation of GA includes three operators: selection operator, crossover operator and mutation operator. The selection operator uses roulette selection method to select a certain number of chromosomes in the population with the probability calculated by formula (17). The crossover operator is used to replace and recombine the genes of the selected chromosomes; mutation operator is to change a gene value of chromosome. The specific operation process is shown in Algorithm 2.

$$P_i = \frac{TCQN(X_i)}{\sum_{i=1}^m TCQN(X_i)} \quad (17)$$

4.3 Greedy Algorithm for JMCSCSTC

The basic idea of greedy algorithm is to make the best choice in the current situation all the times when solving the problem. To address the problem as shown in (11), we first get the energy exhaustion time set SE of all nodes, and arrange the elements in it from small to large to form a new set SE' . The node sequence number set corresponding to this set is R_k . MC charges according to the set order, and the corresponding fitness is $TCQN(R_k)$. The specific algorithm steps are shown in Algorithm 3.

Algorithm 3: Greedy algorithm for JMCSCSTC

-
- Input: Known parameters in the system;
- Step 1: Obtain the energy exhaustion time set SE of all nodes;
- Step 2: Get the set SE' by sorting the elements in SE from small to large;
- Step 3: The charging sequence R_k is obtained according to the node serial number corresponding to the element in SE' , the fitness $TCQN(R_k)$ is calculated.
-

5 Evaluation Performance

To evaluate the performance of the algorithm mentioned above, we have carried out a lot of simulation experiments. Firstly, some parameters before the experiment are described in Sect. 5.1, and four experiments are carried out in the following four sections. These four experiments consider the influence of four parameters on the charging performance, which are the number of network nodes and the maximum carrying energy of nodes in Sects. 5.2 and 5.3, and the MC moving speed and charging speed in Sects. 5.4 and 5.5.

5.1 Simulation Settings

In all simulation experiments, the positions of nodes are randomly generated, they are evenly distributed in 20 closed area. The initial residual energy and energy consumption rate of each node are also randomly generated, and they are different. Other initial parameter settings are shown in the Table 2.

The four experiments in this section compare the effects of four parameters. In addition, the concept of waiting time is defined in JMCSCSTC, in order to explore the impact of waiting time on the charging performance, we define a wireless mobile charging sequence scheduling (WMCSS) problem that does not consider waiting time. In each experiment, the charging performance of these two problems is compared. In the following simulation diagram, GREEDY-JMCSCSTC represents solving JMCSCSTC problem with greedy algorithm, and so on.

For each experiment, we have carried out many experiments, and the experimental results take the average value of the results.

Table 2. Notation definition

Parameters	Description
Coverage area	20 m 20 m
Nodes distribution	Random distribution
Original number of sensor node	$N = 20$
Maximum energy capacity of node	100 J
The moving speed of MC	8 m/s
The charging speed of MC	100 J/s

5.2 Performance with the Size of the Network

It can be seen from Fig. 3 (a) (c) that with the increase of the number of nodes, $TCQN$ and total energy exhaustion time will increase, and the WMCSS problem without considering the waiting time will increase much more than JMCSCSTC

problem which considering the waiting time, considering wait time is significantly better than not. It is shown that considering the waiting time can effectively improve the charging performance when the number of nodes increases. In addition, it can also be seen that GA algorithm and PSO algorithm are better than greedy algorithm in improving $TCQN$ and total energy exhaustion time for WMCSS problem.

Figure 3 (b) shows that the total moving distance of MC also increase with the number of nodes increases, and for WMCSS problem, the greedy algorithm lags behind PSO and GA algorithms in optimizing the total distance of MC with the increase of the number of nodes.

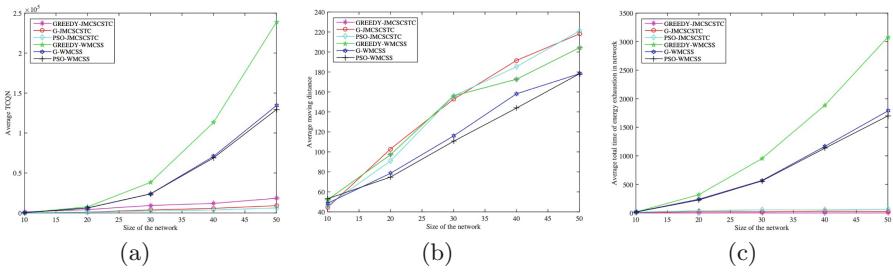


Fig. 3. Performance with different maximum energy capacity. (a) $TCQN$. (b) Moving distance (c) Total time of energy exhaustion in network.

5.3 Performance with Different Maximum Energy Capacity

It can be seen from Fig. 4 (a) (c) that for WMCSS problems, the $TCQN$ and energy exhaustion time increases with the increase of the maximum energy capacity of the node, because with the increase of the maximum energy capacity of the node, the charging time of the node will increase correspondingly, and the probability of energy exhaustion will increase greatly for the subsequent nodes, leading to the decrease of charging performance. For JMCSCSTC problem, increasing node maximum carrying energy also affects $TCQN$ and total energy consumption time, but the impact is much less than WMCSS problem, indicating that the addition of waiting time will effectively reduce the impact of node maximum carrying energy on charging performance.

From Fig. 4 (b), it can be seen that the total moving distance of MC fluctuates with the increase of the maximum carrying energy of the node, and a dramatic increase in MC moving distance occurs when the maximum carrying energy of a node is between 100 and 120. From Fig. 4 (b), it is obvious that for WMCSS problem, greedy algorithm lags behind GA algorithm and PSO algorithm in the optimization of MC total distance.

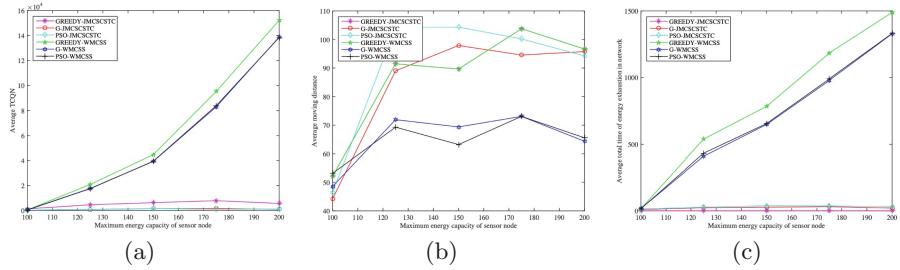


Fig. 4. Performance with different numbers of nodes. (a) $TCQN$. (b) Moving distance (c) Total time of energy exhaustion in network.

5.4 Performance with Different Moving Speed

From Fig. 5 (a) (c), it can be seen that with the change of MC moving speed, the charging performance of JMCSCSTC problem is better than that of WMCS problem as a whole, indicating that the addition of waiting time will improve the charging performance of the system. The WMCS problem experiment in Fig. 5 (a) (c), with the increase of MC's moving speed, the total energy exhaustion time of nodes and $TCQN$ decreased significantly, because with the increase of MC's moving speed, the moving time between nodes decreased, and the exhausted nodes can be charged earlier, resulting in the decrease of $TCQN$ and total energy exhaustion time.

Figure 5 (b) shows that the total moving distance generally decreases with the increase of MC moving speed, but the changing process is fluctuating. For WMCS problem, PSO algorithm is better than GA algorithm and greedy algorithm at the moving distance between 8 and 10.5.

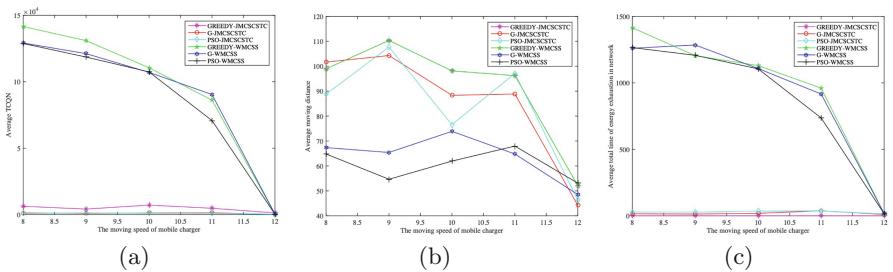


Fig. 5. Performance with different moving speed of MC. (a) $TCQN$. (b) Moving distance (c) Total time of energy exhaustion in network.

5.5 Performance with Different Charging Speed

From Fig. 6 (a) (c), we can see that $TCQN$ and total energy exhaustion time decrease with the increase of MC charging speed, because MC charging speed

affects the charging time of nodes. In contrast, the greater MC charging speed is, the shorter node charging time is, and MC can charge subsequent nodes earlier, which will lead to the reduction of the total energy exhaustion time and $TCQN$.

It can be seen from Fig. 6 (b) that with the increase of charging speed of MC, the total moving distance of MC will increase first and then decrease. When charging speed $U = 14$, the total moving distance of MC will decrease significantly when U continues to increase. From the algorithm point of view, the optimization results of greedy algorithm lag behind the other two algorithms as a whole for WMCSS problem.

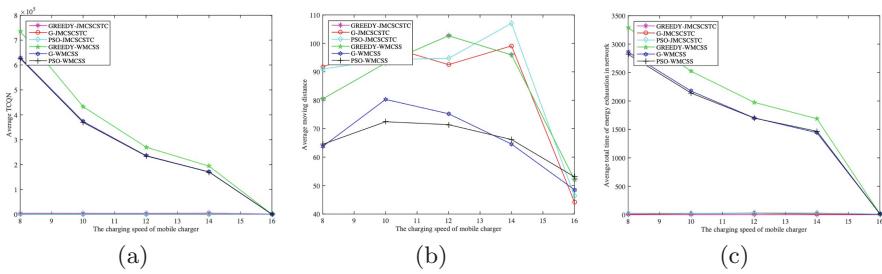


Fig. 6. Performance with different charging speed of MC. (a) $TCQN$. (b) Moving distance (c) Total time of energy exhaustion in network.

6 Conclusion

In this paper, the joint mobile charging sequence and charging start time control (JMCSCSTC) problem is proposed. Then, three algorithms are proposed to address this problem. Compared with the previous research, this problem allows nodes to be exhaustion and can resume working when they get energy replenishment later. Then, to achieve better charging efficiency and charging performance, the concept of waiting time is introduced. In this paper, the system models are established firstly, and the problems are described. Then three algorithms are introduced. Finally, plenty of simulations are carried out to evaluate the rationality of the algorithm.

From the results, the introduction of the concept of waiting time can get better charging performance, and the proposed algorithm also gets better optimization results. However, compared with the actual system, more in-depth research is needed. For example, it can be considered to charge some energy to the nodes, rather than the maximum energy of all nodes, so as to reduce the time that the subsequent nodes are in the state of energy exhaustion.

References

1. Mohamed, R.E., Saleh, A.I., Abdelrazzak, M., Samra, A.S.: Survey on wireless sensor network applications and energy efficient routing protocols. *Wirel. Pers. Commun.* **101**, 1019–1055 (2018)
2. Khalil, I.M., Khreishah, A., Ahmed, F., Shuaib, K.: Dependable wireless sensor networks for reliable and secure humanitarian relief applications. *Ad Hoc Netw.* **13**, 94–106 (2014)
3. Shu, Y.C., Lien, I.C.: Analysis of power output for piezoelectric energy harvesting systems. *Smart Mater. Struct.* **15**(6), 1499 (2006)
4. Kurs, A., Karalis, A., Moffatt, R., Joannopoulos, J.D., Fisher, P., Soljacic, M.: Wireless power transfer via strongly coupled magnetic resonances. *Science* **317**(5834), 83–86 (2007)
5. Shi, Y., Xie, L., Hou, Y.T., Sherali, H.D.: On renewable sensor networks with wireless energy transfer. In: Proceedings of IEEE INFOCOM 2011, pp. 1350–1358, Shanghai (2011)
6. Liang, W., Xu, W., Ren, X.: Maintaining sensor networks perpetually via wireless recharging mobile vehicles. In: IEEE Conference on Local Computer Network. IEEE (2014)
7. Zhao, M., Li, J., Yang, Y.: A framework of joint mobile energy replenishment and data gathering in wireless rechargeable sensor networks. *IEEE Trans. Mobile Comput.* **13**(12), 2689–2705 (2014)
8. He, L., Kong, L., Gu, Y., Pan, J., Zhu, T.: Evaluating the on-demand mobile charging in wireless sensor networks. *IEEE Trans. Mobile Comput.* **14**(9), 1861–1875 (2015)
9. Feng, Y., Guo, L., Fu, X., Liu, N.: Efficient mobile energy replenishment scheme based on hybrid mode for wireless rechargeable sensor networks. *IEEE Sens. J.* **19**(21), 10131–10143 (2019)
10. Xu, W., Liang, W., Lin, X., Mao, G.: Efficient scheduling of multiple mobile chargers for wireless sensor networks. *IEEE Trans. Veh. Technol.* **65**(9), 7670–7683 (2016)
11. Mo, L., Kritikakou, A., He, S.: Energy-aware multiple mobile chargers coordination for wireless rechargeable sensor networks. *IEEE Internet Things J.* **6**(5), 8202–8214 (2019)
12. Dai, H., Wu, X., Chen, G.: Minimizing the number of mobile chargers for large-scale wireless rechargeable sensor networks. *Comput. Commun.* **46**, 54–65 (2014)
13. Zhang, S., Wu, J., Lu, S.: Collaborative mobile charging. *IEEE Trans. Comput.* **64**(3), 654–667 (2015)
14. Han, G., Li, Z., Jiang, J., Shu, L., Zhang, W.: MCRA: a multi-charger cooperation recharging algorithm based on area division for WSNs. *IEEE Access* **5**, 15380–15389 (2017)



Model-Based Battery SOC Estimation Based on GA-UKF Algorithm

Zhijie Zhong and Junyong Zhai^(✉)

School of Automation, Southeast University,
Nanjing 210096, Jiangsu, People's Republic of China
jyzhai@seu.edu.cn

Abstract. A model-based State-of-Charge (SOC) estimation approach of Lithium-Ion battery is put forward based on parameter identification using Genetic Algorithm (GA) and state prediction using an Unscented Kalman Filter (UKF) in this paper. Firstly, the second-order equivalent circuit model (ECM) is established to characterize the performance of battery. Secondly, model parameters are identified through parameter identification based on GA. Then, the UKF algorithm is applied to predict SOC. Finally, The experimental results show that the proposed GA-UKF algorithm can predict the battery SOC accurately.

Keywords: SOC estimation · Genetical algorithm · Unscented Kalman Filter

1 Introduction

Recently, Lithium-Ion batteries have gradually become the research focus in the field of energy storage due to its excellent performance [1]. However, The batteries show obvious aging phenomenon with service time increasing, leading to the inaccurate state estimation such as Remaining-Useful-Lifetime (RUL) and SOC [2]. For the sake of safe operation, an accurate state estimation method is required to efficiently manage the battery pack.

Among all the states of Lithium-ion battery, SOC is the most significant one. The SOC is defined as the following formula:

$$SOC = \frac{Q_r}{Q_n} \times 100\% = (1 - \frac{Q_f}{Q_n}) \times 100\% \quad (1)$$

where Q_r indicates the current remaining capacity of the battery; Q_n denotes the rated capacity of the battery, and Q_f indicates the current consumed capacity of the battery.

It is difficult to measure the SOC of Lithium-ion battery directly through an electronic sensor [3]. Therefore, the estimation of SOC is implemented with the assistance of other measurable variables like terminal voltage and current. However, precise estimation of SOC is a challenging task because of the complex

internal chemical dynamics of Lithium-ion battery. The common SOC estimation consists of the ampere-hour integration method, the open circuit voltage method and the model-based method [4,5]. Among all the available methods, the model-based method is the most suitable approach for estimation.

The overall process of model-based method include battery modelling and algorithm implementation. The battery modeling research can be separated into electrochemical model, equivalent circuit model and black box model [6–9]. Parameters of selected models can be offline identified through different kinds of algorithms, such as GA [10]. As to algorithm implementation, adaptive filter algorithm represented by Kalman Filter (KF) family algorithm has been put into widespread research. The fundamental principle of KF family algorithm is to recursively estimate the state of current time step based on the state estimated previously in the last time step and the current measurable variables [11–15], providing accurate SOC estimates for the model.

The rest of the paper is listed as follows: In Sect. 2, the ECM is established. Also, the U_{ocv} -SOC relationship curve is calibrated on the basis of constant current charging and discharging experiment, and parameter identification is thus committed using GA. In Sect. 3, UKF algorithm is utilized to perform accurate SOC estimation. In Sect. 4, the Main Discharge Cycle Condition (MDCC) is applied for experimental verification of the proposed GA-UKF algorithm. The estimation results are shown to validate the effectiveness. Finally, the conclusion is given in Sect. 5.

2 Lithium-Ion Battery Modelling

A superb battery modelling consists of accurate model structure and precise cell model parameters. The second order ECM is adopted as model structure in this paper. In order to identify model parameters accurately, the relationship curve of U_{ocv} -SOC is firstly calibrated and GA is then adopted to find optimal model parameters.

2.1 Model Description

The ECM has the advantages of low computational complexity and high precision. In addition, it can partially describe the internal reaction process of the battery cell. The structure of the second-order ECM is shown in Fig. 1.

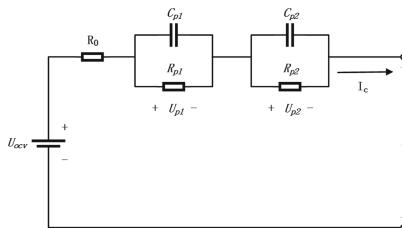


Fig. 1. Schematic diagram of second-order ECM.

The physical meaning of each variable are listed as follows:

The current, I_C : The input of the battery cell model.

The open circuit voltage, U_{ocv} : It's been proved that U_{ocv} has an one-to-one correspondence with SOC. The numerical relationship is usually performed based on the U_{ocv} -SOC curve, which is offline calibrated by constant current discharging and charging experiment.

The parameters of the battery cell: Ohmic internal resistance R_0 , polarization resistance R_{p1} , polarization capacity C_{p1} , diffusion resistance R_{p2} and diffusion capacity C_{p2} . R_{p1} and C_{p1} are used to describe the diffusion phenomenon, having a larger time constant. R_{p2} and C_{p2} are used to describe the polarization phenomenon, having a smaller time constant. The parallel voltages U_{p1} and U_{p2} are the polarization voltages of the corresponding RC links.

The sampling time period, t : The sampling time is usually set as 1 s.

2.2 Calibration of U_{ocv} -SOC

A Lithium-Ion battery with a rated capacity of 33.271 Ah is selected to conduct the constant current charging and discharging experiment. It is discharged at a constant current of 1/3C for 18 min at 25, and then let stand for 1 h, and cycled 10 rounds until the SOC is zero. The data of open circuit voltage U_{ocv} , which is equal to the terminal voltage U after each cycle, is then recorded. Sixth degree polynomial fitting is performed using the fitting tool of MATLAB after total 11 cycles so that the relationship curve of U_{ocv} -SOC is calibrated as

$$U_{ocv} = 11.0784 * SOC^6 - 25.2760 * SOC^5 + 17.5379 * SOC^4 - 1.1590 * SOC^3 - 2.3857 * SOC^2 + 1.2627 * SOC + 3.4222 \quad (2)$$

2.3 Parameter Identification Based on GA

Identifying the model parameters, including R_0 , R_{p1} , C_{p1} , R_{p2} and C_{p2} , offline accurately is another significant task of battery modelling. As GA evaluates large quantity of solutions in the search space, reducing the risk of getting caught in a local optimal solution, it is adopted in this paper.

As is shown in Fig. 2, GA is a stochastic algorithm for finding the optimal solution. The most significant feature of GA is being heuristic: it does not have a rigorous theoretical justification [8], but is justified practically, i.e., it should give a practical suitable result. GA refers to the principles and terminological apparatus of genetics. The search for the optimal or suboptimal solution is executed in the process of individual population evolution. And sequential conversion is carried out based on genetic operators of reproduction, crossover and mutation from one finite set of solutions to another.

The specific steps are listed as follows:

- Utilize a random binary sequence to generate the initial population, a total of 1000 groups, which is, the initial value of the parameters R_0 , R_{p1} , C_{p1} , R_{p2} , C_{p2} . After that, set the probability of mutation and hybridization to 0.05

and 0.5 respectively and select the maximum iteration number as 10. And then, encode the initial value and convert it into a corresponding binary code, which represents the individual's gene.

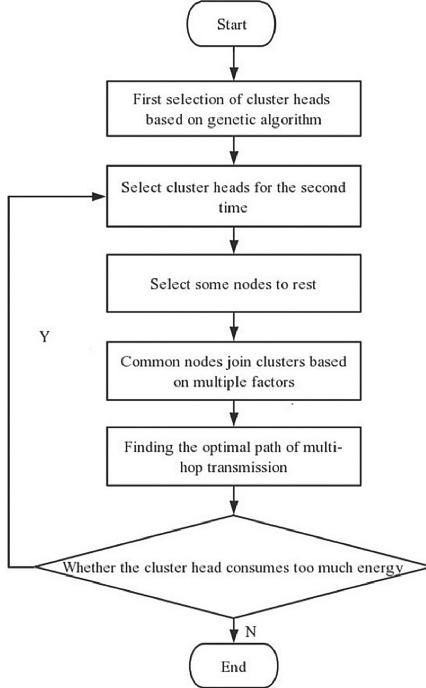


Fig. 2. Schematic diagram of GA.

2. Cross the current population, and exchange part of the gene fragments. The new individuals produced after the exchange are the offspring. The mutation behavior during the simulation of biological evolution is randomly generated in the offspring genes.

3. Decode the crossed and mutated offspring to obtain the offspring values.
 4. Take the state quantity as $x_k = [u_{1k} \ u_{2k} \ SOC_k]^T$, the input quantity as current i_k , and the output quantity as terminal voltage U . Establish state equation and output equation as Eq. (3) and (4):

$$x_{k+1} = \begin{bmatrix} 1 - \frac{T_s}{R_{p1}C_{p1}} & 0 & 0 \\ 0 & 1 - \frac{T_s}{R_{p2}C_{p2}} & 0 \\ 0 & 0 & 0 \end{bmatrix} x_k + \begin{bmatrix} \frac{T_s}{C_{p1}} \\ \frac{T_s}{C_{p2}} \\ -\frac{T_s}{Q_n} \end{bmatrix} i_k \quad (3)$$

$$U_k = [-1 \ -1 \ 0] x_k + U - I \times R_0 \quad (4)$$

5. The fitness function is selected as the RMSE between the actual terminal voltage achieved by the experiment and the estimated terminal voltage value achieved by the algorithm. The fitness function of the k th child is

$$V_- J_k = \sum_{i=1}^{14000} (U_{real} - U_{model})^2 \quad (5)$$

6. Determine whether the individual meets the set requirements through the fitness function (the minimum fitness function of 1000 individuals in the population $V_- J_{min} \leq 0.008$). If there is an individual that meets the set condition, it is chosen as the parameter for model identification; if none of the individuals meets the condition, individuals that have a greater probability of being retained are selected to generate new populations. Repeat steps (2)–(5) to simulate the next round of evolution until the individual meets the condition of adaptability or iteration exceeds the set number. The number of k th children in the next generation is shown as:

$$n_k = \lfloor \frac{V_- J_k}{\sum_{k=1}^{1000} V_- J_k} \times 1000 \rfloor \quad (6)$$

where $\lfloor m \rfloor$ means the integer not greater than m .

3 SOC Estimation Based on UKF

In this section, UKF as a representative in KF family algorithm is applied to improving the accuracy of estimation on the basis of parameter identification. The UKF does not require the derivatives of the state and measurement functions, making it convenient to merge the battery model into UKF to estimate SOC.

The core of UKF is to simulate the real battery system in Eq. (2) and (3) through unscented transformation. Unscented transformation is a numerical sampling technique, which has a minimal set of sigma points to estimate the mean and variance of state variables under nonlinear transformation. The sigma points are then projected through the system state equation and output equation in order to produce new sigma points. The newly estimated mean and covariance are formed based on their statistics.

The specific calculation process is implemented as follows:

1. Determine the initial value of the state value x_0 and the initial value of the posterior state error covariance P_0 on the basis of the battery system shown in Eq. (2)–(3).
2. Calculate sigma points.

$$\begin{cases} x_k^0 = \hat{x}_k \\ x_k^i = x_k + \sqrt{(L + \lambda) P_{k-1}}, i = 1, \dots, L \\ x_k^i = x_k - \sqrt{(L + \lambda) P_{k-1}}, i = L + 1, L + 2, \dots, 2L \end{cases} \quad (7)$$

where L is the length of the state vector. In this paper, the length of the state vector is 3, and the weight value is calculated as follows:

$$\begin{cases} \lambda = \alpha^2 L - L \\ W_m^0 = \frac{\lambda}{L+\lambda}, W_m^i = \frac{1}{2(L+\lambda)}, i = 1, 2, \dots, 2L \\ W_c^0 = \frac{\lambda}{L+\lambda} + 1 - \alpha^2 + \beta, W_c^i = \frac{1}{2(L+\lambda)}, i = 1, 2, \dots, 2L \end{cases} \quad (8)$$

where $\alpha = 0.01$, $\beta = 2$.

3. Update the priori state \bar{x}_{k+1} and the prediction value of variance P_{xx} as

$$\bar{x}_{k+1} = \sum_{i=0}^{2L} W_m^i x_k^i \quad (9)$$

$$P_{xx} = \sum_{i=0}^{2L} W_c^i (x_k^i - \bar{x}_{k+1})(x_k^i - \bar{x}_{k+1})^T + Q_k \quad (10)$$

where Q_k is the system noise covariance matrix.

4. Update the observed value \hat{y}_{k+1} and the predicted value of the observed variance P_{yy} as

$$\hat{y}_{k+1} = \sum_{i=0}^{2L} W_m^i y_k^i \quad (11)$$

$$P_{yy} = \sum_{i=0}^{2L} (W_c^i (y_k^i - \hat{y}_{k+1})(y_k^i - \hat{y}_{k+1})^T) + R_k \quad (12)$$

where R_k is the observation noise covariance matrix.

5. Update covariance P_{xy} , posterior state value \hat{x}_{k+1} and posterior state error covariance P_k as

$$P_{xy} = \sum_{i=0}^{2L} W_c^i (x_k^i - \bar{x}_{k+1})(y_k^i - \hat{y}_{k+1})^T \quad (13)$$

$$K_k = \frac{P_{xy}}{P_{yy}} \quad (14)$$

$$\hat{x}_{k+1} = \bar{x}_{k+1} + K_k (y_{k+1} - \hat{y}_{k+1}) \quad (15)$$

$$P_k = P_{xx} - K_k P_{xy}^T \quad (16)$$

4 Estimation Results

The working condition of parameter identification is selected as constant voltage discharging condition. Set the number of iterations to 11 and perform parameter identification based on GA. The voltage and error are shown in Fig. 3 and 4 respectively.

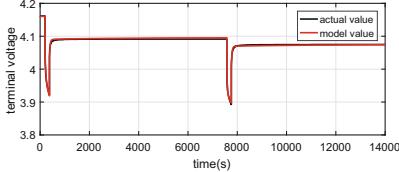


Fig. 3. The voltage.

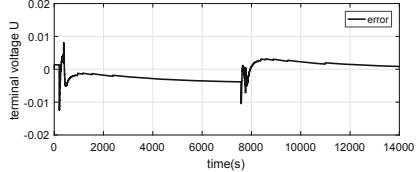


Fig. 4. The error of voltage.

The result of parameter identification is shown in Table 1. Furthermore, the time consumed in the iterations is shown in Table 2. Since the algorithm eliminates elites inferior children, the calculation cost decreases with the number of iterations.

Table 1. The result of parameter identification

R_0/Ω	R_{p1}/Ω	C_{p1}/F	R_{p2}/Ω	C_{p2}/F
0.0039	0.0019	1.491e+04	0.0028	1.102e+06

Table 2. Time consumed in parameter identification

Iteration	Time	Iteration	Time
1	55.688390	7	41.356170
2	55.569559	8	38.600193
3	55.569910	9	34.864351
4	48.540558	10	33.565022
5	55.433895	11	32.237357
6	44.120852		

Based on parameter identification, UKF is applied to estimating the SOC of batteries. The MDCC is adopted to demonstrate the estimation result. The sampling time is set to 0.1 s. The period of the whole working condition is 31201.8 s. The initial SOC of the battery is randomly selected as 0.8. The estimating results are shown as Fig. 5, 6, 7, 8.

Figure 5 illustrates the curve of the actual terminal voltage and the estimated terminal voltage. And Fig. 6 illustrates the error between the actual value and estimated value. The error during the whole MDCC process is within 0.05 V. Figure 7 shows the curve of the actual SOC and the estimated SOC. The estimated SOC approaches the true SOC quickly even if there is a large error in the beginning period. Figure 8 demonstrates the error of SOC. At the beginning of the period, the error reaches approximately 20%. Within only 1000 s, it drops to only 5%. And it remains stable (no more than 2%) after about 6000 s.

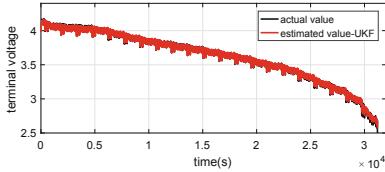


Fig. 5. The voltage of MDCC

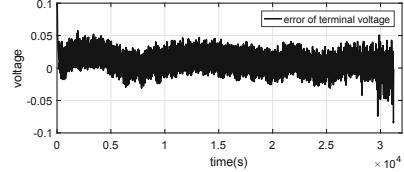


Fig. 6. The error of terminal voltage

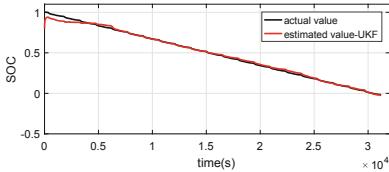


Fig. 7. The actual and estimated SOC

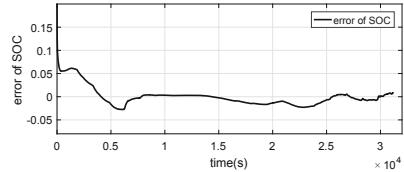


Fig. 8. The error of SOC

5 Conclusion

In this paper, a method combining GA with UKF is put forward to estimate the SOC of Lithium-ion batteries. GA finds the optimal solution of battery parameters through global search. UKF is carried out to improve the accuracy of SOC estimation. The MDCC is applied to verifying the estimation result of the proposed GA-UKF algorithm. The experimental result showed that the proposed method provided good SOC estimations within 2% even if there was an initial SOC error of approximately 20%.

However, the time of parameter identification is relatively long due to the calculation of invalid terms when applying genetic algorithm to parameter identification. It is necessary to optimize GA to reduce calculation cost and calculation time in subsequent studies. Furthermore, the data-driven model combined with deep learning method such as convolutional neural network is worth studying as it avoids the complicated feature extraction process.

References

1. Yang, C., Wang, X., Fang, X.Q.: An online SOC and capacity estimation method for aged lithium-ion battery pack considering cell inconsistency. *J. Energy Storage* **29**, 101250 (2020)
2. Barre, A., Deguilhem, B., Grolleau, S.: A review on lithium-ion battery ageing mechanisms and estimations for automotive applications. *J. Power Sources* **241**, 680–689 (2013)
3. Doyle, M., Fuller, T.F., Newman, J.S.: Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell. *J. Electrochem.* **140**(6), 1526–1533 (1993)
4. Hu, Y., Yurkovich, S., Guezennec, Y.: Electro-thermal battery model identification for automotive applications. *J. Power Sources* **196**(1), 449–457 (2011)
5. He, H., Xiong, R., Guo, H.: Online estimation of model parameters and state-of-charge of LiFePO₄ batteries in electric vehicles. *Appl. Energy* **89**(1), 413–420 (2012)
6. Blanco, C., Snchez, L., Gonzalez, M.: An equivalent circuit model with variable effective capacity for batteries. *IEEE Trans. Veh. Technol.* **63**(8), 3592–3599 (2014)
7. Luo, J., Peng, J., He, H.: Lithium-ion battery SOC estimation study based on Cubature Kalman filter. *Energy Procedia* **158**, 3421–3426 (2019)
8. Yazdanpour, M., Taheri, P., Mansouri, A.: A circuit-based approach for electrothermal modeling of lithium-ion batteries. In: Thermal Measurement, Modeling and Management Symposium (2016)
9. Guo, Y., Zhao, Z., Huang, L.: SoC estimation of lithium battery based on improved BP neural network. *Energy Procedia* **105**, 4153–4158 (2017)
10. Chen, Z., Mi, C., Fu, Y.: Online battery state of health estimation based on Genetic Algorithm for electric and hybrid vehicle applications. *J. Power Sour.* **240**(1), 184–192 (2013)
11. Shrivastava, P., Soon, T.K., Idris, M.Y.I.B., Mekhilef, S.: Overview of model-based online state-of-charge estimation using Kalman filter family for lithium-ion batteries. *Renew. Sustain. Energy Rev.* **113**, 109233 (2019)
12. Irsyad, N.H., Rize, H.S., Deddy, K.: State of Charge (SoC) estimation of LiFePO₄ battery module using support vector regression. In: Joint International Conference on Electric Vehicular Technology and Industrial, Mechanical, Electrical and Chemical Engineering (ICEVT and IMECE). IEEE (2015)
13. Plett, G.L.: Reviews and some perspectives on different methods to estimate state of charge of lithium-ion batteries. *J. Autom. Saf. Energy* **10**(03), 249–272 (2019)
14. Zhang, Y.: Parameter Identification and SOC Estimation of Power Battery for Electric Vehicle. Jilin University, Changchun (2014)
15. Zhang, S.: Research on Modeling and State Estimation of Vehicles LiFePO₄ Batteries. Wuhan University of Technology, Wuhan (2015)
16. Yang, F., Zhang, S., Li, W., Miao, Q.: State-of-charge estimation of lithium-ion batteries using LSTM and UKF. *Energy* **201**, 117664 (2020)



Study on the Vulnerability of Power Grid Cascade Failures Based on Complex Network Theory

Yirong Long and Chaoyang Chen()

School of Information and Electrical Engineering,
Hunan University of Science and Technology, Xiangtan, China
chaoyangg@gmail.com

Abstract. Reliable, safe, efficient and environmental-friendly power supply ensures the favourable operation of the whole society. With the continuous increase of electricity consumption and with the sustainable development of distributed generation, comprehensive analysis of vulnerability of power grids has become urgent. This paper, based on complex network theory, studies failures on the transmission lines, which little research had been done on before, other than nodes. This paper proposes a new vulnerability assessment algorithm, considering two types of different stress that the power grid undergoes: the increase of electricity consumption and the fluctuations caused by distributed generation. This paper simulates IEEE 30-bus system and IEEE 118-bus system to prove this assessment algorithm efficient, and is different from typical research method using network attacks. It also shows that power grid vulnerability has negative correlation with network average degree. This paper shed new light on the vulnerability analysis of power grid based on complex network theory.

Keywords: Complex network · Power grid · Cascade failure · Distributed energy resources · Network attacks

1 Introduction

Modern life cannot be separated from electricity. For one thing, the development of world's electricity consumption and power generation are both at a high speed. National life, social development, power grid expansion and distributed generation are all over the world. Especially with the rapid development of electric vehicles, electrification in the transportation field will be accelerated in an all-round way. The world's electricity consumption and power generation are developing at a high speed. To meet the soaring electricity demands of social life and development, power grid expansion and distributed generation projects are being constructed all over the world. But the working output of distributed generation systems are strongly affected by environment, and its stability and reliability are low. For example, the external environment such as light intensity

and wind speed will surely affect solar power supply and wind power. The function of distributed generation is always uncertain. Nowadays, a large number of distributed energy sources have been connected to the traditional power grids, which changes the structure and operation state of the power transmission and distribution network, causing the power flow in the power system to fluctuate, and even change in multiple directions. The network losses and voltages of power lines are also affected by this randomness, which leads to the difficulty of vulnerability evaluation of the power system. The cascading failures of power grid have relatively short time scales and very bad consequences, deteriorating in the form of avalanche. It is of great importance to comprehensively study the vulnerability analysis of the power distribution network with increasing load and with load fluctuations.

Most research on vulnerability of power grid based on complex network theory usually analyze the structure of the network without considering in details the electrical and physical properties of the power grid, focusing on the intrinsic property of grid topologies and compare them to other types of typical complex networks [1–5]. There are many alternative methods available for solving the assessment of the classical properties. And many research study the behavior of the network under several kinds of network attack strategies, which are always categorized into random attack and targeted attack, and are simulated by removing different categories of nodes in the network, and analyze the topological changes in the network with expanding black-outs and cascading failures [6–10]. A challenging problem which arises in this domain is that the consequences of load increase and uncertainties of the distributed generation in the power grid are rarely examined directly. A new approach is therefore needed for quantify the practical stress on the power grid. The key contribution of this paper is the new vulnerability assessment algorithm it provides. This paper will fill the missing brick in the wall of research already done, and use the insight gained through the analysis to help the power grid planning.

2 Complex Network Theory and Power Grid

Small world theory [1] demonstrates the small world characteristics of the power grid. Many long-distance high-voltage transmission lines in the grid makes the grid have “long-distance connection”. This paper applies the following principles to simplify IEEE 30-bus system and IEEE 118-bus system to topology models including N nodes and M edges as shown in Fig. 1 and Fig. 2: Ignore the parallel lines and the branches linked to ground. Transformer lines and transmission lines are regarded as edges. The generators, the load, the end of the transformers serve as nodes. Transmission lines or transformer branches are taken as edges. This paper defines the topologies are undirected and weighted by reactance, and suggests that the power flow is transported along the lines with the least reactance value from generations to loads.

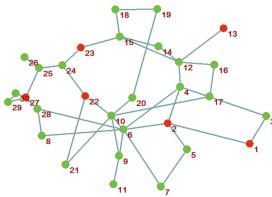


Fig. 1. Topology of the IEEE 30-bus system

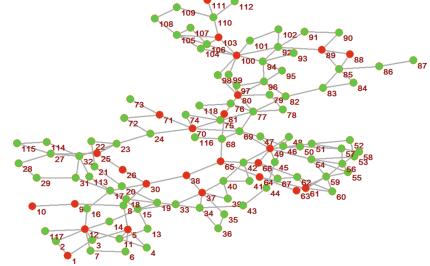


Fig. 2. Topology of the IEEE 118-bus system

3 Vulnerability Analysis of Power Grid Under Load Increase and Fluctuation Based on Complex Network Theory

3.1 Cascade Model for Vulnerability Analysis

DC Power Flow Model. There's a wide variety of load redistribution mechanisms in dynamic vulnerability analysis based on complex networks. Most cascade models for power grids are purely topological models based on the local redistribution of power loads upon failure and disregard the long-range nature of electricity [6,8]. This paper uses DC power flow model [11], defining generation nodes as PU nodes, A special generation node as slack node, load nodes and transformer nodes as PQ nodes.

$$P_l = B_l A \theta \quad (1)$$

where P_l is the vector of branch flow. B_l is branch admittance diagonal matrix. A is the node-arc incidence matrix. θ is the vector of nodal phase angles.

Load-Capacity Model. There's a lot of researches focusing on nodes in dynamic vulnerability analysis based on complex networks [7–9]. But in actual life, there is a higher probability that failures on transmission lines would occur, such as short circuit and disconnection of transmission. This paper supplements this short board by improving node-based Motter-Lai cascade model [12]:

$$C_l = (1 + \omega)L_0 \quad (2)$$

where C_l are the capacity of transmission lines. ω is the tolerance parameter of transmission lines, which is defined by $\omega = 0.1$ in this paper, considering the capacity is strictly limited by cost in the artificial network. And L_0 are the initial power flow of the lines.

3.2 Vulnerability Analysis Model Based on Load Increase and Load Fluctuations

The overall network under stress cascade model is shown in Fig. 3. This paper studies the mechanism that the load and the fluctuation carried by each transmission line in the power grid increases until the line fails. According to the initial power flows and sources (simulating the two kinds of stress imposed to the power grid), initial power flows are calculated. If the load on a line goes beyond its capacity, the line trips and power flows are recalculated on the new topology. Such procedure is repeated until the network is stable.

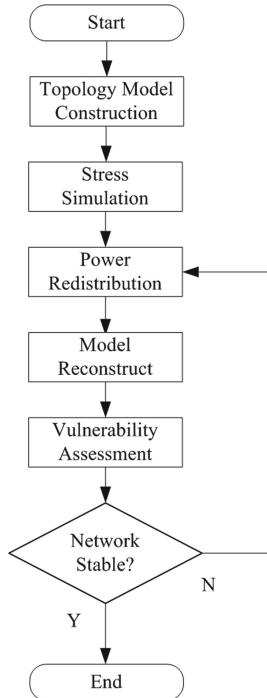


Fig. 3. Power grid under stress cascade model

Load Increase. The load increase is quantified by uniformly increase L_l in transmission load:

$$L_l(\alpha) = (1 + k'_l \alpha) P_l^0. \quad (3)$$

where P_l^0 is the initial value of line load calculated by DC power flow model. α parameters the load increase of the line. When it is equal to 0, this transmission line does not bear any load increase. k'_l is the product of the ratio between the degree of nodes at both ends of the line and the average degree of the network $\langle k \rangle = \frac{2M}{N}$. k'_l quantifies the influence of the degrees of the nodes connected to

the line on the increase of the load borne by the line in the power grid and has a positive correlation.

$$k'_l = \frac{k_j k_k}{\langle k \rangle^2} \quad (4)$$

where k_j and k_k are the degrees of node j and k , which are the two sides of the line l , and are equal to the numbers of edges those two sides of the edge respectively linked to.

Load Fluctuations. The energy fluctuation is quantified as the random increase of load:

$$L_l(\sigma) = (1 + k'_l \sigma \xi_l) P_l^0 \quad (5)$$

where ξ_l is a random variable distributed in $[-1, 1]$, mimicking power fluctuation of the power grid. σ parameters the strength of the load fluctuation of the line in the power grid. For the above formula (5),

$$L_l(\sigma) \in [(1 - k'_l \sigma) P_l, (1 + k'_l \sigma) P_l].$$

Vulnerability Assessment Index. This paper uses the line loss rate as the fault index, reflecting the structural vulnerability of the power grid.

$$f = \frac{m_t}{M} \quad (6)$$

where M is the total number of all edges in the network. m_t is the total number of currently failed (overloaded) edges.

3.3 Results

The results on the IEEE 30-bus system and IEEE 118-bus system verify the effectiveness of the proposed model. The results of the two formulas are shown below.

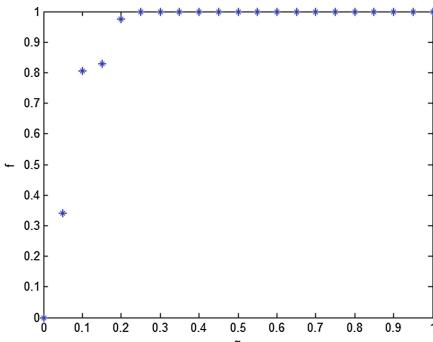


Fig. 4. IEEE 30 bus system under load increase

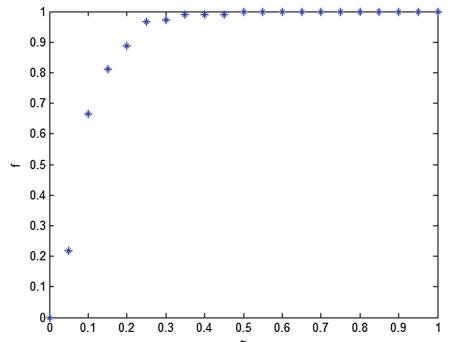


Fig. 5. IEEE 118 bus system based under load increase

Load Increase. As shown in Fig. 6 and Fig. 5, both networks deteriorates as α increases, which proves the load increase aggravate the vulnerability of the network. The collapse of the networks is very abrupt, and it is respectively about 0.25 and 0.5 when the IEEE 30-bus system and IEEE 118-bus system collapse. In view of the average degree of IEEE 30-bus system and 118-bus system are 2.73 and 3.04. It is deduced that the higher the $\langle k \rangle$ is, the less vulnerable the network under load increase is.

Load Fluctuations. As shown in Fig. 6 and Fig. 7, both of the networks deteriorates as σ increases, which proves the fluctuations increase aggravate the vulnerability of the network. Nearly collapsed in the end, but there were still a few isolated islands compared to the load increase situation, there are much smoother transitions in the load fluctuation case; nevertheless, the transition becomes steeper with small $\langle k \rangle$, indicating the possibility of $\langle k \rangle$ strengthening the robustness of the network under load fluctuations. Also, in this case the breakdown due to load fluctuations is less severe since a finite fraction of power lines survives and the system is in an islanded state, compared with the case that the network is under load increase.

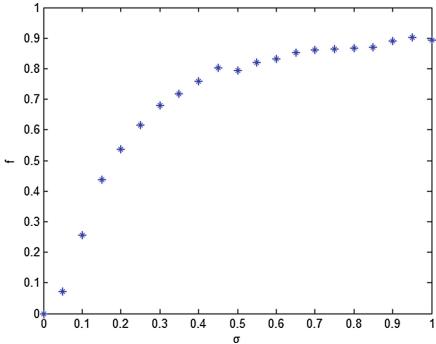


Fig. 6. IEEE 30-bus system under load fluctuation

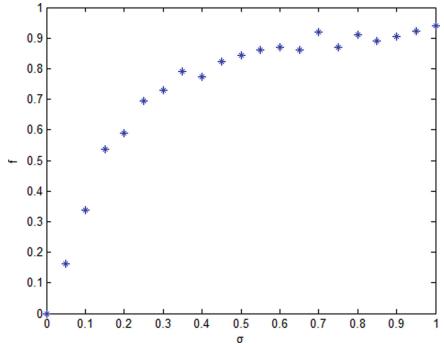


Fig. 7. IEEE 118-bus system under load fluctuation

4 Vulnerability Analysis of Network Under Attack Based on Complex Network Theory

This section simulates random attack and targeted attack, focusing on several security issues of N-k contingency and vulnerability assessment to compare with the algorithm proposed before.

4.1 Vulnerable Lines of Power Grid

This paper proposes the line weighted betweenness index to reflect the vulnerability of lines, and uses Floyd-Warshall algorithm [13] to obtain the shortest path of all generation-load nodes pairs. The minimum values of the sum among them are the final intermediate value of the shortest path. It is believed that the higher the weighted intermediate value of the line, the more vulnerable it is.

4.2 Discussion

The abscissa indicates the percentage number of lines in each attack network, i.e. 5% of lines in the network are removed for the first time, 10% are removed for the second time, and so on. The ordinate indicates the change in network efficiency after the failure. Figure 8 and Fig. 9 show the IEEE-30 bus system and IEEE 118-bus system under network attacks. The overall tendency of the targeted attack is much sharper than random attack. And the abruptness of IEEE 118-bus system is much obvious than that of IEEE 30-bus system, indicating the larger scale of network is more sensitive to incipient failures. Nevertheless, it is after the branches have been respectively removed by 35% and 40% when the network collapsed. Considering that the $\langle k \rangle$ of IEEE 30-bus system and 118-bus system are respectively 2.73 and 3.04. It can be assumed that $\langle k \rangle$ has positive correlation with network robustness, strengthening the robustness of the network under attacks.

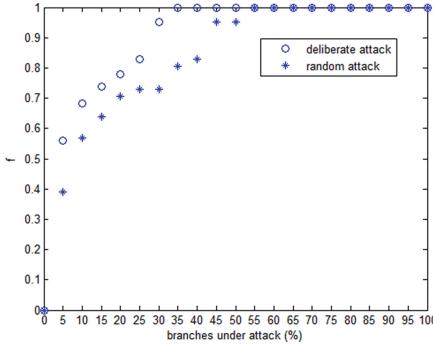


Fig. 8. IEEE 30 bus system under network attacks

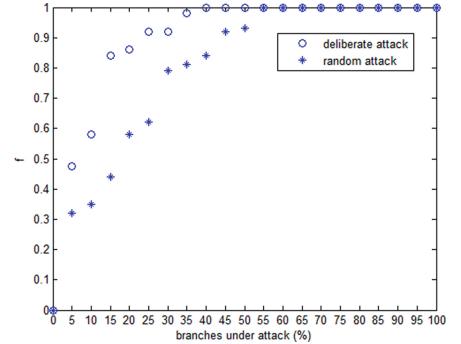


Fig. 9. IEEE 118 bus system under network attacks

5 Summary

This paper vividly demonstrates the vulnerability of power grid cascades under different types of stress and attacks. The new vulnerability assessment algorithm comprehensively compensates for the lack of research on vulnerability of power

grid based on complex network theory, by focusing on lines and the stress from load increase and distributed generations. At the same time, this paper suggests in the related results that the vulnerability of the network is decreased with increasing the average degree of the system by comparing the results of IEEE 30-bus system and 118-bus system. This paper put forward a guidance in the grid planning of upgrading the current network infrastructure towards future topologies optimized for energy exchange in the super grid.

In the future, the power grid will certainly be a blend of two modes: one is a large interconnected power grid, eliminating barriers between regions and energy markets; The second is a small self-contained distribution network including distributed power supply and energy storage, which forms an active distribution network through intelligent control. Only by combining these two modes can an efficient and economical environmental-friendly power system be created. The impact on power system caused by the load increase of power grid and the interconnection of distributed generations shouldnt be underestimated and it calls for more in-depth research. In general, changing the network topology can improve the power grid's response to stress and attacks. After that, specific planning research can be carried out in combination with economic factors and factors with the best robustness of the power grid, considering the average degree shown in [14, 15] is about 4. Using complex network theory to study and adjust the parameters of the power grid can enhance the robustness of the power grid and reduce the impact of such instantaneous changes. The vulnerability analysis of power grid shown in this paper provides an intuitive risk reference system for planners of power grid expansion projects.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China under Grants (61973110, 61503133), the Province Natural Science Foundation of Hunan under Grant (2016JJ6043).

References

1. Watts, D.J., Strogatz, S.H.: Collective dynamics of small world networks. *Nature* **393**, 440–442 (1998). <https://doi.org/10.1038/30918>
2. Barabasi, A.L., Albert, R.: Emergence of scaling in random networks. *Science* **286**, 509–512 (1999). <https://doi.org/10.1126/science.286.5439.509>
3. Watts, D.J.: *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton University Press, Princeton (2003)
4. Scala, A., Lucentini, P.G.D.S.: The equal load-sharing model of cascade failures in power grids. *Phys. A Stat. Mech. Appl.* **462**, 737–742 (2015). <https://doi.org/10.1103/PhysRevE.77.026102>
5. Das, H., Jena, A.K., Rath, P.K., Muduli, B., Das, S.R.: Grid computing-based performance analysis of power system: a graph theoretic approach. In: Intelligent Computing, Communication and Devices, pp. 259–266. Springer, New Delhi (2015). https://doi.org/10.1007/978-81-322-2009-1_30
6. Sole, R.V., Rosas-Casals, M., Corominas-Murtra, B., Valverde, S.: Robustness of the European power grids under intentional attack. *Phys. Rev. E* **77**, 026102 (2008). <https://doi.org/10.1103/PhysRevE.77.026102>

7. Chen, Z., Wu, J., Xia, Y., Zhang, X.: Robustness of interdependent power grids and communication networks: a complex network perspective. *IEEE Trans. Circ. Syst. II: Express Briefs* **65**, 115–119 (2018). <https://doi.org/10.1109/TCSII.2017.2705758>
8. Nesti, T., Zocca, A., Zwart, B.: Emergent failures and cascades in power grids: a statistical physics perspective. *Phys. Rev. Lett.* **120**, 258301 (2018). <https://doi.org/10.1103/PhysRevLett.120.258301>
9. Han, Y., Guo, C., Shiying, M.A., Song, D.: Modeling cascading failures and mitigation strategies in PMU based cyber-physical power systems. *J. Mod. Power Syst. Clean* **6**, 944–957 (2018). <https://doi.org/10.1007/s40565-018-0407-3>
10. Slednev, V., Bertsch, V., Ruppert, M., Fichtner, W.: Highly resolved optimal renewable allocation planning in power systems under consideration of dynamic grid topology. *J. Mod. Power Syst. Clean* **96**, 281–293 (2018). <https://doi.org/10.1016/j.cor.2017.12.008>
11. Glover, J.D., Sarma, M.S.: Power system analysis and design: with personal computer applications. *Int. J. Circ. Theor. App.* **7**, 277–288 (1994). <https://doi.org/10.1002/cta.4490070302>
12. Motter, A.E., Lai, Y.: Cascade-based attacks on complex networks. *Phys. Rev. E* **66**, 065102 (2002). <https://doi.org/10.1103/PhysRevE.66.065102>
13. Floyd, R.W.: Cascade-based attacks on complex networks. *Commun. ACM* **5**, 345 (1962). <https://doi.org/10.1145/367766.368168>
14. Pagani, G.A., Aiello, M.: Power grid network evolutions for local energy trading. arXiv preprint [arXiv:1201.0962](https://arxiv.org/abs/1201.0962) (2012)
15. Pagani, G.A., Aiello, M.: Power grid complex network evolutions for the smart grid. *Phys. A Stat. Appl.* **396**, 248–266 (2014). <https://doi.org/10.1016/j.physa.2013.11.022>



Multiple Unmanned Armored Vehicle Formation Transform

Miqi Huang, Daxin Tian, and Xuting Duan^(✉)

Beijing Advanced Innovation Center for Big Data and Brain Computing
School of Transport Science and Engineering, Beihang University, Beijing, China
duanxuting@buaa.edu.cn

Abstract. The formation of armored vehicle formation achievement and transforming are important research contents of vehicle formation control. Especially for the situation of unmanned armored vehicle formation. This paper combines the virtual navigator method and the behavior-based bait-predator method, and applies it to the formation achievement and transforming control of four unmanned armored vehicle formations. The target position is used as a bait to realize flexible transforming between linear, triangular and diamond formations. This method can prevent the formation leader from losing control due to the incapacitation of the formation leader, and does not specify the fixed position of the vehicle in the formation, which greatly guarantees the stability of the formation and the flexibility of formation transforming. The simulation experiment also proves that the method has good effect on the formation and formation transforming of vehicle formation.

Keywords: Unmanned armored vehicle · Formation transform · Virtual leader · Bait -predator

1 Introduction

Under the background of the new military revolution in the world, the cooperative operation mode of armor group composed of multiple armor units is more complex. Among them, the formation, maneuver and transformation scheme of armored vehicle formation is the key topic of the armor group cooperative operation research. In order to give full play to the firepower advantage and mobility advantage of armored forces effectively, the armored vehicle formation must have fast response ability in the battlefield environment, which can reduce the potential battle threat and loss to a minimum as far as possible. In the future high-tech battlefield form, unmanned armored vehicle is the main force to undertake fire strike in ground war, so it is of great significance to study the cooperative operation scheme.

In the research of unmanned armored vehicle queue transforming control method, action selection method can quickly get control strategy by making rules [1] to select action and weight, but it is difficult to quantify and analyze the

action mathematically. The imaginary rigid body [6] method is easy to describe the whole behavior, having the formation feedback mechanism and high control accuracy, whose feedback mechanism and high precision control need high-quality communication and high-power computing. The control method based on directed graph [3] has a great dependence on the stability of information flow. The artificial potential energy method [4] is widely used, but its control strategy has no advantage in the actual battlefield situation.

The virtual leader method [2] used in this paper is an improved pilot following method, which generates a virtual formation member as the navigator, and other vehicles follow the navigator's movement, transforming the formation control problem into following the agent to track the navigator's orientation and The problem of location has the advantages of avoiding the failure of the navigator and making it uncontrollable, and improving the flexibility of formation control. At the same time, combined with the behavior-based decoy-predator method, it can effectively avoid collisions between vehicles and does not specify the fixed position of the vehicles in the formation, improve the formation speed and greatly improve the flexibility of formation transform.

2 Research Contents

In this paper, virtual leader and behavior-based bait predator method are combined to study the formation transforming control of unmanned armored vehicle formation. It is assumed that a formation consists of four vehicles, which are respectively expressed as $R_i(i = 1, 2, 3, 4)$. The target formation is formed and transformed based on virtual leader R_0 under the distributed control architecture, according to the task requirements.

The formation vehicles described in this paper can communicate with each other and have the function of self-positioning; each vehicle can receive command information and communicate with the adjacent vehicles; all vehicles are in stable state; virtual leader is generated by relying on one vehicle, and each target position and its own running speed of the formation can be broadcasted in real time, and all vehicles can be covered; When the virtual leader vehicle is disabled, it can be transformed to other vehicles in the formation to continue to undertake this function, so as to avoid the loss of formation.

3 Virtual Leader and Behavior-Based System

3.1 Virtual Leader System

The leader follower [8] method is to designate an agent in the formation as the leader, and other agents follow the leader's movement. As an optimization of the leader follower method [3], the virtual leader method is to set up a virtual leader, set a virtual structure based on the leader, and the vehicle forms and transforms by tracking the corresponding virtual points. In this paper, a vehicle is selected to generate a virtual leader in the vehicle formation, and this virtual leader is used as the formation benchmark for the formation of the vehicle formation and

the formation of the formation of the target position corresponding to all formation transforming. The formation vehicles can transform the formation by tracking the target position generated by the virtual leader.

3.2 The Behavior-Based System

Each vehicle in the formation is regarded as a predator, and each target position determined by the virtual leader is regarded as a bait [7]. The bait has an attractive effect on the predator, which makes the predator move in the target direction. The predator can obtain the food in the bait position. At this time, the predator does not need to look for the bait, which means the target formation is achieved. The bait can release the pheromone in the environment, of which the concentration will decrease with the increase of distance. The attractiveness of pheromones to predators also decreases with the decrease of pheromone concentration. The same concentration of pheromone has the same attraction to predators. The bait will spread the food within its moving range, and the nibbled bait will no longer attract other predators.

The predator can catch the bait. The predator will move to the direction with the largest pheromone concentration to generate attractiveness, whose direction is from the predator to the bait. The predator starts to eat bait when entering the food range. That predator stays in the center of the food and the target position stands for the state of satiety, where desire for food is always reduced and the bait effect gotten will gradually decrease. At the same time, the bait can no longer satisfy other predators, that is, it cannot produce the attractive effect.

Inferiority force keeping the predator to stay away from the bait is produced when finding there is a closer predator to the bait. The inferiority force can avoid two predators running to the same target at the same time. The damping region is designed to weaken the predator's oscillation near the bait, which can make the predator stay in the center of the bait as soon as possible. When entering the damping region, the predator suffer the resistance whose direction is opposite to that of the predator's velocity relative to the target bait.

Collision avoidance behavior is produced by predators to ensure their own safety. When predators meet with other predators in the process of predation, they need to keep a safe distance to ensure safety and avoid obstacles. When the distance between the two predators is less than the safe distance, they will suffer the opposite collision avoidance force respectively. If a predator has occupied a bait, it is no longer affected by other predators' collision avoidance.

4 The Armored Vehicles Group Model

4.1 The Vehicles Kinematics Model

In this paper, the quadratic integral model is used to describe the vehicle kinematic model. Each vehicle is represented as

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_t = u_i/m \end{cases}$$

where v_i, x_i, u_i, m respectively is the speed, position, control input and quality of vehicle $R_i, v_i, x_i, u_i \in R^2$. the vehicle is considered as a mass point that can detect other vehicles and obstacles within the range of sensors. the state vector of the vehicle is taken as:

$$\dot{x}_i = [x_i \ \dot{x}_i]^T$$

The state equation of vehicle R_i is:

$$\dot{x}_i = \begin{bmatrix} O & I \\ O & O \end{bmatrix} x_i + \begin{bmatrix} O \\ I \end{bmatrix} u_i$$

Where O, I is the corresponding zero matrix and unit matrix of the same dimension.

Since the vehicle has certain speed limit and limited braking capacity, the maximum speed of the vehicle is set as v_{\max} . The design vehicle speed has the following limitations:

$$v_i = \begin{cases} v_{\max} \frac{v_i^*}{\|v_i^*\|}, & \|v_i^*\| > v_{\max} \\ v_i^* = v_{i-\tau} + \int_0^\tau a(t)dt, & \|v_i^*\| \leq v_{\max} \end{cases}$$

4.2 The Formation Description

According to the method of virtual leader, the virtual leader R_0 is designed as the leader of the formation, always serves as the position calibration point of the formation vehicle and guides the movement direction and speed of the formation. $R_i (i = 1, 2, 3, 4)$ represent four vehicles in the formation to help to study the maintain and transform the formation. Set the standard formation matrix of a formation matrix as S , which describes the angle and distance of each position in the formation relative to the virtual leader. The angle and distance of the standard vehicle position relative to the leader is in the second to fifth row of matrix. The first row is the parameter of virtual leader itself, without deviation, so its parameter is constant $[0, 0]$. The first column β_{0i} represents the angle between the position vector $X_{i0} (X_{i0} = X_{R_i} - X_{R0})$ and X-axis. The second column l_{0i} is the deviation distance.

$$S = \begin{bmatrix} \beta_{00} & l_{00} \\ \beta_{01} & l_{01} \\ \beta_{02} & l_{02} \\ \beta_{03} & l_{03} \\ \beta_{04} & l_{04} \end{bmatrix}$$

The formation S_{X0} is represented as

$$S_{X0} = \begin{bmatrix} \beta_{00} + \theta kl_{00} \\ \beta_{01} + \theta kl_{01} \\ \beta_{02} + \theta kl_{02} \\ \beta_{03} + \theta kl_{03} \\ \beta_{04} + \theta kl_{04} \end{bmatrix}$$

Where θ and k are respectively the rotation angle and scaling factor relative to the standard S_i ($\theta \in R, k > 0$), the position of bait (target point) is:

$$\begin{cases} x_{0i} = x_0 + kl_{0i} \sin(\beta_{0i} + \theta) & i = 1, 2, 3, 4 \\ y_{0i} = y_0 + kl_{0i} \cos(\beta_{0i} + \theta) \end{cases}$$

This paper designs three standard formations: straight formation, triangle formation, and diamond formation, where the angle unit is rad and the distance unit is m.

$$S_{LO} = \begin{bmatrix} 0 & 0 \\ 0 & 3 \\ 0 & 6 \\ \pi & 2 \\ \pi & 6 \end{bmatrix}, S_{TO} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{4\sqrt{3}}{3} \\ \pi & \frac{4\sqrt{3}}{3} \\ \frac{2\pi}{3} & \frac{2\sqrt{3}}{3} \\ -\frac{2\pi}{3} & \frac{4\sqrt{3}}{3} \end{bmatrix}, S_{HO} = \begin{bmatrix} 0 & 0 \\ 0 & 4 \\ \frac{\pi}{2} & 2 \\ -\frac{\pi}{2} & 2 \\ \pi & 4 \end{bmatrix}$$

5 The Protocol of Control System

5.1 The Input of Control System

The input of control system is designed as u_i according to the formula of the force and speed. ($F = ma$), includes the attractive force F_{Ti} that attracts the vehicle to the target position, the desire force F_{Di} that makes the vehicle actively move to the target position, the inferiority force F_{Ii} that avoids multiple vehicles taking the same position as the target position, the resistance F_{Ri} that ensures the vehicle reaches the target position at a steady speed, and the collision avoidance force F_{Ai} that avoids collision between vehicles and obstacles:

$$u_i = F_{Ti} + F_{Di} + F_{Ii} + F_{Ri} + F_{Ai}$$

5.2 The Vehicle Control Force

The Attractive Force. The attractive force is affected by the starvation state and the distance between the vehicle and target location. In order to keep the vehicle moving towards the target position, it needs to be attracted even in the far area. a constant small amount is set as the attraction force of the vehicle farther away from the target formation. the attraction of the predator R_i is:

$$F_{Ti} = k_s \max(f(\|s_{ij}\|)) \frac{s_{ij}}{\|s_{ij}\|}$$

Where, s_{ij} is the location vector between the vehicle R_i and bait B_j . ($s_{ij} = X_{R_i} - X_{B_j}$) The concentration of pheromone emitted by the bait is related to the distance. The more the concentration of pheromone is, the greater the attractiveness force is. $f(\|s_{ij}\|)$ is a function of the concentration of pheromone. The diffusion formula of information concentration $\tau_{ij} = \delta q(i) \frac{L \tan \frac{\alpha}{2} - \rho}{L \tan \frac{\alpha}{2}}$ can draw the

attractiveness formula [5], k_s is the hunger degree of the predator. This article takes $f(d)$ as a linear function and k_s as a piecewise function related to distance:

$$f(d) = \begin{cases} k_1 d + c_1 & 0 \leq d < R_e \\ c_0 & d \geq R_e \end{cases}$$

$$k_s = \begin{cases} \frac{d}{R_f} & 0 \leq d < R_f \\ 1 & d \geq R_f \end{cases}$$

R_e is the diffusion radius of pheromone, c_1 is the concentration of pheromone in the location of bait, k_1 is the decay rate of pheromone with distance, c_0 is a small amount of normal value, R_f is the food around the bait Distribution radius. ($c_1 > 0, k_1 < 0, c_0 = k_1 R_e + c_1$)

The Desire Force. When having the tendency to move away from the target position, the predator needs a force to keep the vehicle moving toward the target position. So, it needs a judgment function δ_i to judge when the desire force is generated.

s_i is the position vector of the vehicle R_i relative to the target position, v_i is the current speed vector of the vehicle R_i , $\text{dig}(s_i, v_i)$ is the angle between the vectors s_i and v_i , the direction of the force is as shown in the Fig. 1, when $\text{dig}(s_i, v_i) \in (0, \pi/2)$, F_{Di} is not needed. In other cases, F_{Di} is needed to strengthen the vehicle's tendency to move to the target position. Then, the judgment functions δ_i and F_{Di} are defined as:

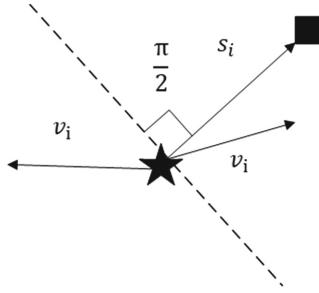


Fig. 1. Schematic of the direction of the force

$$\delta_i = \begin{cases} 0 & \text{dig}(s_i, v_i) \in (0, \frac{\pi}{2}) \\ 1 & \text{other} \end{cases}$$

$$F_{Di} = \sum_{i=1}^n \delta_i k_d F_{Ti}$$

The Inferiority Force. When two vehicles R_1 and R_k determine a certain position as the target position at the same time, it need to select one vehicle to

drive to this target position. At this time, it can be considered that when two predators lock the same target bait at the same time, one state is better than the other, and the other will have inferiority force, that is, must be away from the target position to search for other target positions. In this process, the size of the distance from the target position is used as the criterion for judgment, that is, if $d_{ij} = \|X_{R_j} - X_{B_j}\|$, $d_{kj} = \|X_{R_k} - X_{B_j}\|$, $d_{ij} > d_{kj}$

At this time, the R_k state is better than R_i . Then define the inferiority force generated by the vehicle R_i as

$$F_{Ii} = \begin{cases} \frac{k_i}{d_{kj}} (F_{Ti} + F_{Di}) & d_{kj} < d_{ij} \\ 0 & \text{other} \end{cases}$$

Where k_i is Inferiority factor

The Resistance. When the vehicle is closer to the target position, the state of the vehicle needs to be controlled to be closer to the state of the virtual leader, that is, the vehicle needs to stay at the target position at a steady speed. At this time, when the predator enters the damping zone, it is affected by resistance, and the resistance is related to the speed of the predator relative to the bait. The direction is always opposite to the speed of the vehicle, the resistance is defined as

$$F_{Ri} = \begin{cases} k_r (v_i - v_{Bj}) & 0 \leq d < R_r \\ 0 & d \geq R_r \end{cases}$$

Where k_r is the damping coefficient and R_r is the radius of the damping zone. ($k_r < 0$).

The Collision Avoidance Force. To prevent collisions between vehicles and collisions between vehicles and obstacles, collision avoidance forces are defined. When the distance between vehicles is less than the safe distance, that is, a collision may occur, the collision avoidance force begins to work. The collision avoidance force of the vehicle R_i can be defined as

$$F_{AiK} = \begin{cases} k_\alpha \|v_{R_i}\| \frac{s_{ik}}{\|s_{ik}\|} \cos \theta & \|s_{ik}\| < R_{safe} \\ 0 & \|s_{ik}\| > R_{safe} \end{cases}$$

$$F_{Ai} = \sum_{k \in N(R_i)} F_{Aik}$$

k_α is the collision avoidance coefficient, v_{R_i} is the speed of the vehicle R_i , R_{safe} is the safety radius, that is, the maximum distance at which the collision avoidance force acts, s_{ik} is the position vector between adjacent vehicles, $\theta = \text{dig}(v_{R_i}, s_{ik})$ is the angle between the velocity vector V_{R_i} and the position vector s_{ik} , R_{safe} is determined according to the vehicle's own speed and braking capacity, $N(R_i)$ is the set of all vehicles entering the safety radius of R_i . When the vehicle R_i is already at the target position, the target position will no longer attract other vehicles, and the vehicle will no longer be affected by the collision avoidance force of other vehicles. At this time, R_i is equivalent to a static obstacle, which still affects other vehicles.

6 The Simulation Experiment and the Result Analysis

In this experiment, four vehicles are used to form and transform formations. According to the virtual leader system and the behavior-based bait-predator system model, each vehicle is numbered $R_i(i = 1, 2, 3, 4)$. Select R_i to generate virtual leader R_0 . The virtual leader broadcasts the formation target location to all vehicles in real time. The maximum speed of the designed vehicle in the experiment is $v_{\max} = 8 \text{ m/s}$, $R_f = 8 \text{ m}$, $R_{\text{safe}} = 5 \text{ m}$, $R_e = 70 \text{ m}$, $k_d = 8$, $k_1 = 1.5$, $k_r = 3$. The formations used in the experiment are straight formation and triangle formation, Diamond formation.

6.1 The Formation Forming

In the initial state, all vehicles are in random positions, and it is necessary to quickly form a designated formation in formation under movement. A straight line formation is formed by random positions as shown as Fig. 2:

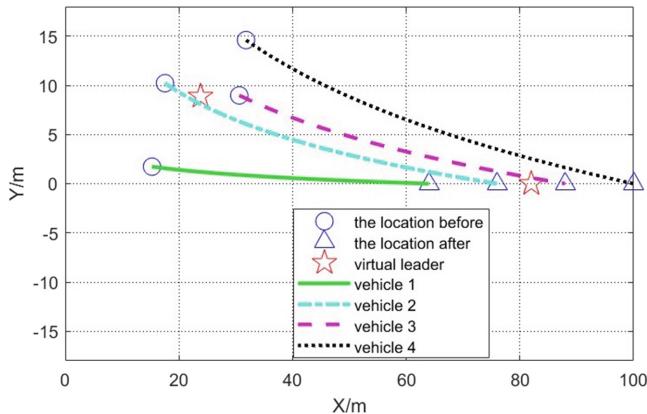


Fig. 2. Straight line formation is formed by random positions

As shown in the figure, in the free-barrier environment, the vehicle position and speed information are randomly generated. During the movement of the virtual leader, the vehicle searches for the target position, and the formation is achieved through the sum of various forces and there is no collision during the whole process.

6.2 The Formation Transforming

Under the overall movement of the formation, the transform between the formations is as shown in the Fig. 3 and Fig. 4. The figures show the vehicles positions before the formation transform and the vehicles positions after the transform. The lines show the vehicle's driving trajectory during the transforming process:

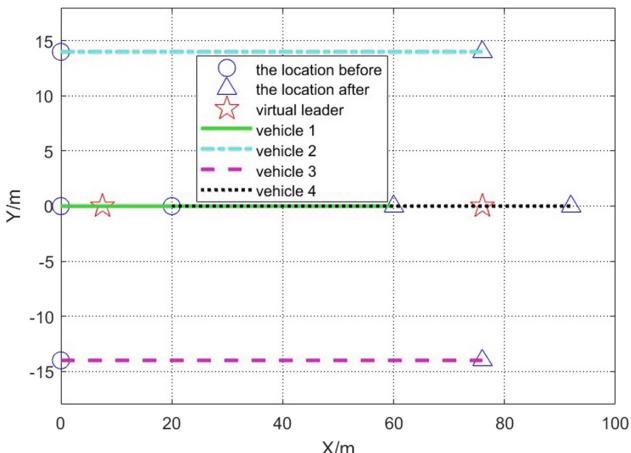


Fig. 3. Diamond formation is formed by Triangle formation

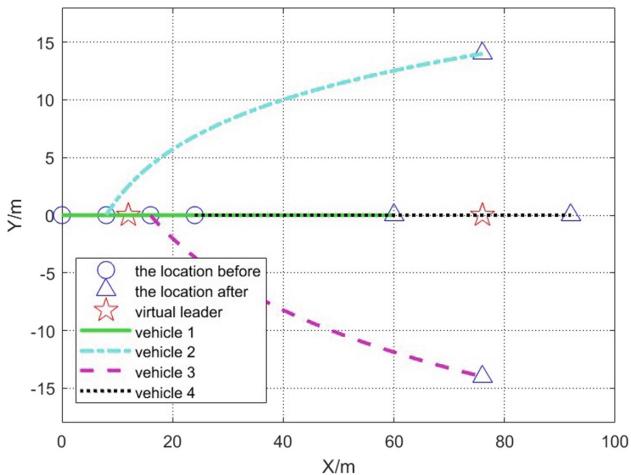


Fig. 4. Diamond formation is formed by straight line formation

7 Conclusion

This paper combines the virtual leader method strategy and the behavior-based bait-Predator system to design a control method. This control method can be applied to the distributed formation and multi-unmanned vehicle formations translation. Use the virtual leader to calibrate the target location and the interaction between the predator and the bait to form and transform formations without specifying the correspondence between vehicle and the target location. This method can avoid the formation chaos caused by the failure of the leader vehicle, and there is no need to plan the movement locus of each vehicle in

advance. Simulation experiments show that the control strategy can be well applied to distributed vehicle formation transforming and show the effectiveness and feasibility of the method. However it needs further improvement if it is to be applied to a real environment.

Acknowledgment. This research was supported in part by the National Natural Science Foundation of China under Grant No. 61672082 and 61822101, Beijing Municipal Natural Science Foundation No. L 191001.

References

1. Balch, T., Arkin, R.C.: Behavior-based formation control for multirobot teams. *IEEE Trans. Robot. Autom.* **14**(6), 926–939 (1998)
2. Beard, R.W., Lawton, J., Hadaegh, F.Y.: A coordination architecture for spacecraft formation control. *IEEE Trans. Control Syst. Technol.* **9**(6), 777–790 (2001)
3. Hu, J., Bhowmick, P., Lanzon, A.: Distributed adaptive time-varying group formation tracking for multiagent systems with multiple leaders on directed graphs. *IEEE Trans. Control Netw. Syst.* **7**(1), 140–150 (2019)
4. Khatib, O.: Real-time obstacle avoidance for manipulators and mobile robots. In: *Autonomous Robot Vehicles*, pp. 396–404. Springer (1986)
5. Liu, K.: Robot guidance planning based on pheromone diffusion guided by artificial potential field, pp. 06–0268–05 (2019)
6. Mehrjerdi, H., Ghommam, J., Saad, M.: Nonlinear coordination control for a group of mobile robots using a virtual structure. *Mechatronics* **21**(7), 1147–1155 (2011)
7. Wu, J.: Distributed formation transform control based on bait-predator system (21), 18–22 (2014)
8. Zhao, W., Yu, W., Zhang, H.: Observer-based formation tracking control for leader-follower multi-agent systems. *IET Control Theory Appl.* **13**(2), 239–247 (2018)



Event-Triggered Control for Linear Multi-agent Systems Under DoS Attacks

Lianghui Sun¹, Jianting Lyu^{1(✉)}, Dai Gao², and Xin Wang¹

¹ School of Mathematical Science,
Heilongjiang University, Harbin, People's Republic of China

jvjianting@hlju.edu.cn

² School of Mechatronics Engineering,
Harbin Institute of Technology, Harbin, People's Republic of China

Abstract. Concerning about general linear multi-agent systems with denial of service (DoS) attacks, this paper investigates the stability analysis and design procedure of fully distributed event-triggered control. For directed topologies, we propose an event-triggered controller and determine the scheduling of controller updating times subject to DoS attacks. Theoretical analysis shows that the proposed event-triggered controller guarantee the leader-following consensus and the exclusion of Zeno behavior. Finally, one case study is conducted to illustrate the effectiveness of the proposed method.

Keywords: Multi-agent systems · Event-triggered control · Denial-of-service (DoS) attack

1 Introduction

The coordination control problems in multi-agent systems have become one of the significant research issues in many areas such as artificial intelligent, mathematics, biology and control engineering. Many profitable achievements have emerged in studying consensus or coordinated tracking problem for multi-agent systems, for example, [1–5], to name a few.

With the development of the network, the transmission of many important information depends on the network communication technology, so the attack and defense issues of network security have become one of the focuses of attention. Generally speaking, attacks on communication networks are categorized as the denial of service (DoS) attacks and the deception ones. The DoS attacks are the more common ones, in which, the legitimate members of a multi-agent system are unable to access information resources due to the malicious cyber threat actors. Thus, one of the key problems for multi-agent control systems is to achieve consensus under DoS attack [6, 7], and some interesting results have been obtained [8–10].

Recently, a new method which is called event-triggered strategy is widely applied in dealing with the scenario that the bandwidth of the communication

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 326–335, 2021.

https://doi.org/10.1007/978-981-15-8458-9_35

channel in a multi-agent system is limited. The event-triggered control method has obvious advantages compared to the traditional control methods, which can effectively avoid continuous communication and reduce energy consumption. Therefore, in [11–17], different forms of event-triggered control schemes and algorithms are presented. Note that the above results all studied in the ideal network states, though, for the unreliable communication environment under DoS attacks, the research results of using event-triggered methods to the systems under DoS attacks are relatively few.

Compared with the consensus of multi-agent system [1–5], this paper studies the consensus of linear multi-agent system in the unreliable network environment with DoS attack. Different from the control strategy under DoS attack studied in [8–10], this paper considers an event-triggered control scheme, which effectively reduces unnecessary communication. Although reference [18] also studies the event-triggered strategy under DoS attack, the control strategy needs to obtain adjacent information through continuous communication when verifying the triggering conditions. The event-triggered strategy studied here can effectively avoid this problem. In this paper, leader-following consensus is studied, which is different from that of literature [19–21]. It is verified that the event-triggered strategy designed in this paper does not show Zeno behavior.

The rest of the paper is organized as follows. In Sect. 2, the necessary preliminaries are introduced. Controller design and stability analysis for leader-following multi-agent systems is presented in Sect. 3. Section 4 gives the simulation results, and finally the conclusions are drawn in Sect. 5.

2 Problem Formulation

Consider the following leader-following multi-agent system:

$$\begin{aligned} \text{Leader : } & \dot{x}_0(t) = Ax_0(t), \\ \text{Followers : } & \dot{x}_i(t) = Ax_i(t) + Bu_i(t), \quad i = 1, \dots, N, \end{aligned} \quad (1)$$

where $x_0(t) \in R^n$ is the state of the leader, $x_i(t) \in R^n$ is the state of follower agent i , $u_i(t) \in R^m$ is the control input of follower agent i . A and B are the known matrices with compatible dimensions, and the pair (A, B) is stabilizable.

Among the $N+1$ agents the communication topology is modeled by a directed graph $\bar{\mathcal{G}} = \mathcal{G} \cup \{0\}$, one member of which is the leader (labeled by 0) and the rest N agents are the followers (related to a directed graph \mathcal{G}). In this graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}\}$, the nodes set $\mathcal{V} = \{1, 2, \dots, N\}$, the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ and the adjacency matrix $\mathcal{A} = [a_{ij}] \in R^{N \times N}$. When $(j, i) \in \mathcal{E}$, $a_{ij} > 0$, and otherwise, $a_{ij} = 0$. Of the directed graph \mathcal{G} the Laplacian matrix is defined as $\mathcal{L} = (l_{ij}) \in R^{N \times N}$, in which the elements are $l_{ij} = -a_{ij} \leq 0, i \neq j$ and $l_{ii} = \sum_{j=1, j \neq i}^N a_{ij} \geq 0$. The adjacency matrix of the leader is defined as $\mathcal{B} = \text{diag}\{b_1, b_2, \dots, b_N\}$, in which $b_i > 0$ when the i -th agent has a direct edge from it to the leader, and otherwise $b_i = 0$. Next, the expression of the information exchange matrix can be obtained as $H = \mathcal{L} + \mathcal{B}$, and for a directed graph, H is asymmetric.

Assumption 1. *The digraph $\bar{\mathcal{G}}$ contains a directed spanning tree with leader as root.*

Lemma 1. *On the basis of the above definitions, there exists a positive vector $\xi = [\xi_1, \xi_2, \dots, \xi_N]^T \in R^N$, that satisfies $H^T \xi = 1_N$ and*

$$\Sigma H + H^T \Sigma > 0, \quad (2)$$

in which $\Sigma = \text{diag}\{\xi_1, \xi_2, \dots, \xi_N\}$.

The time sequence of DoS attacks as $\{t_m\}_{m \in N}$. There exists the m -th attack duration $\Delta_m > 0$ and there satisfies $t_{m+1} > t_m + \Delta_m$ for all $m \in N$. In addition, the set $[t_m + \Delta_m, t_{m+1})$ represents as time interval between two consecutive attacks. Therefore, the time sets that with and without DoS attacks are respectively represented as follows

$$\mathcal{A}_a(t_0, t) = \cup [t_m, t_m + \Delta_m) \cap [t_0, t], \quad m \in N \quad (3)$$

$$\mathcal{A}_b(t_0, t) = [t_0, t] \setminus \mathcal{A}_a(t_0, t). \quad (4)$$

Assumption 2. *For $\forall t > 0$, denote $T_a(t_0, t)$ as the total time interval for systems under DoS attacks during $[t_0, t]$, and there exist $T_0 \geq 0$ and $\tau_a > 1$ such that*

$$T_a(t_0, t) \leq T_0 + \frac{t - t_0}{\tau_a}. \quad (5)$$

Assumption 3. *For $\forall t > 0$, let $N_a(t_0, t)$ denote the number of DoS attacks occurring over $[t_0, t]$, and there exists $\omega > 0$ such that*

$$N_a(t_0, t) \leq \omega(t - t_0). \quad (6)$$

Definition 1. *For any initial $x_i(0)$, the leader-following consensus of multi-agent systems is obtained, if*

$$\lim_{t \rightarrow \infty} \|x_i(t) - x_0(t)\| = 0, \quad i = 1, 2, \dots, N. \quad (7)$$

In this paper, for the directed topological structure, our objective is to cope with the leader-following consensus problem of multi-agent systems under DoS attacks by designing event-triggered information transmission and controller.

3 Main Result

3.1 Event-Triggered Controller Design

For each follower agent i , the state estimate is defined as $\hat{x}_i(t) = e^{A(t-t_k^i)} x_i(t_k^i)$, $\forall t \in [t_k^i, t_{k+1}^i)$. The measurement error $e_i(t)$ is defined as

$$e_i(t) = \hat{x}_i(t) - x_i(t), \quad i = 1, \dots, N. \quad (8)$$

The event-based controller for each follower agent i is proposed as

$$u_i(t) = \begin{cases} K \sum_{j=1}^N a_{ij}(\hat{x}_i(t) - \hat{x}_j(t)) + b_i(\hat{x}_i(t) - x_0(t)), & \text{if } t \in \mathcal{A}_b(t_0, t) \\ 0, & \text{if } t \in \mathcal{A}_a(t_0, t) \end{cases} \quad (9)$$

where $K \in R^{m \times n}$ is the feedback gain matrix.

The event time t_k^i is defined as

$$t_k^i = \inf\{t > t_{k-1}^i : f_i(t) > 0\}, \quad (10)$$

where triggering function for each follower agent i is given by

$$f_i(t) = \|e_i(t)\| - c \left\| \sum_{j=1}^N a_{ij}(\hat{x}_i(t) - \hat{x}_j(t)) + b_i(\hat{x}_i(t) - x_0(t)) \right\|, \quad (11)$$

with c is a positive constant.

Remark 1. It should be noted that the update of the controller depends on the successful transmission of information. When the network works normally, the communication channels effectively provide the system the triggering function (10) by delivering information among the nodes, which leads to the controller updates if the condition is satisfied, or vice versa. When there are DoS attacks happening which obstruct the routine information exchange thus impede the updates of controller, the control input is assigned as zero until the communication channels are clear again.

3.2 Stability Analysis

Theorem 1. Consider the leader-following multi-agent systems (1). There exists $0 < \lambda < \alpha_1$ such that $\tau_a > (\alpha_1 + \alpha_2)/(\alpha_1 - \lambda)$ and $N_a(t_0, t) \leq \frac{\lambda}{\ln \mu}(t - t_0)$, where $\alpha_1 > 0$, $\alpha_2 > 0$, and $\mu = \max\{\xi_{\max}, \frac{1}{\xi_{\min}}\}$. Suppose that event-triggered function (11) and control protocol (9) are applied with $0 < c < \sqrt{\frac{1}{\lambda_{\max}(\bar{P}^{-1})}}$ and $K = M\bar{P}^{-1}$. Then the leader-following consensus problem can be solved if there exist matrices $\bar{P} > 0$ and M such that the following inequalities hold.

$$\begin{bmatrix} \Xi \Sigma H \otimes BM & 0 & H^T \otimes \bar{P} \\ * & -I_N \otimes \bar{P} & H^T \otimes \bar{P} & 0 \\ * & * & -\frac{1}{2N} I_{N \times n} & 0 \\ * & * & * & -\frac{1}{2N} I_{N \times n} \end{bmatrix} < 0 \quad (12)$$

and

$$A\bar{P} + \bar{P}A^T - \alpha_2\bar{P} < 0, \quad (13)$$

where $\Xi = \Sigma \otimes A\bar{P} + \Sigma \otimes \bar{P}A^T + \Sigma H \otimes BM + H^T\Sigma \otimes M^T B^T + \alpha_1(\Sigma \otimes \bar{P})$.

Proof. Define the state tracking error as $\delta_i(t) = x_i(t) - x_0(t)$. By combining (1) and (9), the dynamics of the error closed-loop system can be expressed as

$$\dot{\delta}(t) = \begin{cases} (I_N \otimes A + H \otimes BK)\delta(t) + (H \otimes BK)e(t), & \text{if } t \in \mathcal{A}_b(t_0, t) \\ (I_N \otimes A)\delta(t), & \text{if } t \in \mathcal{A}_a(t_0, t) \end{cases} \quad (14)$$

where $\delta(t) = (\delta_1^T(t), \dots, \delta_N^T(t))^T$, $e(t) = (e_1^T(t), \dots, e_N^T(t))^T$.

Case 1, when $t \in \mathcal{A}_b(t_0, t)$, consider the Lyapunov function candidate

$$V_1(t) = \delta^T(t)(\Sigma \otimes P)\delta(t). \quad (15)$$

Taking the time derivative of $V_1(t)$ gives

$$\begin{aligned} \dot{V}_1(t) &= 2\delta^T(t)(\Sigma \otimes P)[(I_N \otimes A)\delta(t) + (H \otimes BK)\delta(t) + (H \otimes BK)e(t)] \\ &= 2\delta^T(t)(\Sigma \otimes PA)\delta(t) + 2\delta^T(t)(\Sigma H \otimes PBK)\delta(t) \\ &\quad + 2\delta^T(t)(\Sigma H \otimes PBK)e(t) \\ &\leq \eta^T(t)\Pi_1\eta(t) - \alpha_1\delta^T(t)(\Sigma \otimes P)\delta(t) + \lambda_{\max}(P)e^T(t)e(t), \end{aligned} \quad (16)$$

where

$$\begin{aligned} \Pi_1 &= \begin{bmatrix} \Xi_1 & \Sigma H \otimes PBK \\ * & -I_N \otimes P \end{bmatrix}, \quad \eta^T(t) = [\delta^T(t) \ e^T(t)], \\ \Xi_1 &= \Sigma \otimes PA + \Sigma \otimes A^T P + \Sigma H \otimes PBK + H^T \Sigma \otimes K^T B^T P + \alpha_1(\Sigma \otimes P). \end{aligned}$$

By the triggering condition (11), one has

$$\begin{aligned} \|e_i(t)\| &\leq c \left\| \sum_{j=1}^N a_{ij}(\hat{x}_i(t) - \hat{x}_j(t)) + b_i(\hat{x}_i(t) - x_0(t)) \right\| \\ &= c \left\| \sum_{j=1}^N a_{ij}(e_i(t) - e_j(t)) + b_i e_i(t) + \sum_{j=1}^N a_{ij}(\delta_i(t) - \delta_j(t)) + b_i \delta_i(t) \right\| \\ &\leq c \|(H \otimes I_n)e(t) + (H \otimes I_n)\delta(t)\|, \end{aligned} \quad (17)$$

which implies

$$\sum_{i=1}^N \|e_i(t)\|^2 \leq 2Nc^2(\|(H \otimes I_n)e(t)\|^2 + \|(H \otimes I_n)\delta(t)\|^2). \quad (18)$$

Furthermore, it follows from (18) that

$$\|e(t)\|^2 \leq \eta^T(t) \begin{bmatrix} 2Nc^2(H^T H \otimes I_n) & 0 \\ 0 & 2Nc^2(H^T H \otimes I_n) \end{bmatrix} \eta(t). \quad (19)$$

According to formula $0 < c < \sqrt{\frac{1}{\lambda_{\max}(P^{-1})}}$, it can be derived from (16) and (19) that

$$\dot{V}_1(t) \leq \eta^T(t)\Pi_2\eta(t) - \alpha_1\delta^T(t)(\Sigma \otimes P)\delta(t), \quad (20)$$

where

$$\begin{aligned}\Pi_2 &= \begin{bmatrix} \Xi_2 & \Sigma H \otimes PBK \\ * & -I_N \otimes P + 2N(H^T H \otimes I_n) \end{bmatrix}, \\ \Xi_2 &= \Sigma \otimes PA + \Sigma \otimes A^T P + \Sigma H \otimes PBK + H^T \Sigma \otimes K^T B^T P + \alpha_1(\Sigma \otimes P) \\ &\quad + 2N(H^T H \otimes I_n).\end{aligned}$$

By Schur complement lemma, (12) is equivalent to

$$\Pi_3 = \begin{bmatrix} \Xi_3 & \Sigma H \otimes BM \\ * & -I_N \otimes \bar{P} + 2N(H^T H \otimes \bar{P}\bar{P}) \end{bmatrix} < 0,$$

where

$$\begin{aligned}\Xi_3 &= \Sigma \otimes A\bar{P} + \Sigma \otimes \bar{P}A^T + \Sigma H \otimes BM + H^T \Sigma \otimes M^T B^T + \alpha_1(\Sigma \otimes \bar{P}) \\ &\quad + 2N(H^T H \otimes \bar{P}\bar{P}).\end{aligned}$$

Multiplying both sides of Π_2 by $\text{diag}\{I_N \otimes \bar{P}, I_N \otimes \bar{P}\}$, we get that $\Pi_2 < 0$ is equivalent to $\Pi_3 < 0$. By (20), one has

$$\dot{V}_1(t) < -\alpha_1 V_1(t). \quad (21)$$

Case 2, when $t \in \mathcal{A}_a(t_0, t)$, consider the Lyapunov function candidate

$$V_2(t) = \delta^T(t)(I_N \otimes P)\delta(t). \quad (22)$$

From (13), the time derivative of $V_2(t)$ is given by

$$\dot{V}_2(t) = \delta^T(t)[I_N \otimes (A^T P + PA)]\delta(t) \leq \alpha_2 V_2(t). \quad (23)$$

For $[t_{m-1} + \Delta_{m-1}, t_m] \subseteq \mathcal{A}_b(t_0, t)$ and $[t_m, t_m + \Delta_m] \subseteq \mathcal{A}_a(t_0, t)$, it follows from (21) and (23) that

$$V(t) = \begin{cases} e^{-\alpha_1(t-t_{m-1}-\Delta_{m-1})} V_1(t_{m-1} + \Delta_{m-1}), & \text{if } t \in [t_{m-1} + \Delta_{m-1}, t_m) \\ e^{\alpha_2(t-t_m)} V_2(t_m), & \text{if } t \in [t_m, t_m + \Delta_m) \end{cases} \quad (24)$$

If $t \in [t_{m-1} + \Delta_{m-1}, t_m]$, from (24) it follows that

$$\begin{aligned}V(t) &\leq e^{-\alpha_1(t-t_{m-1}-\Delta_{m-1})} V_1(t_{m-1} + \Delta_{m-1}) \\ &\leq \mu e^{-\alpha_1(t-t_{m-1}-\Delta_{m-1})} V_2((t_{m-1} + \Delta_{m-1})^-) \\ &\leq \mu e^{-\alpha_1(t-t_{m-1}-\Delta_{m-1})} e^{\alpha_2(t_{m-1}+\Delta_{m-1}-t_{m-1})} V_2(t_{m-1}) \\ &\leq \dots \\ &\leq \mu^m e^{-\alpha_1|\mathcal{A}_b(t_0, t)|} e^{\alpha_2|\mathcal{A}_a(t_0, t)|} V_1(t_0).\end{aligned} \quad (25)$$

If $t \in [t_m, t_m + \Delta_m]$, from (24) it follows that

$$\begin{aligned}V(t) &\leq e^{\alpha_2(t-t_m)} V_2(t_m) \\ &\leq \mu e^{\alpha_2(t-t_m)} V_1(t_m^-) \\ &\leq \mu e^{\alpha_2(t-t_m)} e^{-\alpha_1(t_m-t_{m-1}-\Delta_{m-1})} V_1(t_{m-1} + \Delta_{m-1}) \\ &\leq \dots \\ &\leq \mu^{m+1} e^{-\alpha_1|\mathcal{A}_b(t_0, t)|} e^{\alpha_2|\mathcal{A}_a(t_0, t)|} V_1(t_0).\end{aligned} \quad (26)$$

According to Assumption 3, $N_a(t_0, t) = m$ for $t \in [t_{m-1} + \Delta_{m-1}, t_m)$ and $N_a(t_0, t) = m+1$ for $t \in [t_m, t_m + \Delta_m)$. Thus, for $\forall t \geq t_0$, it follows from (25) and (26) that

$$V(t) \leq \mu^{N_a(t_0, t)} e^{-\alpha_1 |\mathcal{A}_b(t_0, t)|} e^{\alpha_2 |\mathcal{A}_a(t_0, t)|} V(t_0). \quad (27)$$

In addition, for $t > t_0$, we have $|\mathcal{A}_b(t_0, t)| = t - t_0 - |\mathcal{A}_a(t_0, t)|$. By Assumption 2, it has

$$\begin{aligned} & -\alpha_1 |\mathcal{A}_b(t_0, t)| + \alpha_2 |\mathcal{A}_a(t_0, t)| \\ &= -\alpha_1(t - t_0 - |\mathcal{A}_a(t_0, t)|) + \alpha_2 |\mathcal{A}_a(t_0, t)| \\ &\leq -\alpha_1(t - t_0) + (\alpha_1 + \alpha_2)[T_0 + \frac{t - t_0}{\tau_a}]. \end{aligned} \quad (28)$$

According to (27) and (28), we get

$$\begin{aligned} V(t) &\leq \mu^{N_a(t_0, t)} e^{-\alpha_1(t - t_0) + (\alpha_1 + \alpha_2)[T_0 + \frac{t - t_0}{\tau_a}]} V(t_0) \\ &\leq e^{(\alpha_1 + \alpha_2)T_0} e^{N_a(t_0, t) \ln \mu} e^{(-\alpha_1 + \frac{\alpha_1 + \alpha_2}{\tau_a})(t - t_0)} V(t_0). \end{aligned} \quad (29)$$

According to $\tau_a > (\alpha_1 + \alpha_2)/(\alpha_1 - \lambda)$ and $N_a(t_0, t) \leq \frac{\lambda}{\ln \mu}(t - t_0)$, let $\gamma = \alpha_1 - (\alpha_1 + \alpha_2)/\tau_a - \lambda > 0$. Then (29) can be expressed as

$$V(t) \leq e^{(\alpha_1 + \alpha_2)T_0} e^{-\gamma(t - t_0)} V(t_0). \quad (30)$$

Remark 2. Theorem 1 indicates the proposed event-triggered controller can guarantee the leader-following consensus in the systems with DoS attack, and can effectively deplete the continuous communication among adjacent nodes, thus the controller's update frequency drops significantly.

Theorem 2. *There exist no Zeno behaviour in the multi-agent systems (1).*

Proof. Let $q_i(t) = \sum_{j=1}^N a_{ij}(\hat{x}_i(t) - \hat{x}_j(t)) + b_i(\hat{x}_i(t) - x_0(t))$ and $\zeta(t) = \frac{\|e(t)\|}{\|q(t)\|}$. When any two consecutive triggering are induced by the triggering condition (11), we can estimate the time derivative of $\zeta(t)$ to obtain $\dot{\zeta}(t) \leq \frac{\|\dot{e}(t)\|}{\|q(t)\|} + \frac{\|e(t)\|\|\dot{q}(t)\|}{\|q(t)\|^2}$. With the help of (8) and definition of $q_i(t)$, one obtains

$$\|\dot{e}(t)\| \leq \|I_N \otimes A\| \|e(t)\| + \|I_N \otimes BK\| \|q(t)\|, \quad (31)$$

$$\|\dot{q}(t)\| \leq \|I_N \otimes A\| \|q(t)\|. \quad (32)$$

Combining (31) and (32), one has $\dot{\zeta}(t) \leq c_1 \zeta(t) + c_2$, where $c_1 = 2\|I_N \otimes A\|$, $c_2 = \|I_N \otimes BK\|$. Then $\zeta(t)$ satisfies the bound $\zeta(t) \leq \varphi(t, \varphi_0)$, in which $\varphi(t, \varphi_0)$ is the solution of $\dot{\varphi} = c_1 \varphi + c_2$, $\varphi(0, \varphi_0) = \varphi_0$, given by $\varphi(t, \varphi_0) = -\frac{c_2}{c_1} + (\frac{c_2}{c_1} + \varphi_0)e^{c_1 t}$. Then the inter-execution times are bounded by the solution $\iota \in \mathbb{R}^+$ of $\varphi(\iota, 0) = c\sqrt{N}$. Since $\varphi(\iota, 0) = \frac{c_2}{c_1}(e^{c_1 \iota} - 1)$ we obtain $\iota = \frac{\ln((c_1/c_2)c\sqrt{N}+1)}{c_1}$, which is strictly positive. So Zeno behaviour is excluded for follower agent i .

4 Simulation Example

In this section, a example is presented to illustrate the effectiveness of the theoretical results. The network topology associated with five agents is shown in Fig. 1.

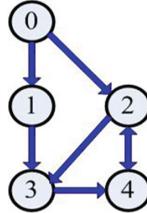


Fig. 1. Network topology associated with five agents

The parameters are selected as

$$A = \begin{bmatrix} -0.059 & 0.496 & -0.868 \\ -5.513 & -0.939 & 0.665 \\ 0.068 & 0.026 & -0.104 \end{bmatrix}, \quad B = \begin{bmatrix} 0.006 & 0.006 & 0.004 & 0 & 0.09 \\ 1.879 & 1.328 & 0.029 & 0.675 & 0.217 \\ -0.109 & -0.096 & -0.084 & 0.007 & -2.974 \end{bmatrix}.$$

The initial conditions and related parameters with controller design are selected as: $x_0(0) = [0, -0.5, 0.2]^T$, $x_1(0) = [-3, -6.8, 0.5]^T$, $x_2(0) = [0, -4.5, 0.2]^T$, $x_3(0) = [1, 2, -0.2]^T$, $x_4(0) = [-1, -2.8, -0.3]^T$, $\alpha_1 = 0.25$, $\alpha_2 = 1.2$, $c = 0.9$. Matrix M and K by solving (12) in Theorem 1 are derived as follows

$$M = \begin{bmatrix} 0.1726 & 0.2036 & -1.2700 \\ -0.2207 & -0.2037 & 1.5008 \\ 0.1561 & 0.0795 & -0.9218 \\ -0.0520 & -0.1684 & 0.6140 \\ -0.0037 & -0.0035 & 0.0256 \end{bmatrix}, \quad K = \begin{bmatrix} 0.3384 & 0.1913 & -1.9550 \\ -0.4131 & -0.2232 & 2.3079 \\ 0.2700 & 0.1336 & -1.4145 \\ -0.1387 & -0.0978 & 0.9499 \\ -0.0070 & -0.0038 & 0.0393 \end{bmatrix}.$$

The DoS jamming signal, the controller update time instant and the corresponding time interval and state of follower x_1 are presented in Figs. 2.

The consensus tracking errors are exhibited in Fig. 3, which the multi-agent systems (1) are subjected to DoS attacks. Based on the above analysis, the event-triggered controller adopted in this paper can effectively reduce the number of communications, and the leader-following consensus can be achieved under DOS attack.

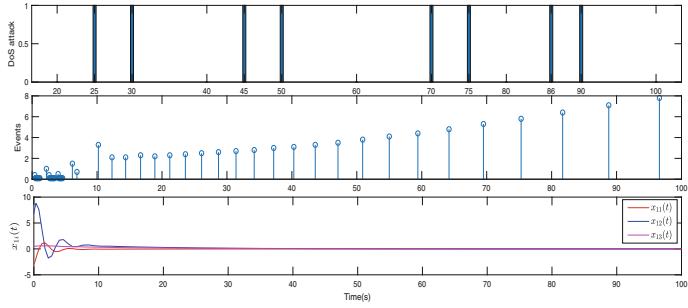


Fig. 2. Signal of DoS attacks and event instants and state of follower x_1

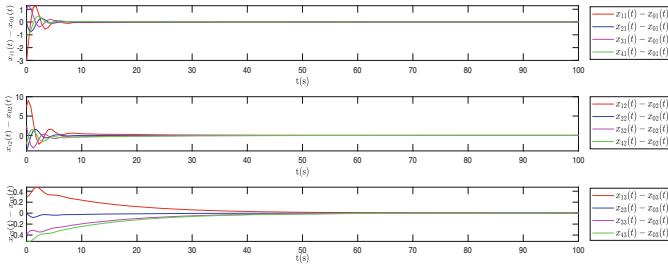


Fig. 3. Trajectories of tracking errors $x_i - x_0, i = 1, \dots, 4$

5 Conclusion

One kind of typical leader-following consensus problem of linear multi-agent system under DoS attack has been developed in this article. When the duration and frequency of DoS attack meet certain conditions, it is proved that the proposed event-triggered control strategy guarantees the leader-following consensus. Furthermore, the Zeno behavior is ruled out for communication channels, nor does it exist in controller updates. Future work will try to extend the event-based consensus control results to nonlinear multi-agent system under DoS attack.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (Grant nos. 61873087 and 61703148).

References

1. Li, Z., Wen, G., Duan, Z., Ren, W.: Designing fully distributed consensus protocols for linear multi-agent systems with directed graphs. *IEEE Trans. Autom. Control* **60**(4), 1152–1157 (2015)
2. Qin, J., Ma, Q., Shi, Y., Wang, L.: Recent advances in consensus of multi-agent systems: a brief survey. *IEEE Trans. Industr. Electron.* **64**(6), 4972–4983 (2017)
3. Ding, L., Han, Q., Guo, G.: Network-based leader-following consensus for distributed multi-agent systems. *Automatica* **49**(7), 2281–2286 (2013)

4. Ni, W., Cheng, D.: Leader-following consensus of multi-agent systems under fixed and switching topologies. *Syst. Control Lett.* **59**(3), 209–217 (2010)
5. Rezaee, H., Abdollahi, F.: Adaptive leaderless consensus control of strict-feedback nonlinear multiagent systems with unknown control directions. *IEEE Trans. Syst. Man Cybern. Syst.* 1–10 (2020)
6. Feng, S., Tesi, P.: Resilient control under denial-of-service: robust design. *Automatica* **79**(79), 42–51 (2017)
7. De Persis, C., Tesi, P.: Input-to-state stabilizing control under denial-of-service. *IEEE Trans. Autom. Control* **60**(11), 2930–2944 (2015)
8. Zhang, D., Liu, L., Feng, G.: Consensus of heterogeneous linear multiagent systems subject to aperiodic sampled-data and dos attack. *IEEE Trans. Cybern.* **49**(4), 1501–1511 (2019)
9. Feng, Z., Hu, G., Wen, G.: Distributed consensus tracking for multi-agent systems under two types of attacks. *Int. J. Robust Nonlinear Control* **26**(5), 896–918 (2016)
10. Feng, Z., Wen, G., Hu, G.: Distributed secure coordinated control for multiagent systems under strategic attacks. *IEEE Trans. Cybern.* **47**(5), 1273–1284 (2017)
11. Cheng, B., Li, Z.: Coordinated tracking control with asynchronous edge-based event-triggered communications. *IEEE Trans. Autom. Control* **64**(10), 4321–4328 (2019)
12. Du, S., Liu, T., Ho, D.W.: Dynamic event-triggered control for leader-following consensus of multiagent systems. *IEEE Trans. Syst. Man Cybern. Syst.* **50**(9), 3243–3251 (2020)
13. Hu, W., Yang, C., Huang, T., Gui, W.: A distributed dynamic event-triggered control approach to consensus of linear multiagent systems with directed networks. *IEEE Trans. Cybern.* **50**(2), 869–874 (2020)
14. Cheng, B., Li, Z.: Fully distributed event-triggered protocols for linear multiagent networks. *IEEE Trans. Autom. Control* **64**(4), 1655–1662 (2019)
15. Yang, D., Ren, W., Liu, X., Chen, W.: Decentralized event-triggered consensus for linear multi-agent systems under general directed graphs. *Automatica* **69**, 242–249 (2016)
16. Zhu, W., Jiang, Z.-P., Feng, G.: Event-based consensus of multi-agent systems with general linear models. *Automatica* **50**(2), 552–558 (2014)
17. Guo, G., Ding, L., Han, Q.-L.: A distributed event-triggered transmission strategy for sampled-data consensus of multi-agent systems. *Automatica* **50**(5), 1489–1496 (2014)
18. Xu, W., Ho, D.W., Zhong, J., Chen, B.: Event/self-triggered control for leader-following consensus over unreliable network with DoS attacks. *IEEE Trans. Neural Netw.* **30**(10), 3137–3149 (2019)
19. Yang, Y., Li, Y., Yue, D.: Event-trigger-based consensus secure control of linear multi-agent systems under DoS attacks over multiple transmission channels. *Sci. Chin. Inf. Sci.* **63**(5), 105–118 (2020)
20. Xu, Y., Fang, M., Wu, Z., Pan, Y., Chadli, M., Huang, T.: Input-based event-triggering consensus of multiagent systems under denial-of-service attacks. *IEEE Trans. Syst. Man Cybern. Syst.* **50**(4), 1455–1464 (2020)
21. Hu, S., Yue, D., Xie, X., Chen, X., Yin, X.: Resilient event-triggered controller synthesis of networked control systems under periodic DoS jamming attacks. *IEEE Trans. Cybern.* **49**(12), 4271–4281 (2019)



A Celestial Navigation Approach Based on Star Pattern Recognition

Boning Wang, Wei Shao^(✉), Hanxue Zhao, Lingfei Dou, and Wenlong Yao

College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, China
greatshao@126.com

Abstract. Star pattern recognition plays an important role in celestial navigation. Aiming at the shortcomings of the traditional star map recognition algorithm such as large storage space and high recognition error rate, an improved star pattern recognition navigation algorithm is proposed. Firstly, the navigation star database is constructed based on the idea of star angular distance. Then a new companion is introduced to eliminate the mismatched results and reduce the mismatched rate. Finally, the position information of the spacecraft is solved by the star position information in the field of view of the star sensor. Simulation results demonstrate the effectiveness and high accuracy of the proposed navigation method.

Keywords: Star tracker · Star pattern recognition · Celestial navigation

1 Introduction

In this section, the background, research significance and common methods of star map recognition for celestial navigation are introduced. Aiming at the shortcomings of the existing methods, the improved methods and research objectives proposed in this paper are summarized.

Celestial navigation is the use of celestial sensors to realize the observation of natural celestial bodies, according to the observation of the celestial body's position information for autonomous positioning and navigation [1]. Since the extraterrestrial objects are far away from the earth, the communication between the earth and the spacecraft has a long delay, so the probe can only rely on its own equipment to complete the navigation task [2]. The star tracker has the advantages of high pointing accuracy, small size, strong autonomy and no attitude accumulated error, etc., making the star tracker system a key system to realize the autonomous attitude control of the spacecraft [3]. It relies on the autonomous star map recognition technology to realize real-time autonomous monitoring of the spacecraft attitude. Therefore, the in-depth study of the key technologies of star sensors is of great significance to the attitude control of spacecraft.

For celestial navigation, stars can be regarded as an ideal point light source with certain spectral characteristics, which is located at infinity and is nearly stationary. The processing of star catalogue and star map is the basic work of star map recognition [4]. The basic idea of star map recognition is to compare the star map features captured by the star sensor with the star patterns in the navigation feature library to determine the current attitude of the spacecraft. Triangle algorithm is the most widely used and mature star map recognition algorithm [5]. This method is relatively simple, but generally requires a large storage space, the recognition algorithm is not real-time, and is prone to noise interference and redundancy matching [6]. In order to solve these problems, many scholars put forward an improved method to increase the dimension of features. The pyramid algorithm uses the fourth observation star to form a polygon structure and the k-vector method to search the navigation star feature library for matching results, which reduces the probability of redundant matching and mismatching to some extent [7]. The raster algorithm divides the observation star and its adjacent star into two-dimensional raster according to certain rules, and then converts the two-dimensional raster into a one-dimensional 0–1 string as the recognition feature, which greatly reduces the complexity of the algorithm [8]. At the same time, the intelligent recognition methods represented by neural network algorithm, genetic algorithm and ant colony algorithm are also applied in the process of star map recognition [9, 10]. However, this kind of algorithm requires a large number of sample set training features, which is more complex than the star map feature recognition algorithm and difficult to acquire features, so it is difficult to be applied in practice.

In this paper, a fast and efficient star pattern recognition algorithm is proposed based on the idea of star angular distance and the theory of partitioned star list. The reliability and real-time performance of the algorithm were verified by using sky2000 database. The content of this paper is arranged as follows: In Sect. 2, the attitude observation model of spacecraft in celestial coordinate system is established. In Sect. 3, the construction of the navstar database and the selection process of matching results are introduced. In Sect. 4, the algorithm proposed in this paper is simulated and verified by taking the simulated star map as an example. Finally, the paper closes with conclusions in Sect. 5.

2 Establishment of Celestial Navigation Model

In this section, the model of celestial navigation based on star tracker is established, and the formula for solving the position of spacecraft in celestial coordinate system is derived.

In astronomy, the celestial sphere is an ideal sphere with the center of the earth and an infinite radius. It is stipulated in the celestial coordinate system ($\Theta - \alpha\delta$) that right ascension α should start at the vernal equinox point and be measured in a counterclockwise direction, ranging from 0° to 360° . It is stipulated that the declination δ is measured from the equator to the north and south respectively, and the range is $0^\circ \sim \pm 90^\circ$. In this paper, we use right ascension

and declination to represent the position of stars in celestial coordinates. At the same time, due to the fixed physical position of the star sensor and the spacecraft, in order to simplify the derivation process, it is assumed that the star sensitive coordinate system ($O - xyz$) coincides with the spacecraft's own coordinate system.

The position information of star P_i in celestial coordinate system is represented by right ascension α_i and declination δ_i , and is converted into coordinates of cartesian coordinate system through Eq.(1).

$$A_i = \begin{bmatrix} \cos\alpha_i \cos\delta_i \\ \cos\alpha_i \sin\delta_i \\ \sin\delta_i \end{bmatrix} = \begin{bmatrix} I_{i1} \\ I_{i2} \\ I_{i3} \end{bmatrix} \quad (1)$$

The relationship between celestial coordinate system and star tracker coordinate system is shown in Fig. 1, then the direction vector of star P_i in star tracker coordinate system can be expressed as:

$$B_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = M^T A_i \quad (2)$$

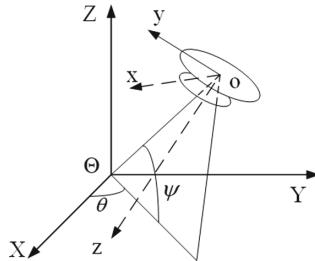


Fig. 1. Diagram of celestial coordinate system and star tracker coordinate system

Where, M is the transformation matrix between celestial coordinate system and star tracker. The attitude angle of the star tracker about the xyz axis is represented by the roll angle ϕ , pitch angle ψ and yaw angle θ respectively, then:

$$M = \begin{bmatrix} b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi - \cos\theta\sin\phi\sin\psi & -\sin\theta\sin\phi - \cos\theta\cos\phi\sin\psi & -\cos\theta\cos\psi \\ -\cos\theta\cos\phi - \sin\theta\sin\phi\sin\psi & \cos\theta\sin\phi - \sin\theta\cos\phi\sin\psi & -\sin\theta\cos\psi \\ \sin\phi\cos\psi & \cos\phi\cos\psi & -\sin\psi \end{bmatrix} \quad (3)$$

Obviously, the expression can be:

$$\begin{cases} \theta = \arctan\left(\frac{c_3}{b_3}\right) \\ \phi = \arctan\left(\frac{d_1}{d_2}\right) \\ \psi = \arcsin(-d_3) \end{cases} \quad (4)$$

The unit direction vector of the $z-axis$ of the star sensor is denoted as $\vec{z} = (0, 0, 1)$, and the angle between the direction vector of the star P_i and the direction of the $z-axis$ under the coordinate of the star tracker is denoted as σ_i , then:

$$\vec{z}^T \cdot B_i = |\vec{z}| \cdot |B_i| \cdot \cos\sigma_i = \cos\sigma_i \quad (5)$$

The distance between the image point of star P_i on the photosensitive surface and the center of the photosensitive surface is denoted as a_i , then

$$\cos\sigma_i = \frac{f}{\sqrt{a_i^2 + f^2}} \quad (6)$$

Where f represents the focal length of the star tracker. Substituting Eq. (5) into Eq. (6) yields:

$$\vec{z}^T \cdot \vec{M}^T \cdot A_i = b_3 I_{i1} + c_3 I_{i2} + d_3 I_{i3} = \frac{f}{\sqrt{a_i^2 + f^2}} \quad (7)$$

It is obtained when at least three non-collinear stars $P_i(\alpha_i, \delta_i)(i = 1, 2, 3)$ are observed:

$$\begin{cases} f_i(b_3, c_3, d_3) = b_3 I_{i1} + c_3 I_{i2} + d_3 I_{i3} - \frac{f}{\sqrt{a_i^2 + f^2}} \\ f_4(b_3, c_3, d_3) = b_3^2 + c_3^2 + d_3^2 - 1 \end{cases} \quad (8)$$

The direction of $z-axis$ in the star tracker coordinate system is determined by yaw angle θ and pitch angle ψ , and

$$\begin{cases} \theta = \alpha \\ \psi = \delta \end{cases} \quad (9)$$

By substituting into Eq. (4), the position coordinates of the spacecraft in the celestial coordinate system at the time can be obtained.

3 Star Pattern Recognition Algorithm

3.1 The Construction of the Navigation Star Database

Based on the *sky2000* catalog, it is assumed that the magnitude sensitivity limit of the star sensor is 5.5 mv, and 5103 stars whose brightness is higher than (or equal to) 6 Mv are selected to form the catalog. Its distribution is shown in Fig. 2.

The triangle algorithm uses three stars to form an observation triangle, and finds a unique matching isomorphism pattern for the observation triangle in the navigation star feature library. Three observation stars that are not on the same line in the field of view are selected to form a triangle, with each star forming a side and each side corresponding to an angular distance. If the coordinates of the

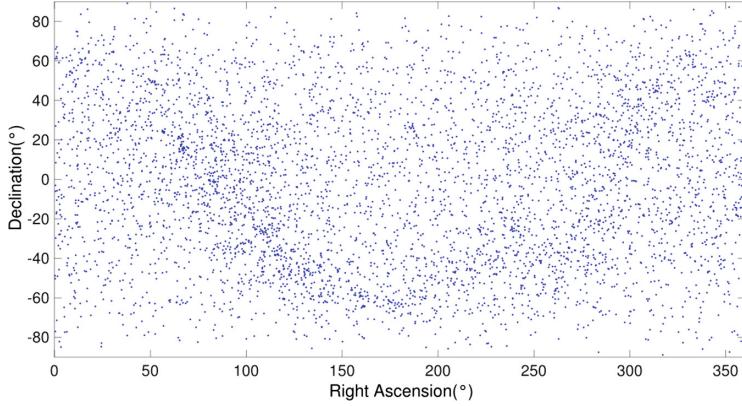


Fig. 2. Distribution of navigation stars

navigation stars i and j are (α_i, δ_i) and (α_j, δ_j) , respectively, then the angular distance in the celestial coordinate system is defined as:

$$d(i, j) = \arccos \left(\frac{\vec{A}_i \cdot \vec{A}_j}{|\vec{A}_i| \cdot |\vec{A}_j|} \right) \quad (10)$$

where \vec{A}_i and \vec{A}_j are the direction vectors of the navigation star i and j in the celestial coordinate system, respectively. The spacecraft uses the star map image acquired by the star tracker for autonomous navigation and obtains the coordinates of the navigation star in the image plane coordinate system. Let the coordinates of observed stars 1 and 2 on the star map image be (x_1, y_1) and (x_2, y_2) respectively, then the angular distance in the star tracker plane coordinate system can be defined as:

$$d_{12}^c = \arccos \left(\frac{\vec{s}_1 \cdot \vec{s}_2}{|\vec{s}_1| \cdot |\vec{s}_2|} \right) \quad (11)$$

Where \vec{s}_1 and \vec{s}_2 are the direction vectors of observed stars 1 and 2 in the plane coordinate system of the star sensor, and f is the focal length of CCD .

$$\vec{s}_1 = \frac{1}{\sqrt{x_1^2 + y_1^2 + f^2}} \begin{pmatrix} x_1 \\ y_1 \\ -f \end{pmatrix}, \vec{s}_2 = \frac{1}{\sqrt{x_2^2 + y_2^2 + f^2}} \begin{pmatrix} x_2 \\ y_2 \\ -f \end{pmatrix} \quad (12)$$

As shown in Fig. 3, three stars (i, j, k) are selected from the simple star list as the matching objects, where $d_{jk} > d_{ij} > d_{ik}$. The observed star $(1, 2, 3)$ on the imaging plane of the star sensor is the object to be matched, where $d_{23}^c > d_{12}^c > d_{13}^c$.

Therefore, based on the information of the selected 5,103 stars, C_{5103}^3 triangles are generated to form the navigation catalogue.

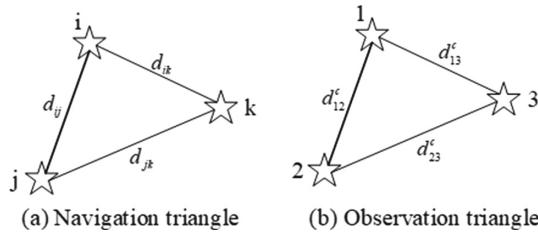


Fig. 3. Schematic diagram of triangle matching principle

3.2 Star Pattern Recognition Process

In the navigation star database, a large amount of feature information between stars is stored, which can easily cause data congestion and confusion in the matching process and increase the mismatching rate. When multiple matching results occur, global matching needs to be performed again, resulting in data redundancy and increasing matching time. Therefore, this paper designs a method of partitioning the star list to reduce the redundant time in the matching process.

As shown in Fig. 4, in the celestial coordinate system, the right ascension is divided into 10 groups ($S_1 - S_{10}$) with 90° as a group, and declination is divided into 12 groups ($N_1 - N_{12}$) with 90° as a group, forming a total of 120 regions.

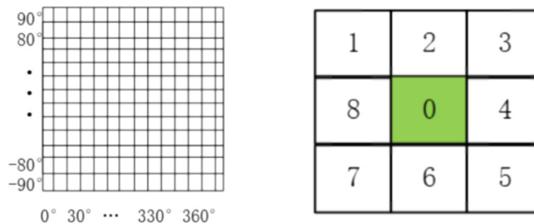


Fig. 4. Navigation star table partition rules

After the partition catalog is adopted, three star map points are selected to construct the triangular shape to be matched, and the first global search is conducted to obtain their coordinates in the celestial coordinate system. Due to the constraint of the field of view angle of the star sensor, in the process of solving the other star image points of the same star map, only the region where the initial search result is located and its eight adjacent regions need to be matched, which reduces the matching range, speeds up the matching speed and enhances the rapidity of the recognition algorithm.

When there is a mismatching phenomenon, the fourth star is introduced as a companion star through the screening of the star image points, and a new triangle is formed with the long side of the mismatched triangle. Then, according to the

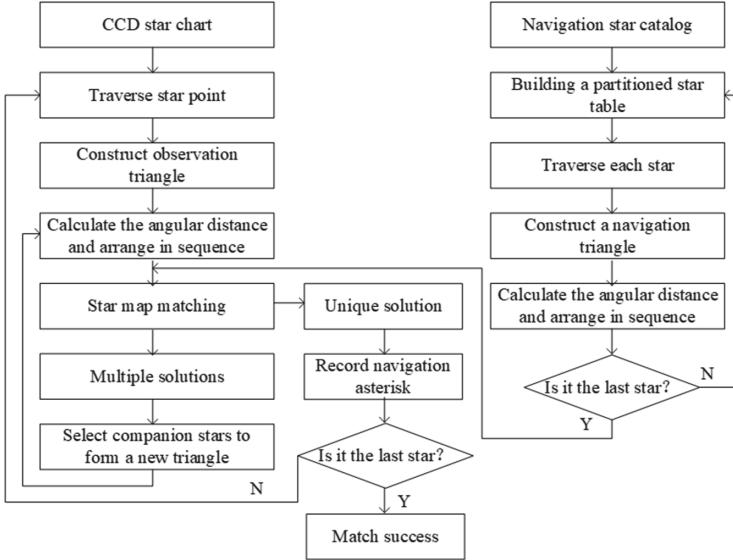


Fig. 5. Star pattern recognition based on improved triangle algorithm

partitioned star list method, all the matching results were recorded in the region, and the new triangle was matched again in the region and its eight adjacent regions, so as to eliminate the mismatched results and improve the success rate of matching. The star map matching flow chart based on the improved triangle algorithm is shown in Fig. 5.

4 Analysis of Simulation Results

The experimental simulation environment was an *Intel Core i5 – 6300 2.30 GHz PC*. The simulated field of view of the star sensor is $12^\circ \times 12^\circ$, and the number of pixels is 512×512 . The star map taken by the star sensor is generated according to the simulation of 5,103 stars whose brightness is higher than (or equal to) 6Mv in the *sky2000* catalog by *Matlab2016a* software.

In the simulation experiment, by setting the parameters of the simulation experiment, including the size of the field of view, right ascension and declination, the corresponding simulated star map is automatically obtained and preprocessed to identify it. In the experiment, 100 groups of test data were randomly generated, and examples of right ascension and declination data of one group of test data are shown in Table 1.

Table 1. Examples of right ascension and declination of test data

Star point number	The X coordinate	The Y coordinate
A01	22.64	489.48
A02	38.14	117.35
A03	75.33	189.14
A04	108.24	409.06
A05	141.67	352.92

The experiment tested the average recognition time of different recognition algorithms under the simulated field of view of $12^\circ \times 12^\circ$, as shown in Table 2. As a comparison, it can be seen that the average recognition time of the algorithm in this paper is 4.720640 s, which is only 30% of that of the triangle recognition algorithm.

Table 2. Comparison of recognition time between different recognition algorithms

Algorithm	Recognition time
Recognition algorithm in this paper	4.720640 s
Triangle recognition algorithm	15.736607 s

After partitioning, the processor only needs to store the information of each area every time it matches. After the current partition information matching is completed, the storage area can be emptied and put into the data of the new area. Therefore, the required storage capacity is related to the total area occupied by each area. The storage capacity after partitioning is only 0.2 of the original required capacity. It can be seen that the algorithm in this paper reduces the running time and storage space of the algorithm.

Meanwhile, in the second section, when the star sensor recognizes the position coordinates of at least three stars, the position and attitude information of the spacecraft itself at the current moment can be solved. As a direct follow-up step of star map recognition, the test results of star map attitude solution are shown in Table 3.

The accuracy of the three groups of simulation results in the table is above 97%, which meets the requirements of the general star sensor posture accuracy.

Table 3. The result of star catalogue attitude calculation

Number		Input quantity (known star)			
Group	Star number	Right ascension α_i	Declination δ_i	Projector distance a_i/mm	
1	1	10.2799	39.4587	0.093	
	2	9.3384	35.3995	5.9031	
	3	9.7912	49.3546	24.43	
2	1	10.3363	39.4587	0.0987	
	2	9.7384	38.3995	0.7139	
	3	9.7912	45.3546	7.913	
3	1	10.2074	37.9742	1.1332	
	2	8.7894	42.7865	2.3588	
	3	9.791	40.3764	0.0458	
Output (current position)		Reference value		Accuracy/%	
Right ascension	Declination	Right ascension	Declination	Right ascension	Declination
9.9759	39.9896	10	40	99.76	99.97
9.9975	39.9983			99.98	99.99
9.7949	39.97			97.95	99.92

5 Conclusion

In this paper, the autonomous navigation of astronomy based on star sensors is studied. Firstly, the mathematical model of estimating the attitude of spacecraft by using stars is analyzed and derived, and then a star map matching method based on the traditional triangle algorithm is designed. In the laboratory environment, the improved algorithm was tested, compared with the traditional triangle algorithm in storage capacity and average operating time, and the estimation accuracy of spacecraft pose was obtained through simulation experiments. It can be concluded from the experimental data that the algorithm in this paper can complete the autonomous celestial navigation task of spacecraft.

Acknowledgments. This work is supported by National Nature Science Foundation under Grant No. 61773227 and 61971253.

References

1. Nguyen, V.S., Im, N.K., Dao, Q.D.: Azimuth method for ship position in celestial navigation. *Int. J. E-Navig. Marit. Econ.* **7**, 55–62 (2017). <https://doi.org/10.1016/j.enavi.2017.06.006>
2. Liu, H.B., Yang, J.K., Wang, J.Q., et al.: Star spot location estimation using Kalman filter for star tracker. *Appl. Opt.* **50**(12), 1735–1744 (2011). <https://doi.org/10.1364/AO.50.001735>
3. Zhu, H., Liang, B., Zhang, T.: A robust and fast star identification algorithm based on an ordered set of points pattern. *Acta Astronaut.* (2018):S0094576517307658. <https://doi.org/10.1016/j.actaastro.2018.04.026>

4. Quan, W., Gong, X., Fang, J., Li, J.: Star map processing algorithm of star sensor and autonomous celestial navigation. INS/CNS/GNSS Integrated Navigation Technology. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-45159-5_5
5. Zhang, L., Zhou, Y., Lin, R., et al.: Fast triangle star pattern recognition algorithm. *J. Appl. Opt.* (2018). <https://doi.org/10.5768/JAO201839.0102005>
6. Nabi, A., Ahmed-Foitih, Z., Cheriet, E.A.: Improved triangular-based star pattern recognition algorithm for low-cost star trackers. *J. King Saud Univ. Comput. Inf. Sci.* (2019). <https://doi.org/10.1016/j.jksuci.2019.01.008>
7. Wu, F., Shen, W., Zhou, J., et al.: Design and simulation of a novel APS star tracker. In: International Conference on Optical Instruments & Technology. <https://doi.org/10.1117/12.806698>
8. Yi, W.J., Liu, H.B., Yang, J.K., et al.: Three-dimensional grid algorithm for all-sky autonomous star identification. In: Proceedings of SPIE the International Society for Optical Engineering, vol. 8420, pp. 08 (2012). <https://doi.org/10.1117/12.977625>
9. Yao, L., Xiao-Xiang, Z., Rong-Yu, S.: A star pattern recognition method based on decreasing redundancy matching. *Chin. Astron. Astrophys.* **40**(2), 277–290 (2016). <https://doi.org/10.1016/j.chinastron.2016.05.010>
10. Hernandez, E.A., Alonso, M.A., Chavez, E., et al.: Robust polygon recognition method with similarity invariants applied to star identification. *Adv. Space Res.* **59**(4), 1095–1111 (2017). <https://doi.org/10.1016/j.asr.2016.11.016>



Consensus Control of High-Dimensional Multi-agent Systems: A Large-Scale System Approach

Bin Zhang^(✉)

School of Automation, Beijing University of Posts and Telecommunications (BUPT),
Beijing 100876, China
zb362301@126.com

Abstract. This paper considers the consensus problem of complex dynamic multi-agent systems. Compared with the traditional model, a new dynamic model with leader is proposed, which is convenient for practical operation. In this paper, whether the state of the model can be achieved is reduced to the existence of solutions of a matrix equation system. Based on the theory of dynamic decentralized feedback for large-scale systems, sufficient conditions for the existence of consensus protocol are given, and the advantages and disadvantages of this kind of model are analyzed.

Keywords: Multi-agent systems · Consensus · Distributed protocol · Large system

1 Introduction

The collective behavior of multi-agent system has been widely used in biology, physics, engineering and other disciplines, and has attracted great attention of researchers in recent years [1–3]. The consensus control of multi-agent system is a representative problem in the field of multi-agent coordination, and it is also the research basis of other distributed control and estimation [4,5]. At present, most of the research work on consensus control of multi-agent system assumes that the agent has first-order and second-order dynamic model, but the control algorithm based on simple model can not solve the practical problems [6]. Therefore, it is of great value in theory and practice to study the consensus control of multi-agent system with complex dynamic model [7]. In view of this point, this paper takes the complex dynamic model as the research object, and does the following work:

The consensus control problem of intelligent system described by general linear time-invariant dynamic equation is considered. According to the needs of practical application, it is divided into zero vector state consistency problem and general constant vector state consistency problem. Based on the conclusion of decentralized dynamic feedback, the sufficient and necessary conditions for multi-agent system to achieve zero state consistency are given. For the general constant

value vector problem, this paper mainly summarizes the previous research. It is difficult to apply the necessary and sufficient condition that multi-agent system can achieve zero state consensus. Researchers can only propose several special protocol forms to simplify the design, which needs to be improved.

Considering the practical application, this paper proposes a new model to study the consensus of intelligent systems. The problem that the model can reach a uniform state is reduced to the problem of the existence of solutions for a system of linear matrix equations. A sufficient condition for the existence of solutions for this system of linear matrix equations is given.

2 Preliminary Results

A. Basic Symbols and Concepts

Unless otherwise specified, the following notation will always be used in this paper: R and C represent the real number field and the complex number field, respectively. The symbols 1_n and 0_n are column vectors with dimension n . For a given matrix X , \bar{X} represents the conjugate matrix; X^T represents the transposition; X^* represents the conjugate transpose; and $N(X)$ represents the zero space.

Definition 1. Let $A = [a_{ij}] \in C^{m \times n}$ and $B = [b_{ij}] \in C^{p \times q}$. Then the following operation

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix} \in C^{mp \times np} \quad (1)$$

is defined as the Kronecker product of A and B .

For the Kronecker product, the following properties are satisfied.

- P1. Let $A \in C^{m \times n}$. Then $0_{p \times q} \otimes A = A \otimes 0_{p \times q} = 0_{mp \times nq} \in C^{mp \times nq}$.
P2. Let λ be a given constant, $A \in C^{m \times n}$, $B \in C^{p \times q}$. Then we have

$$(\lambda A) \otimes B = A \otimes (\lambda B) = \lambda(A \otimes B).$$

- P3. Given matrices A , B , and C , where A and B have the same order. Then

$$\begin{aligned} (A + B) \otimes C &= A \otimes C + B \otimes C \\ C \otimes (A + B) &= C \otimes A + C \otimes B. \end{aligned}$$

- P4. Let $A \in C^{m \times n}$, $B \in C^{p \times q}$, and $C \in C^{r \times s}$. Then we have

$$(A \otimes B) \otimes C = A \otimes (B \otimes C).$$

- P5. Let $A \in C^{m \times n}$, $B \in C^{p \times q}$, $C \in C^{r \times s}$, and $D \in C^{q \times l}$. Then we have

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

P6. Let $A, B \in C^{n \times n}$ be invertible matrices. Then

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

P7. Let $A \in C^{m \times n}$ and $B \in C^{p \times q}$. Then

$$(A \otimes B)^T = A^T \otimes B^T \quad (A \otimes B)^* = A^* \otimes B^*.$$

Given the square matrix $A \in C^{n \times n}$, if the eigenvalues are all in the left half plane, that is, the real part of the eigenvalues is strictly less than zero, then it is called a Hurwitz matrix. If a matrix A contains at least one eigenvalue with a positive real part, then it is called an unstable matrix. The multiplicity of eigenvalues $\lambda(A)$ of A is defined as: the algebraic multiplicity of A is the multiplicity of eigenvalues of $\det(\lambda I_n - A) = 0$ and the geometric multiplicity of A is the subspace dimension of eigenvalues corresponding to $\lambda(A)$. Generally speaking, the algebraic multiplicity of matrix is greater than or equal to its geometric multiplicity. For symmetric matrix, the algebraic multiplicity of eigenvalue is equal to its geometric multiplicity.

B. Preliminary Theorems

Let A and B be n -order and m -order square matrices, and $C \in C^{n \times m}$. For any given matrix $R = [r_{ij}] \in C^{p \times q}$, define

$$(VecR)^T = [r_{11} r_{12} \cdots r_{1q} r_{21} \cdots r_{2q} \cdots r_{p1} \cdots r_{pq}]$$

then for matrix equation

$$AX - XB = C \tag{2}$$

we have the following theorem.

Theorem 1. *The necessary and sufficient condition for the solution of the equation (2) is*

$$\text{rank}[WVecC] = \text{rank}W \tag{3}$$

where $W = A \otimes I_m - I_n \otimes B^T$.

Theorem 2. *The necessary and sufficient condition for the existence and uniqueness of the solution of equation (2) is that there is no common eigenvalue between A and B , and the solution is $VecX = [A \otimes I_m - I_n \otimes B^T]^{-1} \cdot VecC$.*

Theorem 3. *For any given matrix $A = [a_{ij}] \in C^{n \times n}$, there is a unitary matrix U , such that $U^*AU = T$, where T is a upper triangular matrix. The matrix T is called the Schur upper triangle standard form, and can be divided into $T = A + M$. The diagonal matrix A is $A = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, where λ_i is the eigenvalue of matrix A , and M is a strict upper triangular matrix.*

3 Main Results

The first-order multi-agent system has been widely studied and a lot of results have been obtained. However, most of the current research on the consensus problem is focused on the multi-agent system with first-order and second-order integrated dynamics model, and the research on the multi-agent system with general linear dynamics model is relatively less. In this section, the design of distributed consensus controller is given for general linear dynamics model.

A. Description and Classification of the Problems

We consider a multi-agent system composed of N agents. The information is transmitted in one direction between agents, and the corresponding topology is represented by directed graph G . Each agent has the following dynamic characteristics

$$\begin{aligned}\dot{x}_i &= Ax_i + Bu_i \\ y_i &= Cx_i\end{aligned}\tag{4}$$

where $x_i \in R^n$ is the state of the i th agent, $u_i \in R^m$ is the input of the i th agent, and $y_i \in R^r$ is the output of the i th agent. $A \in C^{n \times n}$, $B \in C^{n \times m}$, and $C \in C^{r \times n}$ are the state matrix, input matrix and output matrix of the i th agent, respectively.

In the following, we give the definition of consensus.

Definition 2. For a multi-agent system with linear time-invariant dynamics (4), a set of control variables $\mu = \{u_i, i = 1, 2, \dots, N\}$ is called a distributed consensus protocol iff the design of $u_i(t)$ only depends on the information of the i th agent and its neighbor agent N_i , and μ makes the state of each agent converge to the same point. That is, there is x^* such that $\lim_{t \rightarrow \infty} x_i(t) = x^*$ for any initial point $x_i(0)$, $i = 1, 2, \dots, N$.

As the final consensus state, there are two cases for x^* : zero vector and general constant vector. Zero vector can be used as a special form of general constant vector, but considering the practical application of multi-agent systems often put forward aggregation requirements, for example, let the robot set at a certain point. This kind of problem can be transformed into multi-agent state uniformly converging to zero vector. Considering the independent research value of zero vector situation, this section distinguishes it from general constant value vector as an independent problem. They are called the zero vector consensus problem and the general constant vector consensus problem respectively.

B. Zero Vector Consensus Problem

Consider the following consensus protocol

$$\begin{aligned}\dot{z}_i &= F_i z_i + \sum_{j \in N_i} G_{ij} y_i \\ u_i &= H_i z_i + \sum_{j \in N_i} K_{ij} y_i.\end{aligned}\tag{5}$$

Definition 3. Let $F = [f_{ij}]$ be a $0 - 1$ matrix, where the elements of the matrix are determined as

$$f_{ij} = \begin{cases} 1, & y_j \in N_i; \\ 0, & y_j \notin N_i. \end{cases}$$

For the given matrix F , define a class of feedback matrix

$$\overline{\Pi}_F := \left\{ \Pi^F = \text{block}\{\Pi_{ij} \in R^{m_i \times r_i}\}; \Pi_{ij} = 0 \text{ iff } f_{ij} = 0 \right\}.$$

Then the closed-loop system can be written as follows

$$\begin{pmatrix} \dot{x} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} I_N \otimes A + (I_N \otimes B)K^F(I_N \otimes C) & (I_N \otimes B)H^N \\ G^F(I_N \otimes C) & F^N \end{pmatrix} \quad (6)$$

where $K^F, G^F \in \overline{\Pi}_F$, $F^N = \text{diag}\{F_1, \dots, F_N\}$, and $H^N = \text{diag}\{H_1, \dots, H_N\}$.

The necessary and sufficient conditions for system stabilization are obtained in the study of dynamic decentralized control in large-scale system theory. The corresponding results can be used to analyze multi-agent systems (1).

Definition 4. For a N -channel control system

$$\begin{cases} \dot{x} = Ax + \sum_{i=1}^N B_i u_i = Ax + B^N u \\ y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} C_1 x \\ \vdots \\ C_N x \end{pmatrix} = C^N x \end{cases}$$

we define the fixed-mode as

$$\Lambda(A, B^N, C^N, F) = \cap_{K^F \in \overline{\Pi}_F} \sigma(A + B^N K^F C^N).$$

Lemma 1. The necessary and sufficient condition for the asymptotically stable problem of system (1) is that there exists matrix F such that

$$\Lambda(A, B^N, C^N, F) \subset C^-.$$

With this lemma, we can get the following theorems.

Theorem 4. The necessary and sufficient condition for a multi-agent system to achieve zero vector consistent state under a fixed topology G is

$$\Lambda(I_N \otimes A, I_N \otimes B, I_N \otimes C, F) \subset C^-.$$

The problem of whether the multi-agent system (1) can reach the zero vector consensus state under the distributed control protocol (5) is reduced to the fixed mode problem under a feedback matrix.

C. General Constant Vector Consensus Problem

For general linear dynamic multi-agent system (1), the following state feedback protocol is adopted

$$u_i = K_1 x_i + \sum_{j \in N_i} a_{ij} K_2 (x_j - x_i). \quad (7)$$

Let $x = [x_1^T, \dots, x_n^T]^T$. Then the system can be written in the following compact form

$$\dot{x} = (I_N \otimes (A + BK_1) - L \otimes BK_2)x. \quad (8)$$

Lemma 2. *The consensus problem of system $\dot{x} = Ax$ can be achieved iff the following conditions hold*

- C1. $R(A) = R(A^2)$;
- C2. The eigenvalue of A is 0 or strictly has a negative real part;
- C3. If 0 is an eigenvalue of A , then for any $\xi \in N(A)$, there exists $b \in R^m$ such that $\xi = 1 \otimes b$.

Lemma 3. *The consensus problem of system $\dot{x} = Ax$ can be achieved iff*

$$\dim(N(A)) = \dim(N(A^2)) = \dim(N[C_{11}^T, \dots, C_{nn}^T]^T)$$

and the eigenvalue of A is 0 or strictly has a negative real part.

According to the above two lemmas, the necessary and sufficient conditions for system (8) to reach consensus are given below.

Theorem 5. *Assume that the directed graph of the multi-agent system (1) has spanning tree. Let the eigenvalues of L be $\lambda_1 = 0, \lambda_2, \dots, \lambda_n$. Then the consensus problem can be achieved iff*

- (1) $R(A + BK_1) = R((A + BK_1)^2)$;
- (2) the eigenvalue of $A + BK_1$ is 0 or strictly has a negative real part;
- (3) the matrix $A + BK_1 - \lambda_i BK_2$ is Hurwitz.

Proof. Because the system digraph contains a spanning tree, 0 is a eigenvalue of matrix L , and all other eigenvalues have positive real parts. According to Schur theorem, there is matrix $U \in C^{n \times n}$ such that

$$U^* LU = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

It follows that

$$\begin{aligned} & (U^* \otimes I_m)(I_n \otimes (A + BK_1) - L \otimes BK_2)(U \otimes I_m) \\ &= \begin{pmatrix} A + BK_1 & * & \cdots & * \\ 0 & A + BK_1 - \lambda_2 BK_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A + BK_1 - \lambda_n BK_2 \end{pmatrix}. \end{aligned}$$

Let $P = I_n \otimes (A + BK_1) - L \otimes BK_2$ and $Q = (U^* \otimes I_m)P(U \otimes I_m)$. From Lemma 3, we can see that the consensus problem can be achieved iff $\dim(N(P)) = \dim(N(P^2)) = \dim(N(A + BK_1))$ and the eigenvalue of P is 0 or strictly has a negative real part.

\Rightarrow Suppose that the consensus problem can be achieved. Since $\dim(N(P)) = \dim(N(Q))$ and $\dim(N(P)) = \dim(N(A + BK_1))$, we have $A + BK_1 - \lambda_i BK_2$ is invertible. From $\dim(N(Q)) = \dim(N(P)) = \dim(N(P^2)) = \dim(N(Q^2))$, we have $\dim(N(A + BK_1)) = \dim(N(A + BK_1)^2)$. Therefore, $R(N(A + BK_1)) = R(A(A + BK_1)^2)$.

\Leftarrow If terms (1), (2), and (3) hold. Then we have $\dim(N(Q)) = \dim(N(Q^2)) = \dim(N(A + BK_1))$, the eigenvalue of Q is 0 or strictly has a negative real part. Therefore, the consensus problem can be achieved.

4 Conclusions

In this paper, the consensus problem of multi-agent system is divided into two cases. Different control protocol forms are given, and the necessary and sufficient conditions for the existence of consensus controller are analyzed. In fact, the analysis of zero state consistency problem can be included in the analysis of general state consensus problem, but considering its particularity and wide application value, this paper studies it as an independent problem. In the analysis of zero state consensus, the most general form of control protocol to make the minimum restrictions on the control object is used. In the study of general state consensus, a special control protocol is adopted, which limits the variables. The analysis of this paper is aimed at the multi-agent system with definite topology, and the corresponding analysis will be made for the variable topology in the future.

References

- Jadbabaie, A., Lin, J., Morse, A.S.: Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Autom. Control* **48**(6), 988–1001 (2003)
- Xiao, F., Wang, L.: Consensus problems for high-dimensional multi-agent systems. *IET Control Theory Appl.* **1**(3), 830–837 (2007)
- Li, Z., Jia, Y.: Algebraic criteria for consensus problems of continuous-time networked systems. *Int. J. Control* **41**(4), 643–658 (2009)
- Wang, X., Hong, Y.: Finite-time consensus for multi-agent networks with second-order agent dynamics. In: Proceedings of IFAC World Congress, Seoul, Korea, 6–11 July, pp. 15185–15190 (2008)
- Yang, Q., Zhou, F., Chen, J., Li, X., Fang, H.: Distributed tracking for multiple Lagrangian systems using only position measurements. In: The International Federation of Automatic Control, Cape Town, South Africa, 24–29 August, pp. 287–292 (2014)
- Lin, J., Morse, A., Anderson, B.: The multi-agent rendezvous problem. In: Proceedings of the 42nd Conference on Decision and Control, Hawaii, USA, 9–12 December, pp. 1508–1513 (2003)
- Chong, C.Y., Kumar, S.P.: Sensor networks: evolution, opportunities, and challenges. *Proc. IEEE* **91**(8), 1247–1256 (2003)



A New Method of Power Quality Cloud Evaluation Considering Dynamic Characteristics

Cheng Guo^{1(✉)} and Yan-chao Yin²

¹ Yunnan Electrical Power Research Institute, Kunming 650217, China
gc325@126.com

² Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming 650500, China
yinyc@163.com

Abstract. Aiming at the fuzzy and random problems in the process of power quality evaluation, a multi-dimensional and multi rule dynamic weighted power quality cloud evaluation method is proposed on the basis of considering the dynamic characteristics of power quality evaluation indexes. Firstly, we construct a comprehensive evaluation system of power quality, and also establish a reverse cloud model of each qualitative evaluation index of power quality. Then, the multi-dimensional and multi rule dynamic cloud generator of power quality evaluation is constructed by combining different antecedent cloud and consequent cloud with dynamic weighting of evaluation index, and a reasoning framework and algorithm are designed to analyze the uncertain mapping relationship between each evaluation index and quality level in the value domain. The distribution of space and certainty space completes the mapping transformation from quantitative input to dynamic reasoning, and then to quantitative output. The feasibility and effectiveness of the new comprehensive evaluation method for power quality are proved by the comparative analysis of examples, which provide a new idea and method for the efficient evaluation of power quality.

Keywords: Power quality · Spatiotemporal characteristics · Dynamic weighting · Cloud evaluation

1 Introduction

With the rapid development of national economy and science and technology, and the further improvement of high-power electronic technology and automatic control technology, the concept and specific requirements of power quality for users have changed fundamentally. Due to the power quality has an important impact on the market operation of power system, in the era of moving towards smart grid, power quality evaluation becomes particularly important for enterprises and individual consumers [1].

In recent years, scholars have made many explorations on the comprehensive evaluation of power quality. For these research results, neither the evaluation system, index weight calculation, nor the evaluation method consider the process information of the evaluation object from the perspective of time series dynamics, and then provide objective and reasonable evaluation conclusions. For example, methods based on fuzzy mathematics theory [2], methods based on probability statistics principle [3], evaluation methods based on artificial intelligence algorithm [4], methods based on graph evaluation [5], evaluation methods based on matter-element analysis [6], etc., but these methods can not reflect the development process of evaluation objects, and can not objectively and reasonably evaluate the power quality of objects with similar level. However, a small amount of researches about the dynamic comprehensive evaluation index weight and the comprehensive weighting method [7] accelerate the establishment of power quality dynamic evaluation system such as the second weighting method of information aggregation [8], which is to compare the data in different periods, is not enough to compare the evaluation values in different development stages, nor to evaluate the problem of large amount of discrete information.

In this paper, aiming at the uncertainty of power quality evaluation, and considering the dynamic process information of the evaluation object. The cloud model theory is first introduced, where the reverse cloud is used to cloud the evaluation indexes. The quantitative index information is transformed into the qualitative concept, and a multi-dimensional and multi rule dynamic cloud generator considering the impact of the evaluation process information is established. The dynamic weighting of evaluation indexes are moved by the time sequence constraint and the method of cyclic updating cloud model is used to enhance the dynamic evaluation ability of the cloud generator to realize the qualitative and quantitative transformation of power quality evaluation, which objectively reflects the dynamic, complexity and coupling of power quality evaluation process.

2 Cloud Reasoning Considering Temporal Constraints

2.1 Basic Concepts and Digital Characteristics of Cloud

Let U be a domain $U = \{x\}$, and T is the linguistic value associated with U . The membership degree $C_T(x)$ (or compatibility degree of X and T) of the element x in U for the qualitative concept expressed by T is a random number with stable tendency. The distribution of membership degree in the domain can be considered as membership cloud, which is called cloud for short [9].

The value of $C_T(x)$ is in $[0, 1]$, and cloud is the mapping from domain U to interval $[0, 1]$, which is $C_T(x) : U \rightarrow [0, 1], \forall x \in U, x \rightarrow C_T(x)$. The cloud model describes the fuzziness and randomness of qualitative concepts in an unified quantitative way through three digital characteristics of expectation (Ex), entropy (En) and hyper entropy (He).

2.2 Dynamic Weighting of Evaluation Indexes Considering Time Series Constraints

Based on the objective weighting method, this paper introduces the relative closeness of the evaluation cloud, reflects the relative importance of the index in the evaluation process by the distance between any two evaluation index clouds at different evaluation times, and gives the corresponding weight of each evaluation index in the way of dynamic weighting.

Assuming that there are two evaluation index clouds $C_{t1}(Ex_{t1}, En_{t1}, He_{t1})$, $C_{t2}(Ex_{t2}, En_{t2}, He_{t2})$ in the evaluation domain space at any time, the relative closeness of the two index clouds can be defined as $R_{t1,t2} = |Ex_{t1} - Ex_{t2}|$, and each one-dimensional evaluation index of the two evaluation index clouds calculates its corresponding cloud digital eigenvalue through the reverse cloud $(Ex_{t1i}, En_{t1i}, He_{t1i})$, $(Ex_{t2i}, En_{t2i}, He_{t2i})$, $t = 1, 2, \dots, n$ and $i = 1, 2, \dots, m$.

Then the dynamic weight of each evaluation index is:

$$\lambda_i = \frac{R_{t1i,t2i}}{\sum_{i=1}^m R_{t1i,t2i}}, \quad t = 1, 2, \dots, n \quad (1)$$

If the weight vector obtained by entropy weight method is ω_{O1} , and the weight vector obtained by mean square deviation method is ω_{O2} , the objective weight vector obtained by combining geometric average method will be:

$$\omega_{Oi} = \frac{\sqrt{\omega_{O1i}\omega_{O2i}}}{\sum_{i=1}^m \omega_{O1i}\omega_{O2i}} \quad (2)$$

Using the method of additive integration, the comprehensive weight at any time is obtained as follows:

$$\xi_t = a\lambda_t + b\omega_O \quad (3)$$

$$a = \frac{1}{m-1}[(\delta_1 + 2\delta_2 + \dots + m\delta_m) - \frac{m+1}{m}] \quad (4)$$

$$b = 1 - a \quad (5)$$

The δ_i is the corresponding component of the dynamic weight vector after being arranged in ascending order; and the m is the number of evaluation indicators.

2.3 Uncertain Regular Reasoning of Cloud Assessment

The uncertain rule reasoning of cloud mapping for power quality comprehensive evaluation obtains all evaluation indexes of power quality according to the power quality comprehensive evaluation system and gives them objective and reasonable weights. It constructs the antecedent (X condition) cloud model of power quality evaluation system through cloud processing and the subsequent (Y condition) cloud model of power quality evaluation system through cloud

processing of power quality evaluation rules and results. The constructed front and rear part cloud model is shown in Eqs. (6)–(9).

The n -dimensional X conditional cloud model is

$$(P_{x_{1_i}}, P_{x_{2_i}}, \dots, P_{x_{n_i}}) = R_n(E_{n_{1i}}, E_{n_{2i}}, \dots, E_{n_{ni}}, H_{e_{1i}}, H_{e_{2i}}, \dots, H_{e_{ni}}) \quad (6)$$

$$\begin{aligned} \tilde{\mu} &= (e^{\xi_1} \mu_1) \times (e^{\xi_2} \mu_2) \times \dots \times (e^{\xi_n} \mu_n) \\ &= e^{-\left(\frac{\xi_{t1}(x_1 - Ex_{b11})^2}{2E'n'_{b11}^2} + \frac{\xi_{t2}(x_2 - Ex_{b21})^2}{2E'n'_{b21}^2} + \dots + \frac{\xi_{tn}(x_n - Ex_{bn1})^2}{2E'n'_{bn1}^2}\right)} \end{aligned} \quad (7)$$

The n -dimensional Y conditional cloud model is

$$(P_{y_{1_i}}, P_{y_{2_i}}, \dots, P_{y_{n_i}}) = R_n(E_{n_{y1i}}, E_{n_{y2i}}, \dots, E_{n_{yni}}, H_{e_{y1i}}, H_{e_{y2i}}, \dots, H_{e_{yni}}) \quad (8)$$

$$y_i = nE_y \pm \sqrt{-2 \ln(\mu)}(P_{y_{1_i}} + P_{y_{2_i}} + \dots + P_{y_{n_i}}) \quad (9)$$

where Ex_{ni} , $E_{n_{xi}}$, $H_{e_{xi}}$ are the digital features of the n -dimensional evaluation cloud model of the rule antecedent in (6) and (7), $E_{y_{ni}}$, $E_{n_{yi}}$, $H_{e_{yi}}$ are the digital features of the n -dimensional evaluation cloud model of the rule antecedent in (8) and (9). The cloud model generator based on multi-dimensional multi-rule power quality evaluation can be constructed using multiple n -dimensional X condition generators and n one-dimensional Y condition generators. The reasoning structure is as follows:

$$\begin{aligned} &\text{IF } A_{11}, A_{21}, \dots, \text{ and } A_{n1}, \text{ THEN } C_{11} \\ &\text{IF } A_{12}, A_{22}, \dots, \text{ and } A_{n2}, \text{ THEN } C_{12} \\ &\quad \dots \quad \dots \\ &\text{IF } A_{1m}, A_{2m}, \dots, \text{ and } A_{nm}, \text{ THEN } C_{1m} \end{aligned}$$

Therefore, the reasoning of quality evaluation rules with different structures can be realized by combining the X conditional cloud model with the Y conditional cloud model.

3 Comprehensive Evaluation of Power Quality Based on Multi Dimension and Multi Rule Cloud Model

3.1 Power Quality Evaluation Index Cloud Model

In this paper, nine power quality evaluation indexes are selected to construct the cloud model of pre-evaluation as shown in Table 1. The evaluation results of power quality are divided into five grades, which are excellent, good, medium, qualified and unqualified. The values of 1–5 are given to correspond to the 5 grades. The grade limit values of each index of power quality are shown in Table 1 [8].

Since the change range of evaluation index is bilateral constraint space $[C_{min}, C_{max}]$ and its spatial boundary is stochastic and fuzzy, the cloud processing of data in Table 1 can be based on the $3En$ principle of cloud model [10], and the parameters of cloud model can be determined by the following equation:

$$\begin{cases} E_x = (C_{min} + C_{max})/2 \\ E_n = (C_{max} - C_{min})/2 \\ H_e = k \end{cases} \quad (10)$$

where k is a constant and can be adjusted according to the needs of the evaluation model.

Table 1. Grade limits of power quality evaluation indexes

Grade	Evaluation index limit								
	$b_1/\%$	b_2/Hz	$b_3/\%$	$b_4/\%$	$b_5/\%$	b_6	$b_7/\%$	b_8	b_9
1	≤ 1.2	≤ 0.05	≤ 1	≤ 0.5	≤ 0.2	≥ 0.9	≤ 0.5	≥ 0.95	≥ 0.9
2	≤ 3.0	≤ 0.10	≤ 2	≤ 1.0	≤ 0.5	≥ 0.8	≤ 1.0	≥ 0.85	≥ 0.8
3	≤ 4.5	≤ 0.15	≤ 3	≤ 1.5	≤ 0.8	≥ 0.5	≤ 1.5	≥ 0.80	≥ 0.7
4	≤ 7.0	≤ 0.20	≤ 5	≤ 2.0	≤ 1.0	≥ 0.1	≤ 2.0	≥ 0.70	≥ 0.6
5	≥ 7.0	≥ 0.20	≥ 5	≥ 2.0	≥ 1.0	≤ 0.1	≥ 2.0	≤ 0.70	≤ 0.6

Note: the above classification is based on 380 V voltage. b_1 -Voltage deviation index, b_2 -frequency deviation index, b_3 -harmonic voltage content index, b_4 -voltage fluctuation index, b_5 -voltage flicker index, b_6 -voltage transient index, b_7 -three-phase imbalance index, b_8 -power supply reliability index, b_9 -service index.

After the cloudization of each index, the digital characteristics of each index cloud model of the rule antecedent are shown in Table 2 [8].

The assessment results of power quality are divided into five grades: excellent, good, medium, qualified and unqualified. The corresponding grade cloud model is shown in Table 3.

3.2 Multi Dimension and Multi Rule Qualitative Reasoning Algorithm of Power Quality Evaluation

According to Table 2, we can obtain the reasoning rules of power quality comprehensive evaluation described by qualitative language. Given the power quality evaluation index to evaluate the power quality of an observation point, the qualitative reasoning process based on multi-dimensional multi rule cloud model is as follows:

Input: Digital eigenvalues of n power quality evaluation indexes $(Ex_{b11}, En_{b11}, He_{b11}), \dots, (Ex_{bn1}, En_{bn1}, He_{bn1}) \dots (Ex_{bij}, En_{bij}, He_{bij}), \dots, (Ex_{bnm}, En_{bnm}, He_{bnm})$ n digital eigenvalues of power quality evaluation rule knowledge $(Ex_{U11}, En_{U11}, He_{U11}), \dots, (Ex_{Un1}, En_{Un1}, He_{Un1}) \dots (Ex_{Uij}, En_{Uij}, He_{Uij}), \dots, (Ex_{Unm}, En_{Unm}, He_{Unm})$, given the input $x_i = \mu_i, i = 1, 2, \dots, n$, the number of cloud droplets generated n ;

Table 2. Grade limit cloud model of power quality evaluation indexes

Finger mark	Hierarchical boundary cloud				
	Level 1	Level 2	Level 3	Level 4	Level 5
b_1	(0.6050, 0.2017, 0.0800)	(2.1450, 0.3193, 0.0800)	(3.7050 0.2683, 0.0800)	(5.7950, 0.4350, 0.0800)	(8.5000, 0.5333, 0.0800)
b_2	(0.0255, 0.0085, 0.0800)	(0.0795, 0.0102, 0.0800)	(0.1205 0.0102, 0.0800)	(0.1795, 0.0102, 0.0800)	(0.3000, 0.0367, 0.0800)
b_3	(0.5500, 0.1833, 0.0800)	(1.5000, 0.2000, 0.0800)	(2.5000 0.2000, 0.0800)	(4.0000, 0.3367, 0.0800)	(6.5000, 0.5333, 0.0800)
b_4	(0.2550, 0.0850, 0.0800)	(0.7950, 0.1017, 0.0800)	(1.2050 0.1017, 0.0800)	(1.7950, 0.1017, 0.0800)	(3.0000, 0.3667, 0.0800)
b_5	(0.1050, 0.0350, 0.0800)	(0.3500, 0.0200, 0.0800)	(0.6500 0.0533, 0.0800)	(0.9450, 0.0517, 0.0800)	(1.5000, 0.2000, 0.0800)
b_6	(0.9450, 0.0183, 0.0800)	(0.8500, 0.0200, 0.0800)	(0.6500 0.0533, 0.0800)	(0.3000, 0.0700, 0.0800)	(0.0550, 0.0183, 0.0800)
b_7	(0.2550, 0.0850, 0.0800)	(0.7950, 0.1017, 0.0800)	(1.2050 0.1017, 0.0800)	(1.7950, 0.1017, 0.0800)	(3.0000, 0.3667, 0.0800)
b_8	(0.9745, 0.0085, 0.0800)	(0.9000, 0.0170, 0.0800)	(0.8205 0.0102, 0.0800)	(0.7500, 0.0200, 0.0800)	(0.3550, 0.1183, 0.0800)
b_9	(0.9450, 0.0183, 0.0800)	(0.8500, 0.0200, 0.0800)	(0.7500 0.0200, 0.0800)	(0.6500, 0.0200, 0.0800)	(0.3050, 0.1017, 0.0800)

Table 3. Power quality level cloud model

Number features	Power quality level				
	1	2	3	4	5
Ex	1	2	3	4	5
En	0.10	0.15	0.20	0.20	0.10
He	0.02	0.02	0.02	0.02	0.02

Output: power quality evaluation grade value E_C . According to the number of activation rules, randomly output single or multiple values.

Algorithm: multidimensional multi rule cloud model qualitative reasoning algorithm.

Step 1: judge how many rules are activated for a given input x_i .

Step 2: if a rule is activated.

Step 2.1: generate one-dimensional normal random number En_{bni}' with expected value En_{bni} and standard deviation He_{bni} , use the evaluation index dynamic weighting method considering time sequence constraints to determine the weight coefficient $\xi_{t1}, \xi_{t2} \dots \xi_{tn}$ of each index of power quality at any time, and calculate the membership degree $\tilde{\mu}$ after dynamic weighting:

$$\begin{aligned}\tilde{\mu} &= (e^{\xi_1} \mu_1) \times (e^{\xi_2} \mu_2) \times \dots \times (e^{\xi_n} \mu_n) \\ &= e^{-\left(\frac{\xi_{t1}(x_1 - Ex_{b11})^2}{2En_{b11}^{'2}} + \frac{\xi_{t2}(x_2 - Ex_{b21})^2}{2En_{b21}^{'2}} + \dots + \frac{\xi_{tn}(x_n - Ex_{bn1})^2}{2En_{bn1}^{'2}}\right)}\end{aligned}\quad (11)$$

Step 2.2: according to the acquired evaluation knowledge rule, we can generate the n -dimension normal value as expected value En_{U1i} and standard deviation He_{U1i} , and calculate y_i :

$$y_i = E_{U1i} \pm \sqrt{-2 \ln(\tilde{\mu})} En_{U1i}' \quad (12)$$

Step 2.3: make $(y_i, \tilde{\mu})$ a cloud drop.

Step 2.4: return to step 2.1, cycle several times, and finally output with the average of the expected values of all cloud drops.

Step 3.1: if two rules are activated, each single rule will be repeated step 2.1 to obtain the weighted membership degree $\tilde{\mu}$ of the evaluation index of the activated two single rules.

Step 3.2: take out μ_1 and μ_2 from $\tilde{\mu}_i$. The two single rules randomly generate a one-dimensional normal random value En_{U111} and En_{U112} with expectation En_{U11} and variance He_{U11} according to the (En_{U11}, He_{U11}) of a given subsequent piece. According to the following equations, the values of the two y_1, μ_2 under the condition of μ_1, En_{U111} and the values of two y_2 under the condition of En_{U112} are obtained by inverse calculation:

$$\mu_1 = e^{-\frac{(y_1 - Ex_{U111})^2}{2En_{U111}^2}}, \quad \mu_2 = e^{-\frac{(y_2 - Ex_{U112})^2}{2En_{U112}^2}}$$

Step 3.3: select the outermost two cloud drops (y_1, μ_1) and (y_2, μ_2) , and build a virtual concept. The digital characteristics of the virtual cloud are (Ex, En, He) , and the temporary super entropy is $He = 0$. The expectation of the virtual cloud can be obtained by solving the variance group through the geometric method.

Step 4: if multiple rules are activated, each single rule will output multiple cloud droplets according to steps 2.1 to 2.3, then return to step 2.1, cycle steps 2.1 to 2.3 several times, and finally output the average value of all expected cloud droplets.

Step 5: repeat steps 1–4 until the number of output values Y meet the requirements of power quality evaluation.

4 Case Application

To compare with other evaluation models, the measured data of the same observation points as that in reference [8] are used in this paper, as shown in Table 5, and the voltage level considered is 380 V. Take the first row of data in Table 4 as an example. If x_1 equals 3.212, x_2 equals 0.0922, x_3 equals 1.72, x_4 equals 1.33, x_5 equals 0.473, x_6 equals 0.7963, x_7 equals 0.83, x_8 equals 0.833, x_9 equals 0.83, and we input the multidimensional multi-rule cloud model qualitative inference generator, it will activate rule 2 with the five qualitative rules and the activation intensity μ will be 0.4514. y equals 2.0008 by step 2.2 in Sect. 2.2. Under this condition, the power quality level of observation point 1 is Level 2, and the corresponding language description value is good.

Take the third row of data in Table 5 as an example, and input it into the multi-dimensional multi rule cloud model qualitative inference generator, it will activate qualitative rules 3 and 4, and the maximum value μ_1 of the activation intensity will be 0.6875 and the second largest will be 0.3080. Corresponding to two cloud drops generated in the subsequent cloud generator, take the outermost two cloud drops A(2.8780, 0.6875) and D(4.2319, 0.3080), and calculate it according to step 3.3 of the above power quality evaluation algorithm. So it is obtained that Ex equals 3.0063, the power quality level of observation point 3 is level 3, and the corresponding language description value is general under this condition. It can be seen that the comprehensive evaluation results of power quality obtained by applying the methods proposed in this paper are more accurate compared with the extension cloud method and other methods, which shows that the comprehensive evaluation method of power quality proposed in this paper based on multi-dimensional and multi rule cloud model is feasible and effective.

Table 4. Measured data of observed points

T observation point	Hierarchical boundary cloud								
	$b_1/\%$	b_2/Hz	$b_3/\%$	$b_4/\%$	$b_5/\%$	b_6	$b_7/\%$	b_8	b_9
1	3.212	0.0922	1.72	1.33	0.473	0.7963	0.83	0.833	0.832
2	6.680	0.1562	4.28	1.53	0.847	0.1589	1.36	0.762	0.713
3	4.350	0.1180	2.67	1.95	0.634	0.5156	1.35	0.796	0.864
4	5.330	0.1787	3.36	1.37	0.826	0.5856	1.74	0.740	0.684
5	4.220	0.1892	4.57	1.58	0.828	0.4863	1.83	0.764	0.783

Table 5. Comparison of the evaluation results

Observation point	Results of literature [2]	Results of literature [3]	Results of literature [5]	Results of literature [6]	Results of literature [9]	Results of this paper
1	Level 2	Level 2	Level 2	Level 2.05	Level 2	The calculated value is 2.0008, and the evaluation is level 2
2	Level 4	Level 4	Level 4	Level 4.00	Level 4	The calculated value is 3.9388, and the evaluation is level 4
3	Level 3	Level 3	Level 3	Level 3.00	Level 3	The calculated value is 3.0063, and the evaluation is level 3
4	Level 4	Level 4	Level 4	Level 3.46	Level 4	The calculated value is 3.9520, and the evaluation is level 4
5	Level 4	Level 4	Level 4	Level 4.00	Level 4	The calculated value is 3.9824, and the evaluation is level 4

5 Conclusion

This paper studies the power quality comprehensive evaluation system, evaluation index and cloud reasoning evaluation method considering the dynamic characteristics of time and space. The multivariate combination of the former cloud and the latter cloud of inference rule is first realized by dynamically weighting the evaluation index. We design a multi-dimensional multi-rule dynamic cloud generator for power quality evaluation and give a reasoning framework and algorithm. The dynamic evaluation ability of cloud generator is continuously improved in the cycle updating of cloud model, and the transformation from quantitative to qualitative to quantitative is realized. It avoids the defect that the classification grade of power quality is expressed as clear boundary, and objectively reflects the dynamic, complex and coupling of the process of power quality evaluation.

References

1. Zheng, G., Du, X., Qi, Z.: Research and implementation of online location device for small current single-phase ground fault. *Power Syst. Prot. Control* **40**(8), 135–139 (2012)
2. Tu, C., Yang, Y., Xiao, F., Lan, Z., Li, Y.: Power quality control strategy of output side of main inverter in microgrid under nonlinear load. *J. Electr. Technol.* **33**(11), 2486–2495 (2018)
3. Lin, Z., Chen, Y., Cai, J., Li, T.: Influence and control measures of three-phase unbalanced operation of low-voltage distribution network. *J. Electric Power Sci. Technol.* **24**(3), 63–67 (2009)
4. Shu, S., Luo, J., Wang, Q., Sun, B.: Design and implementation of portable power quality monitoring system based on DSP and arm. *Power Syst. Prot. Control* **38**(24), 185–189 (2010)
5. Wang, B., Pan, Z., Zhao, J., Shi, L.: Synchronization and device design of power quality monitoring data. *Power Syst. Autom.* **11**, 45–49 (2002)

6. Li, Q., et al.: Research on field verification technology of digital electric energy measurement system. *Electric Meas. Instrum.* **47**(10), 25–28 (2010)
7. Wu, L., Liu, H., Yue, C., Zhang, S., Wu, J.: Reactive compensation in field verification test of large current transformer. *High Volt. Technol.* **36**(6), 1560–1565 (2010)
8. Lu, Z.: Discussion on field verification of electric energy meter. *Electric Meas. Instrum.* **48**(1), 1–5 (2011)
9. Zhang, Q., Zuo, Q., He, G., Zhang, B.: Research on time synchronization test of intelligent substation network. *Power Syst. Prot. Control* **38**(21), 237–240 (2010)
10. Song, S., Shao, Y., Du, Y.: Summary of sampling method research. *Data Acquis. Process.* **31**(3), 452–463 (2016)



A Control Method of Fixed-Wing UAVs Under Close Formation

Yafei Zhao^(✉) and Yanjie Zhao

China Academy of Electronics and Information Technology, 100041 Beijing, China
zhaoafei001@163.com

Abstract. UAVs with a close formation will improve the flying efficient and save energy. However, when UAVs fly within 1.5 wingspan distance, the nonlinear coupling phenomenon of the aerodynamics, motion and control will be obvious and cannot be neglected. This paper introduces a nonlinear robust control method to solve multi-UAVs control problem under close formation. In the simulation, formation position tracking error is less than 5% wingspan, which is proved that close formation can be maintained effectively.

Keywords: Close formation · UAVs · Airflow filed · Nonlinear robust control

1 Introduction

When UAVs are flying in close formation, the interact of the airflow field will lead to the unsteady aerodynamic and motion parameters of UAVs and the nonlinear change of the aerodynamic force, which will lead to the obvious nonlinear coupling phenomenon of the aerodynamics, motion and control. The main characteristics are strong coupling, the interaction between the aerodynamic characteristics of UAVs. The aerodynamic and motion parameters of the aircraft change with time, which directly affects the aerodynamic characteristics and flight performance of UAVs. One solution of the problem is to convert the aerodynamic interference into the power which is beneficial to the formation flying, which is to reduce the flying resistance of follower drones. A natural phenomenon is that nearby birds can use this updraft to reduce flight resistance and save physical strength. At the same time, each bird in the formation can save physical strength by changing its position. Therefore, efficient use of aerodynamic interference between formations can prolong the group of birds.

In order to save fuel and extend cruise time, Altshuler Y et al. [1] Studied the method of adjusting the position of UAV in the front and rear end of the formation in order. Nadia H [2] further carried out research from the perspective of mission flexibility and safety. Through the graphic theory, they studied that when the target of the formation flying mission changes, or when the threat is encountered, the geometry of the UAV formation is changed in real time, so

as to achieve the multi-purpose and multi task of the formation networking. Li Xinhong [3] et al. Of the Chinese Academy of Sciences studied how to supplement the individual in the formation flight at any time or reconfigure the whole, in order to reduce the mission risk. Some researchers put forward a regional cooperation algorithm for the reasonable allocation of the pilot role in a pair of UAVs, and calculated the travel interval of each UAV as a pilot role [4-7]. The algorithm effectively solves the problem of dynamic adjustment of formation, and minimizes the conversion times of leader follower role, so that the fuel consumption of the whole formation and the single UAV is the least at the same time. In 2014, Xiamen University put forward a design idea of formation transformation by controlling formation spacing, which effectively realized the real-time transformation of formation [8]. Some researcher designed a distributed algorithm based on trajectory segmentation for the conversion of multi UAV vertical formation and circular formation, which realized the rapid adjustment of formation [9].

In conclusion, formation coordinated control is the precondition for UAV cluster to be able to fly in formation efficiently. The traditional coordinated control technology has many problems, such as low autonomy, poor coordination, weak anti-interference ability, etc., which can not effectively solve the contradictions of time, space and task level in the intensive formation, especially in the large-scale UAV intensive formation combat environment, it is difficult to achieve the UAV intensive formation combat in the complex battlefield environment. Therefore, it is urgent to explore a complete theory and method of UAV cluster aerodynamic coupling and cooperative control technology, maintain the advantages of UAV cooperative operation, avoid the conflicts and collisions between UAVs, improve the aerodynamic utilization characteristics and cluster mobility efficiency, and provide effective technical support for the realization of the overall combat effectiveness of UAV system.

2 Dynamical Equations

The framework is defined as follows.

In inertial framework, position vector is $\mathbf{P}_I = [x \ y \ z]^T$. When a drone is affected by another tail vortex in a close formation, its kinematic equation is

$$\dot{\mathbf{P}}_I = \underline{\mathbf{C}}_{IW} \mathbf{v}_a + \mathbf{w}_I \quad (1)$$

where $\mathbf{v}_a = [V \ 0 \ 0]^T$ is airspeed vector in wind coordinate framework (Fig. 1), $\mathbf{w}_I = [w_x \ w_y \ w_z]$ is vortex induced effective wake velocity vector in inertial framework, $\underline{\mathbf{C}}_{IW}$ is transfer matrix wind framework to inertial framework.

μ, γ, χ are tilt angle, track angle and heading angle of a drone respectively (Fig. 2). The kinematic equations are rewritten as

$$\dot{x} = \mathbf{V} \cos \gamma \cos \chi + w_x, \quad \dot{y} = \mathbf{V} \cos \gamma \sin \chi + w_y, \quad \dot{z} = -\mathbf{V} \sin \gamma + w_z \quad (2)$$

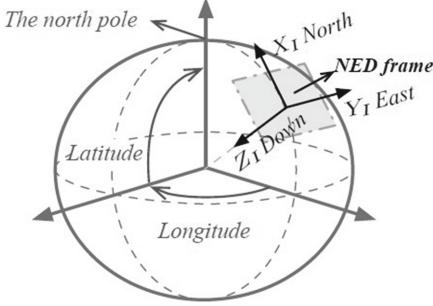


Fig. 1. North East Down coordinate system

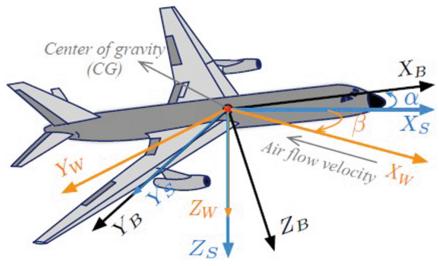


Fig. 2. Body and velocity coordinate systems

Translational dynamics equations are,

$$\begin{cases} \dot{V} = \frac{F_x + \dot{w}}{m} \\ \dot{\gamma} = -\frac{(F_y + \dot{w}_y) \sin \mu + (F_z + \dot{w}_z) \cos \mu}{mV} \\ \dot{\chi} = \frac{(F_y + \dot{w}_y) \cos \mu + (F_z + \dot{w}_z) \sin \mu}{mV \cos \gamma} \end{cases} \quad (3)$$

where F_x, F_y, F_z are three components of in airspeed framework.

The rotational dynamics equations are,

$$\begin{cases} \dot{p} = \frac{(I_y - I_z)I_z - I_{xz}^2}{I_x I_z - I_{xz}^2} pq + \frac{(I_x - I_y + I_z)I_{xz}}{I_x I_z - I_{xz}^2} pq + \frac{I_z}{I_x I_z - I_{xz}^2} (L + \Delta L) + \frac{I_z}{I_x I_z - I_{xz}^2} (N + \Delta N) \\ \dot{q} = \frac{(I_x I_z - I_{xz}^2)I_y}{I_x I_z - I_{xz}^2} pr + \frac{I_x}{(I_x I_z - I_{xz}^2)I_y} (p^2 - r^2) + \frac{1}{(I_x I_z - I_{xz}^2)I_y} (M + \Delta M) \\ \dot{r} = \frac{(I_x - I_y)I_x - I_{xz}^2}{I_x I_z - I_{xz}^2} pq - \frac{(I_x - I_y + I_z)I_x}{I_x I_z - I_{xz}^2} rq + \frac{I_{xz}}{I_x I_z - I_{xz}^2} (L + \Delta L) + \frac{I_x}{I_x I_z - I_{xz}^2} (N + \Delta N) \end{cases} \quad (4)$$

3 Nonlinear Robust Formation Control

In the leader-follower mode, when the leader drone is maneuvering, relative position in the inertial system will change with the attitude of the leader drone. It is time-varying. However, the best relative position change rate which is necessary for precise position control is not available. Therefore, it will be more challenging for follower drone to track their best position in close formation under leader drone maneuver, because the relative position is changing at an unknown rate. The second challenge comes from the complex nonlinear dynamics of the outer and inner loops. The kinematics of the fixed-wing aircraft has nonholonomic constraints, which will increase the difficulty of control design and analysis. If the response speed of inner loop control is not fast enough, the performance of outer loop formation controller cannot be guaranteed. Besides the induced force caused by vortex, the robustness to the moment caused by eddy current should also be considered in this part. In this part, the influence of induced vortex will

be considered in the control design, and the integrated design of aerodynamic and formation controller will be carried out.

A nonlinear robust controller based on disturbance observer is proposed in this research. In order to simplify the analysis and make it clearer, the nonlinear robust formation control proposed in this paper will be discussed in the close formation flight of two drones. However, it should be noted that by following the design procedures described in this section, the proposed control algorithm can be extended to three drones or more.

3.1 Problem Description

In close formation, the relative optimal position of the follower and the leader is constant in the airflow coordinate system of the leader. Let $\mathbf{I}_W = [l_x \ l_y \ 0]^T$ the relative optimal position of the leader in the airflow coordinate system, the range of l_x is $[-2b, -10b]$, l_y is approximated to $\pm 0.95b$. Let x_d , y_d and z_d are relative optimal position of close formation in inertial coordinate frame, thus

$$x_d = x_l + l_x, \quad y_d = y_l + l_y, \quad z_d = z_l + l_z \quad (5)$$

where x_l , y_l and z_l are components in inertial framework. If μ_l , γ_l and χ_l can be obtained, then $\mathbf{I}_l = \underline{\mathcal{C}}_{IW}(\mu_l, \gamma_l, \chi_l) \mathbf{l}_W$ can be also obtained. Derivation of (5) at time will derive following motional mode of the follower,

$$\dot{x}_d = V_l \cos \gamma_l \cos \chi_l + \dot{l}_x, \quad \dot{y}_d = V_l \cos \gamma_l \sin \chi_l + \dot{l}_y, \quad \dot{z}_d = -V_l \sin \gamma_l + \dot{l}_z \quad (6)$$

where V_l , γ_l and χ_l represent the follower's airspeed, track angle and heading angle.

Let $\mathcal{F}(t)$ fully smooth signal, thus command filter is

$$\begin{bmatrix} \dot{\mathcal{F}}_c \\ \ddot{\mathcal{F}}_c \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\zeta_F \omega_F \end{bmatrix} \begin{bmatrix} \mathcal{F}_c \\ \dot{\mathcal{F}}_c \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_F^2 \end{bmatrix} \mathcal{F}(t) \quad (7)$$

where \mathcal{F}_c and $\dot{\mathcal{F}}_c$ are the estimation of \mathcal{F}_t and $\dot{\mathcal{F}}_t$, $\omega_F > 0$ is natural frequency, $\zeta_F > 0$ is damping ratio. Initial conditions of (7) are assumed as $\mathcal{F}_c(0) = \mathcal{F}(0)$ and $\dot{\mathcal{F}}_c(0) = 0$.

The command signal are $\mathbf{I}_c = [l_{cx}, l_{cy}, l_{cz}]^T$ and $\mathbf{I}_c = [i_{cx}, i_{cy}, l_{cc}]$, the refer motional model of follower in close formation is described as

$$\dot{x}_r = V_r \cos \gamma_r \cos \chi_r, \quad \dot{y}_r = V_r \cos \gamma_r \sin \chi_r, \quad \dot{z}_r = -V_r \sin \gamma_r \quad (8)$$

where $x_r = x_l + l_{ce}$, $y_r = y_l + l_g$ and $z_r = z_l + l_{cz}$ are reference position signal, V_r , γ_r and χ_r are reference airspeed, track angle and heading angle, they are calculated by the following equations,

$$\begin{cases} V_r = \sqrt{(\dot{x}_l + l_{ce})^2 + (\dot{y}_l + l_{cy})^2 + (\dot{z}_l + l_{cz})^2} \\ \gamma_r = -\sin^{-1} \left(\frac{\dot{z}_l + l_{cz}}{V_r} \right) \\ \chi_r = \chi_l + \sin^{-1} \left(\frac{-l_{ce} \sin \chi_l + l_g \cos \chi_l}{V_r \cos \gamma_r} \right) \end{cases} \quad (9)$$

where $\dot{x}_l = V_l \cos \gamma_l \cos \chi_l$, $\dot{y}_l = V_l \cos \gamma_l \sin \chi_l$ and $\dot{z}_l = -V_l \sin \gamma_l$.

3.2 Nonlinear Robust Formation Controller Design

Formation controller is designed into inner attitude control and outer position control loops. The command force T_c desired inclination μ_d and desired attack angle α_d will be computed by outer position controller. The inner attitude controller will compute desired μ_d and α_d . By the way, sideslip angle β_f will be controlled to stabilized at zero by inner controller.

The Outer Formation Position Control. Drag D and lift L contain unknown elements and cannot be calculated accurately. Nominal items D and L are written as \bar{D} and \bar{L} can be calculated by aerodynamic data of the wind tunnel. Let $\bar{L} = \bar{L}_0 + \bar{L}_\alpha \alpha_f$, where \bar{L}_0 represents basic lift when $\alpha_f = 0$ and \bar{L}_α lift derivative to α_f , a uncertain model of feasible position control of outer formation is described as

$$\begin{cases} \dot{x}_f = V_f \cos \gamma_f \cos \chi_f + w_x \\ \dot{y}_f = V_f \cos \gamma_f \sin \chi_f + w_y \\ \dot{z}_f = -V_f \sin \gamma_f + w_z \\ \dot{V}_f = \frac{1}{m_f} (T \cos \alpha_f \cos \beta_f - \bar{D}) - g \sin \gamma_f + d_V \\ \dot{\gamma}_f = \frac{(\bar{L} + T \sin \alpha_f) \cos \mu_f}{m_f V_f} - \frac{g}{V_f} \cos \gamma_f + d_\gamma \\ \dot{\chi}_f = \frac{(\bar{L} + T \sin \alpha_f) \sin \mu_f}{m_f V_f \cos \gamma_f} + d_\chi \end{cases} \quad (10)$$

where d_V , d_γ and d_χ are disturbing items of system uncertainties and aerodynamic disturbance. We design the following nonlinear disturbing estimator

$$\begin{cases} \lambda_{wx} = -\frac{\lambda_{wx}}{T_{wx}} - \frac{1}{T_{wx}} \left(\frac{1}{T_{wx}} x_f + V_f \cos \gamma_f \cos \chi_f \right) \\ \dot{\lambda}_{wy} = -\frac{\lambda_{wy}}{T_{wy}} - \frac{1}{T_{wy}} \left(\frac{1}{T_{wy}} y_f + V_f \cos \gamma_f \sin \chi_f \right) \\ \lambda_{wz} = -\frac{\lambda_{wz}}{T_{wz}} - \frac{1}{T_{wz}} \left(\frac{1}{T_{wy}} z_f - V_f \sin \gamma_f \right) \\ \hat{w}_x = \lambda_{wx} + \frac{1}{T_{wx}} x_f \\ \hat{w}_y = \lambda_{wy} + \frac{1}{T_{wy}} y_f \\ \hat{w}_z = \lambda_{wz} + \frac{1}{T_{wz}} z_f \end{cases} \quad (11)$$

where λ_{nx} , λ_{ny} and λ_{ni} are status of disturbing estimator, \hat{w}_x , \hat{w}_y and \hat{w}_z are state estimate of trailing vortex velocity w_x , w_y and w_z , T_{wx} , T_{wy} , $T_{wz} > 0$ are constants. Let $\tilde{w}_x = \hat{w}_x - w_x$, $\tilde{w}_y = \hat{w}_y - w_y$ and $\tilde{w}_z = \hat{w}_z - w_z$, error dynamic model of (11) is

$$\dot{\tilde{w}}_x = -\tilde{w}_x / T_{wx}, \dot{\tilde{w}}_y = -\tilde{w}_y / T_{wy}, \dot{\tilde{w}}_z = -\tilde{w}_z / T_{wz} \quad (12)$$

According to (11), kinematics equations of the follower can be rewritten as,

$$\dot{x}_f = \hat{V}_f \cos \hat{\gamma}_f \cos \hat{\chi}_f, \dot{y}_f = \hat{V}_f \cos \hat{\gamma}_f \sin \hat{\chi}_f, \dot{z}_f = -\hat{V}_f \sin \hat{\gamma}_f \quad (13)$$

where

$$\begin{cases} \hat{V}_f = \sqrt{(V_f \cos \gamma_f \cos \chi_f + \hat{w}_x)^2 + (V_f \cos \gamma_f \sin \chi_f + \hat{w}_y)^2 + (-V_f \sin \gamma_f + \hat{w}_z)^2} \\ \hat{\gamma}_f = -\sin^{-1} \left(\frac{\hat{w}_z - V_f \sin \gamma_f}{\hat{V}_f} \right) \\ \hat{\chi}_f = \chi_f + \sin^{-1} \left(\frac{\hat{w}_y \cos \chi_f - \hat{w}_x \sin \chi_f}{\hat{V}_f \cos \hat{\gamma}_f} \right) \end{cases} \quad (14)$$

thus, we introduce the following tracking error equations,

$$\begin{bmatrix} e_x \\ e_y \\ e_z \end{bmatrix} = \begin{bmatrix} \cos \hat{\chi}_f & \sin \hat{\chi}_f & 0 \\ -\sin \hat{\chi}_f & \cos \hat{\chi}_f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_f - x_r \\ y_f - y_r \\ z_f - z_r \end{bmatrix} \quad (15)$$

and

$$\begin{cases} \dot{e}_x = \hat{V}_f \cos \hat{\gamma}_f - V_r \cos \gamma_r \cos e_\chi + \hat{\chi}_f e_y \\ \dot{e}_y = V_r \cos \gamma_r \sin e_\chi - \dot{\hat{\chi}}_f e_x \\ \dot{e}_z = -\hat{V}_f \sin \hat{\gamma}_f + V_r \sin \gamma_r \end{cases} \quad (16)$$

where $e_x = \hat{\chi}_f - \chi_r$, the required velocity and track angle are

$$\begin{cases} V_d = \frac{-K_x e_x + V_r \cos \gamma_r \cos e_x - \delta V}{\cos \hat{\gamma}_f} \\ \gamma_d = \sin^{-1} \left(\frac{K_z e_z + V_r \sin \gamma_r + \hat{w}_z}{\hat{V}_f} \right) \end{cases} \quad (17)$$

where $K_x, K_z > 0$ are control inputs, and $\delta V = \hat{V}_f - V_f$. the relationship of command signal and required signal is,

$$\begin{bmatrix} \dot{\mathcal{F}}_c \\ \dot{\mathcal{F}}_c \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\zeta_F \omega_F \end{bmatrix} \begin{bmatrix} \mathcal{F}_c \\ \dot{\mathcal{F}}_c \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_F^2 \end{bmatrix} \mathcal{F}_d(t), \quad \mathcal{F} \in \{V, \gamma\} \quad (18)$$

Let $e_V = V_f - V_c$ and $e_\gamma = \gamma_f - \gamma_c$, we can obtain the following error dynamical equations,

$$\begin{cases} \dot{e}_V = \underbrace{(T \cos \alpha_f \cos \beta_f - \bar{D}) / m_f - g \sin \gamma_f}_{u_V} - \dot{V}_c + d_V \\ \dot{e}_\gamma = \underbrace{((\bar{L}_0 + \bar{L}_\alpha \alpha_f + T \sin \alpha_f) \cos \mu_f - m_f g \cos \gamma_f) / (m_f V_f)}_{u_\gamma} - \dot{\gamma}_c + d_\gamma \\ \dot{e}_\chi = \underbrace{((\bar{L}_0 + \bar{L}_\alpha \alpha_f + T \sin \alpha_f) \sin \mu_f) / (m_f V_f \cos \gamma_f)}_{u_\chi} - \dot{\hat{\chi}}_r + d_\chi \end{cases} \quad (19)$$

In error dynamical equations, u_V, u_γ and u_χ are virtual control inputs, $\dot{\hat{x}}_r$ is estimation of $\dot{\chi}_r$, $\dot{\hat{\chi}}_r = \dot{\chi}_r - \dot{\chi}_r$ is estimation error of track angle. $\dot{\hat{\chi}}_f$ is uncertainty of track angular velocity due to tail vortex. Thus, d_χ is redefined as,

$$\begin{aligned} d_x = & \frac{(L - \bar{L} + \Delta L) \sin \mu_f + (Y + \Delta Y - \cos \alpha_f \sin \beta_f) \cos \mu_f}{m_f V_f \cos \gamma_f} + \\ & \dot{\chi}_r + \dot{\chi}_f + \frac{\dot{W}_y \cos \mu_f - \dot{w}_z \sin \mu_f}{m_f V_f \cos \gamma_f} \end{aligned} \quad (20)$$

Note that χ_r is non-smooth signal, its derivation is bounded. We choose the following virtual input to stabilize error dynamical system (17) and (20)

$$u_V^d = u_{V0}^d - \hat{d}_V + \dot{V}_c, u_\gamma^d = u_{\gamma0}^d - \hat{d}_\gamma + \dot{\gamma}_{cc} \text{ and } u_\chi^d = u_{\chi0}^d - \hat{d}_\chi + \dot{\chi}_r \quad (21)$$

where $\hat{d}_V, \hat{d}_\gamma$ and \hat{d}_χ are the estimations of uncertainties and disturbing items d_V, d_γ and d_χ . $u_{V0}^d, u_{\gamma0}^d$ and $u_{\chi0}^d$ are basic virtual control variables as follows,

$$\begin{cases} u_{V0}^d = -K_V e_V - c_V \varepsilon_x \cos \hat{\gamma}_f / H \\ u_{\gamma0}^d = -K_\gamma e_\gamma \\ u_{\chi0}^d = -K_\chi \sin(\frac{e_x}{2}) - c_\chi e_y V_r \cos \gamma_r \cos(\frac{e_x}{2}) / H \end{cases} \quad (22)$$

where $H = \sqrt{\varepsilon_x^2 + \varepsilon_z^2 + 1}$, $K_V > 0$, $K_\gamma > 0$, and $K_\chi > 0$ are control gains, $c_V > 0$ and $c_\chi > 0$ are small constants. Uncertainty and disturbance estimator are designed as

$$\begin{cases} \dot{\lambda}_V = -\frac{\lambda_V}{T_V} - \frac{1}{T_V} \left(\frac{1}{T_V} V_f - K_V e_V - \frac{c_V \varepsilon_x \cos \hat{\gamma}_f}{H} + \dot{V}_c - \hat{d}_V \right) \\ \dot{\lambda}_y = -\frac{\lambda_\gamma}{T_\gamma} - \frac{1}{T_\gamma} \left(\frac{1}{T_\gamma} \gamma_f - K_\gamma e_\gamma + \dot{\gamma}_c - \hat{d}_\gamma \right) \\ \dot{\lambda}_x = -\frac{\lambda_\chi}{T_\chi} - \frac{1}{T_\chi} \left(\frac{1}{T_\chi} \chi_f - K_\chi e_\chi + \dot{\chi}_r - \frac{c_\chi e_y V_r \cos \gamma_r \cos(\frac{e_x}{2})}{H} - \hat{d}_\chi \right) \\ \hat{d}_V = \lambda_V + \frac{1}{T_V} V_f \\ \hat{d}_\gamma = \lambda_\gamma + \frac{1}{T_\gamma} \gamma_f \\ \hat{d}_\chi = \lambda_\chi + \frac{1}{T_\chi} \chi_f \end{cases} \quad (23)$$

where $T_V, T_\gamma, T_\chi > 0$ are constant.

Position controller are proposed as following equations,

$$\begin{cases} T_c = \frac{m_f(u_V^d + g \sin \gamma_f) + \bar{D}}{\cos \alpha_f \cos \beta_f} \\ \alpha_d = \frac{m_f V_f \sqrt{(u_\gamma^d + \frac{g}{V} \cos \gamma_f)^2 + (u_x^d)^2 \cos^2 \gamma_f} - T \sin \alpha_f - \bar{L}_0}{\bar{L}_\alpha} \\ \mu_d = \tan^{-1} \left(\frac{m_f V_f u_\chi^d \cos \gamma_f}{m_f V_f u_y^d + m_f g \cos \gamma_f} \right) \end{cases} \quad (24)$$

where T_c is thrust input, μ_d required flight path angle, and α_d required attack angle.

Inner Attitude Control. This section will introduce inner attitude controller design. Let

$$\mathbf{u}_\tau = -I^{-1} \boldsymbol{\Omega} \times I \boldsymbol{\Omega} + I^{-1} (\bar{\tau}_0 + \bar{\mathbf{M}}_\tau \delta_{\mathbf{u}}) = [u_p, u_q, u_r]^T \quad (25)$$

error dynamics is

$$\dot{\mathbf{e}}_{\Omega} = \mathbf{u}_{\tau} + \mathbf{d}_{\tau} - \dot{\boldsymbol{\Omega}}_c \quad (26)$$

where $\mathbf{d}_{\tau} = I^{-1}(\boldsymbol{\tau} - \bar{\boldsymbol{\tau}} + \Delta\boldsymbol{\tau})$ represents summary of model uncertainties and formation aerodynamic disturbance. Thus, control law \mathbf{u}_{τ} is designed as

$$\mathbf{u}_{\tau}^d = -\mathbf{K}_{\Omega}\mathbf{e}_{\Omega} - \mathbf{C}_{\Omega}\mathbf{G}^T\boldsymbol{\varepsilon}_{\Theta} - \hat{\mathbf{d}}_{\tau} + \dot{\boldsymbol{\Omega}}_c \quad (27)$$

where $\mathbf{K}_{\Omega} = \text{diag}\{K_p, K_q, K_r\} > 0$ and $\mathbf{C}_{\Omega} = \text{diag}\{c_p, c_q, c_r\} > 0$ are diagonal constant gain matrix, uncertainty and disturbance estimator are described as

$$\begin{cases} \hat{d}_{\tau} = \lambda_{\Omega} + T_{\Omega}^{-1}\mathbf{e}_{\Omega} \\ \dot{\lambda}_{\Omega} = -T_{\Omega}^{-1}\lambda_{\Omega} - T_{\Omega}^{-1}(T_{\Omega}^{-1}\mathbf{e}_{\Omega} - \mathbf{K}_{\Omega}\mathbf{e}_{\Omega} - \mathbf{C}_{\Omega}\mathbf{G}^T\boldsymbol{\varepsilon}_{\Theta} - \hat{d}_{\tau}) \end{cases} \quad (28)$$

where $T_{\Omega} = \text{diag}\{T_p, T_q, T_r\} > 0$ is constant matrix, $\boldsymbol{\delta}_c = [\delta_{ac}, \delta_{ec}, \delta_{rc}]^T$ rudder deflection vector, and then

$$\boldsymbol{\delta}_c = \mathbf{M}_{\tau}^{-1}(\mathbf{I}\mathbf{u}_{\tau}^d + \boldsymbol{\Omega} \times \mathbf{I}\boldsymbol{\Omega} - \bar{\boldsymbol{\tau}}_0) \quad (29)$$

4 Simulation

We assume initial conditions of leader drone are

$$x_l(0) = 45 \text{ m}, y_l = -15 \text{ m}, z_l = -5015 \text{ m}, V_l = 200 \text{ m/s}, \gamma_l = \chi_l = 0 \text{ deg} \quad (30)$$

The relative optimal position vector is $\mathbf{I}_W = [-36, 9, 0]^T \text{ m}$, and initial conditions of follower drone is

$$\begin{aligned} x_f(0) &= 45 \text{ m}, y_f(0) = -15 \text{ m}, z_f(0) = -5015 \text{ m} \\ V_f(0) &= 200 \text{ m/s}, \gamma_f(0) = \chi_f(0) = \mu_f(0) = \beta_f(0) = 0 \text{ deg} \\ \alpha_f(0) &= 2.774 \text{ deg}, p = q = r = 0 \text{ rad/s} \end{aligned} \quad (31)$$

In the simulation, trajectories of formation are illustrated as Fig. 3. The formation keeps flying in a line at first 35 s, and then leader drone makes turning maneuver until 145 s. After that, leader drone resumes straight flight. In Fig. 4, attitude and relative positions of leader and follower drone are figured at four different moments. Follower drone is far away from leader drone at first, and then becomes to move fast to track leader drone, as shown in Fig. 5. The relative position between leader and follower drones in three-dimensional graph is shown in 180 s.

In the simulation, another controller is used to demonstrate the advantages of proposed control method. One is proposed nonlinear robust formation controller, another only contains basic nonlinear controller without disturbance observers. The position tracking error of two methods is shown in Fig. 7. According to aerodynamic analysis, if the tracking error cannot be maintained in 10% wingspan in horizontal and vertical directions, the close formation is not effective. As to more effective close formation, this number is less than 5%. By adding

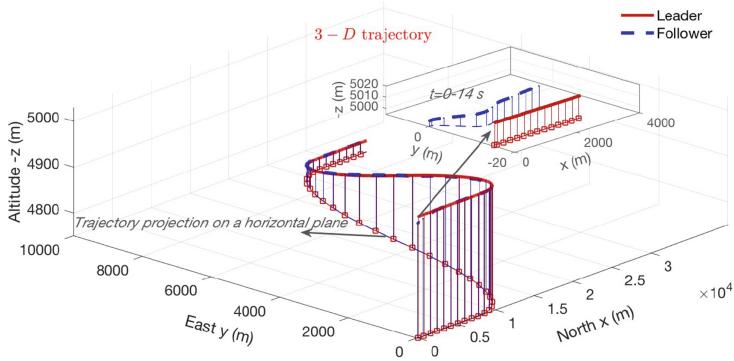


Fig. 3. Trajectories of formation

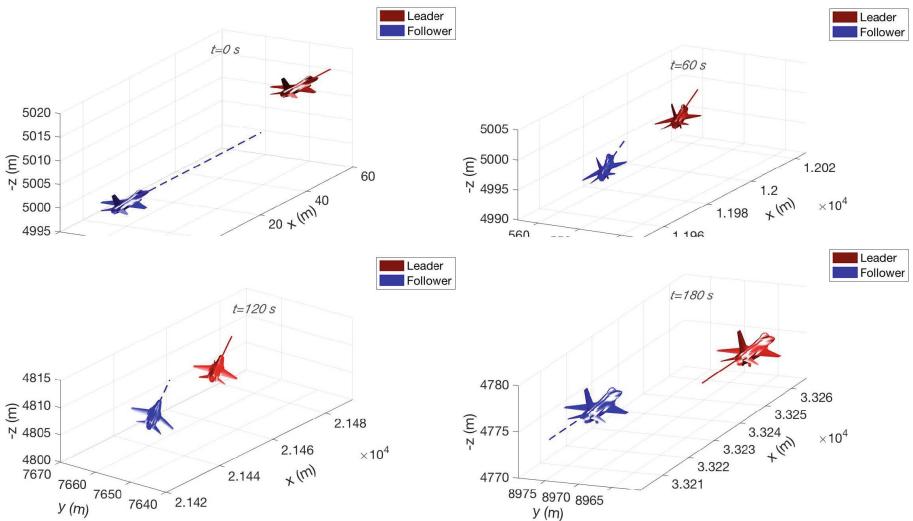


Fig. 4. Relative position between leader and follower drones by 3D illustration

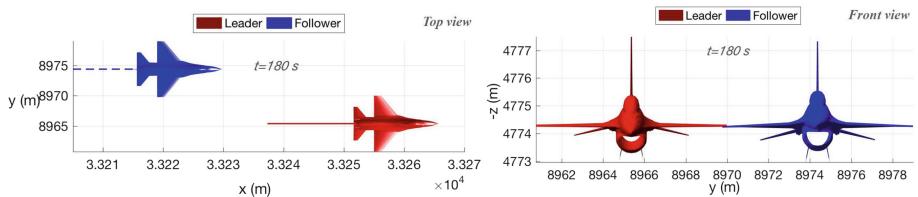


Fig. 5. Relative position between leader and follower drones when $t = 180$ s

to disturbance observer, position error is less than 5%, as shown in Fig. 8. When leader drone maneuvers between 35 s to 145 s we can observe position tracking

error, and it is proved to be less than 5% wingspan, as shown in Fig. 8. It is proved that close formation can be maintained effectively (Fig. 6).

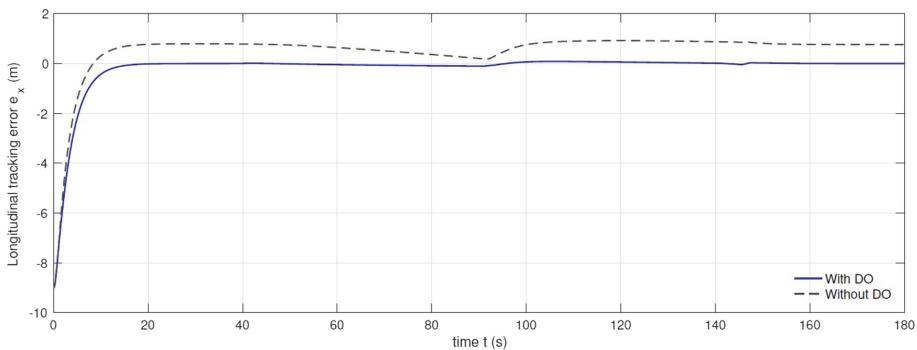


Fig. 6. Longitudinal tracking error between the two different algorithms

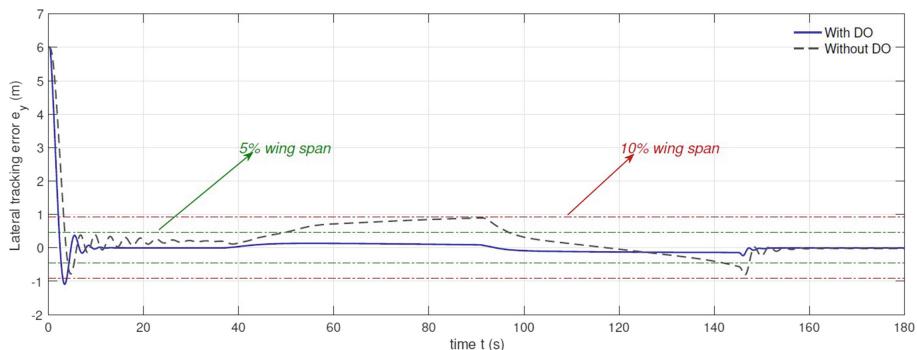


Fig. 7. Variations of longitudinal tracking error

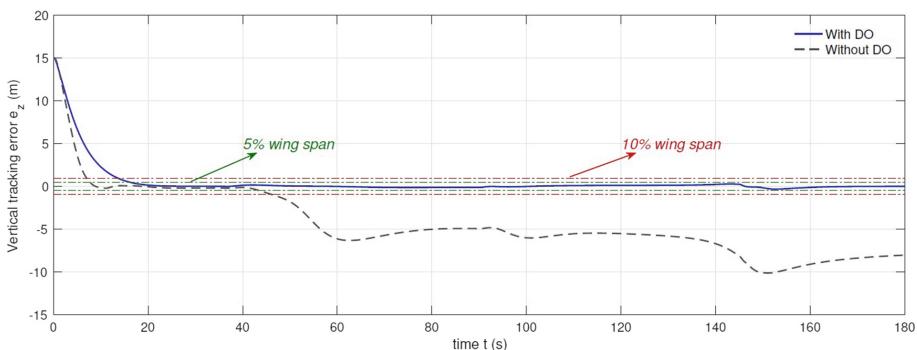


Fig. 8. Variations of vertical tracking error

5 Conclusion

In this paper, a nonlinear robust controller is proposed for autonomous close formation flight under different leader UAV flight maneuvers. Although this is described in the scenario of two UAVs under close formation, the designer can be extended to more than three UAVs. The controller is designed to adapt to tight formation flying under different maneuvers, which can not be solved by novel linear controller. Through the controller designed in this paper, even if the leader adopts different control methods, the follower can also track its best position in the tight formation. At the same time, the outer loop controller and the inner loop controller are studied, so the assumption of well-designed inner loop controller is no longer needed. In the simulation, the proposed design can achieve enough control robustness and high precision under close formation.

Acknowledgements. The work described in this paper was supported by Chinese Postdoctoral Science Foundation, No. 2019M650792.

References

1. Altshuler, Y., Yanovsky, V., Wagner, I.A., et al.: Efficient cooperative search of smart targets using UAV Swarms. *Robotica* **26**(4), 551–557 (2008)
2. Nadia, H., Saber, M., Kais, B.: Cooperative Distributed Model Predictive Control for Nonlinear System Multi-vehicle Formation (2018)
3. Li, X.-H., Wu, X.: The application of UAV remote sensing surveying and mapping technology in engineering surveying and mapping. *World Nonferrous Met.* **17**, 297–298 (2019)
4. Mir, I., Maqsood, A., Eisa, S.A., et al.: Optimal morphing - augmented dynamic soaring maneuvers for unmanned air vehicle capable of span and sweep morphologies. *Aerosp. Sci. Technol.* **79**(Aug), 17–36 (2018)
5. Cantelli, L., Mangiameli, M., Melita, C.D., et al.: UAV/UGV cooperation for surveying operations in humanitarian demining. In: IEEE International Symposium on Safety, Security, & Rescue Robotics. IEEE (2013)
6. Richert, D., Cortés, J.: Optimal leader allocation in UAV formation pairs ensuring cooperation. *Automatica* **49**(11), 3189–3198 (2013)
7. Cantelli, L., Presti, M.L., Mangiameli, M., et al.: Autonomous cooperation between UAV and UGV to improve navigation and environmental monitoring in rough environments. In: 10th International Symposium Humanitarian Demining (HUDEM) (2013)
8. Duan, H.: Optimal formation reconfiguration control of multiple UCAVs using improved particle swarm optimization. In: Guidance, Navigation & Control Conference. IEEE Chinese, IEEE (2014)
9. Bian, L., Sun, W., Sun, T.: Trajectory following and improved differential evolution solution for rapid forming of UAV formation. *IEEE Access* **7**(99), 169899–169613 (2019)



3D Path Planning for UAV with Improved Double Deep Q-Network

Liping Zhao¹, Yaofei Ma^{1,2}, and Jie Zou^{2(✉)}

¹ Beihang University, Beijing 100191, China

² Science and Technology on Electro-optic Control Laboratory, Luoyang, China
zoujie613@126.com

Abstract. Unmanned aerial vehicle (UAV) has been widely used in civil and military fields due to its advantages such as zero casualties, low cost and strong maneuverability. Path planning in 3D obstacle environment is one of the fundamental capabilities of UAV for mission performing. In this paper, we propose a 3D path planning algorithm to learn a target-driven end-to-end model based on an improved double deep Q-network (DQN), where a greedy exploration strategy is applied to accelerate learning. The model takes target and obstacle message as input, and moving command of UAV as output. It can realize path planning successfully for UAV in 3D complex environment. Besides, the experimental results show that improved double DQN has better convergence speed compared with DQN and double DQN.

Keywords: 3D path planning · UAV · Improved double DQN · Greedy exploration strategy · Reinforcement learning

1 Introduction

Path planning is one of the core contents of research in robotics, which aims to give an optimal path based on one or more optimization criteria from location A to location B in various environments. It has been used to solve a wide range of problems, such as GPS navigation, autonomous driving, automatic sweeping, robot arm motion planning, package routing, etc. Excellent path planning algorithms can not only save time, but also avoid the waste of electricity, fuel and other resources [1]. Nowadays, UAV plays an irreplaceable role as a kind of important equipment of air missions. Consequently, the path planning of UAV has become a hot research spot at home and abroad.

Algorithms of path planning can be divided into four categories: traditional algorithms, grid-based algorithms, sampling-based algorithms and intelligent bionic algorithms. Traditional algorithms include simulated annealing algorithm

This work is supported by Science and Technology on Electro-optic Control Laboratory and Aviation Science Fund under Grant 20175151024.

(SA), artificial potential field (APF) algorithm, fuzzy logic algorithm, etc. The combination of APF and iterative regional inflation by semidefinite (IRIS) in [2] achieved the goal of obstacle avoidance during UAV advancement process. Grid-based algorithms divide the environment map into grids which is convenient for the computer to establish model. [3] and [4] applied Dijkstra and A-star to navigation and obstacle avoidance tasks. Sampling-based algorithms such as rapidly-exploring random trees (RRT) and probabilistic roadmaps (PRM) are suitable for dealing with high-dimensional problems due to high speed. [5] proposed control barrier function guided RRT to handle dynamical obstacles. Algorithms whose basic idea come from inspiration of nature build up intelligent bionic algorithms, including genetic algorithm (GA), ant colony (AC) algorithm and so on. GA searches for the optimal solution by simulating selecting, crossing, and mutating in the process of evolution. [6] solved cooperative path planning of UAV with GA. However, most of these algorithms don't take the kinematic constraints of the agent into account. Due to the existence of kinematic constraints, the agent may not be able to move along the planned path. In our method, We first design the actions based on kinematics and dynamics, then plan the path based on the actions, so the planned results can be directly used by agents.

Deep reinforcement learning (DRL) has achieved great success recently. DQN, the combination of deep neural network and Q-learning, have been widely employed. In this paper, a 3D path planning algorithm is proposed for UAV based on improved double DQN, a variant of DQN. The path planning problem is modeled as a sequential decision-making procedure to solve via a target-driven end-to-end planner, which can output moving command of UAV directly from target and obstacle information. Besides, in order to accelerate the training of network, a greedy exploration strategy is applied to double DQN.

The rest of this paper is organized as follows: In Sect. 2, we introduce related work on DRL and path planning with DRL. Section 3 introduces the details of our approach. In order to analyze the effectiveness of our method in UAV path planning, comparative experiments are conducted in Sect. 4. Finally, some conclusions are presented in Sect. 5.

2 Related Work

Literatures about path planning based on DRL have been published intensively over the past few years. In this section, we introduce some work on DRL and applications of it in path planning problems.

2.1 Deep Reinforcement Learning

In recent years, the achievements of DRL has attracted worldwide attention. In 2017, the AI programme AlphaGo Zero from DeepMind won 100-0 against champion-defeating AlphaGo which was published previously [7]. In 2018, OpenAI Five system which was trained with proximal policy optimization (PPO)

algorithm [8] defeated human players in the competitive five-on-five video game Dota 2.

DQN was proposed by DeepMind in 2013 [9] and subsequently improved in 2015 [10]. For the first time, DQN learned game strategies in a high-dimensional state space from scratch. Since then, people have paid extensive attention to deep reinforcement learning. Lots of research has been carried out with the growing focus on it.

Double DQN [11] was proposed to reduce the overestimations of action values, a well-known problem of DQN. In addition, there are still other improved algorithms of DQN [12–14].

2.2 Path Planning with DRL

DRL algorithms have been applied to path planning problems. [15] realized path planning of UAV on 2D plane under both static and dynamic task settings by employing dueling double DQN. In [16], the avoidance obstacle path was planned by dueling DQN in the robot’s navigation. [17] utilized the deep deterministic policy gradient (DDPG) algorithm to realize intelligent path planning of unmanned ships in unknown environment. [18] proposed an end-to-end DRL model which capable to learn to navigate in multiple cities without a map.

Generally, path planning with DRL has the advantage of end-to-end planning, which is also reflected in this paper. In addition, most of the research is carried out on 2D planes while our work realizes 3D path planning which is still not generally seen. Besides, aiming to reduce the disadvantage of DRL’s long training time, we introduce a greedy exploration strategy. The focus of our work is to demonstrate a practicable approach based on DRL for UAV path planning in 3D obstacle environment and improve the exploration strategy simultaneously.

3 Approach

3.1 The Improved Double DQN Algorithm

Double DQN is a very simple and effective one of DQN’s series of improved algorithms. For this reason, we choose double DQN as a basic idea for our work.

The two main techniques of DQN are target network and replay buffer. DQN uses two Q networks with the same structure. Behavior network interacts with environment and updates parameters by gradient descent. Target network is used to generate target Q value. Its parameters are copied from behavior network at regular intervals. Replay buffer stores transition (s_t, a_t, r_t, s_{t+1}) generated by the interaction between agent and environment. Sample random transitions when updating network parameters. Not only can the transitions be used multiple times, but also the correlation among data can be broken.

DQN has the shortcoming of overestimation in most cases. Normally, the overestimation is caused by Q value update rule of taking the maximum Q value of new state, as is shown in (1).

$$y_j = \begin{cases} r_j & \text{for terminal } s_{j+1} \\ r_j + \gamma \max_{a'} \hat{Q}(s_{j+1}, a'; \theta^-) & \text{for non-terminal } s_{j+1} \end{cases} \quad (1)$$

where s_{j+1} is the state of agent; r_j represents the reward of s_j ; γ is discount factor; $\hat{Q}(s_{j+1}, a'; \theta^-)$ is the output of target Q-network with input s_{j+1} and parameter θ^- .

Double DQN improved Q value update rule to reduce overestimation. As is shown in (2), it selects the action corresponding to the maximum Q value of current Q-network to calculate the target Q value rather than use the the maximum Q value of target Q-network.

$$y_j = \begin{cases} r_j & \text{for terminal } s_{j+1} \\ r_j + \gamma \hat{Q}(s_{j+1}, \max_{a'} Q(s_{j+1}, a'; \theta); \theta^-) & \text{for non-terminal } s_{j+1} \end{cases} \quad (2)$$

The compromise between exploration and exploitation has always been a great challenge in RL. Double DQN applies ϵ -greedy strategy, which is simple but not effective enough. Greedy algorithm always makes the best choice at the present time when solving a problem. If it is applied to path planning, completing the planning successfully is difficult because it only focuses on current reward. However, the local optimal solution is better than a randomly selected action often as long as the reward function is designed reasonably. The proportion of good actions has a great influence on the convergence speed of the algorithm. Thus, We consider combining ϵ -greedy with greedy exploration strategy. The detailed steps of the improved double DQN are shown in Algorithm 1.

3.2 Construction of Deep Neural Network

In order to build the neural network, we need to design the state space and action space first to determine the input and output of network. The UAV is modeled based on dynamics and kinematics, etc. [19]. In the environment of path planning, the information can be acquired by UAV includes its own status, target position and obstacle information obtained through sensors.

The state space is designed as a 1×5 vector $(xd, yd, zd, \alpha, d_{uo})$, where (xd, yd, zd) is the distance between UAV and target position in the three directions of xyz , α is the the direction angle of UAV relative to target position on $x-y$ plane and d_{uo} is the distance between UAV and the nearest obstacle. In addition, we take the state of consecutive 5 steps as the input of the network to help understand the task.

UAV can perform actions such as level flight, climbing, diving, left turn and right turn. These types of actions are discretized to adapt to double DQN algorithm. Finally, the discrete action space consists of nine actions, which are 4 meters of level flight, climbing 100 m, diving 100 m, turning left 20° , 45° , 60° and turning right 20° , 45° , 60° .

In conclusion, we build a fully connected network with 4 layers. And nonlinear activation function relu is used in the network. The input layer, two hidden layers and output layer have 25, 64, 32 and 9 nodes, respectively. Figure 1 shows the specific structure of the network.

Algorithm 1: Improved Double DQN

```

Initialize replay memory  $D$  to capacity  $N$ 
Initialize action-value function  $Q$  with random weights  $\theta$ 
Initialize target action-value function  $\widehat{Q}$  with weight  $\theta^- = \theta$ 
Initialize  $\varepsilon = 1$ ,  $\varepsilon_{min} = 0.05$ 
for  $episode = 1, M$  do
    Initialize sequences  $s_1=x_1$ 
    for  $t = 1, T$  do
        generate a random probability  $p \in (0, 1)$ 
        if  $p < \varepsilon/2$  then
            | select a random action  $a_t$ 
        end
        else if  $\varepsilon/2 \leq p < \varepsilon$  then
            | select  $a_t = \arg \max_{a \in A} r(s_t, a)$ 
        end
        else
            | select  $a_t = \arg \max_{a \in A} Q(s_t, a; \theta)$ 
        end
        Execute  $a_t$  in emulator and observe reward  $r_t$  and state  $s_{t+1}$ 
        Store transition  $(s_t, a_t, r_t, s_{t+1})$  in  $D$ 
         $s_t = s_{t+1}$ 
        Sample random minibatch of transitions  $(s_j, a_j, r_j, s_{j+1})$  from  $D$ 
        Set
        
$$y_j = \begin{cases} r_j & \text{for terminal step } j + 1 \\ r_j + \gamma \widehat{Q}(s_{j+1}, \max_{a' \in A} Q(s_{j+1}, a'; \theta); \theta^-) & \text{otherwise} \end{cases}$$

        Perform a gradient descent step on  $(y_j - Q(s_j, a_j; \theta))^2$  with respect to
        the network parameters  $\theta$ 
        Every  $C$  steps reset  $\widehat{Q} = Q$ 
        Decrease  $\varepsilon$  until  $\varepsilon \leq \varepsilon_{min}$ 
    end
end

```

3.3 The Design of Reward Function

The design of reward function is extremely crucial in RL since reward is the only signal that agent can receive in the interaction with environment. Path planning requires the agent to reach the target position with a collision-free path consisting of as few steps as possible. Thus, the reward function is designed to include 3 parts related to obstacle and target information as shown in (3), where d_{ut} is the distance between UAV and target position; d_{uo} is the shortest distance between UAV and obstacle; Δd_{ut} is the difference of d_{ut} between the last moment and the current moment; δ is the minimum standard for arrival.

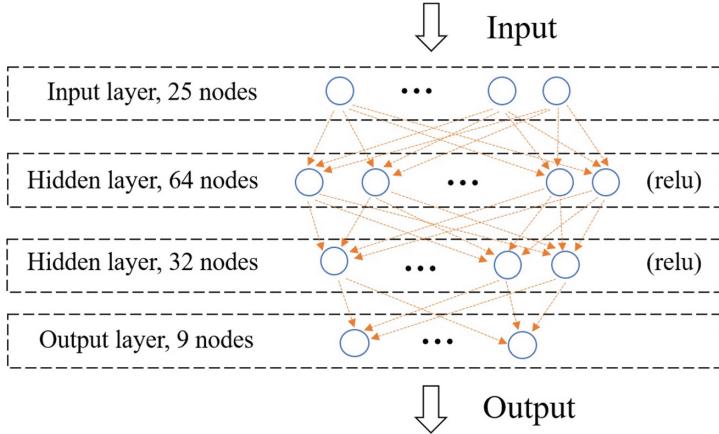


Fig. 1. The structure of deep neural network

$$r = \begin{cases} +10 & \text{if } d_{ut} < \delta \\ -10 & \text{if } d_{uo} \leq 0 \\ -0.15 + 0.001 \times \Delta d_{ut} - \frac{0.005}{d_{uo}} & \text{otherwise} \end{cases} \quad (3)$$

It's necessary to judge whether the UAV reaches the target position and whether a collision occurs at each step. A big positive reward (+10) will be given if the UAV arrives the target position. In contrast, a big negative reward (-10) will be given if the UAV collided with an obstacle. Otherwise, the reward is calculated by a continuous function according to different effect of action choose by the UAV. What is noteworthy is that a small negative reward (-0.15) at each step is designed to motivate the UAV reaches the target position as soon as possible. Moreover, when the UAV satisfies one of the below three conditions: 1) arrival, 2) collision, 3) timeout, an episode will be stopped and the next episode will be started in the training process.

4 Experiments

4.1 Experimental Environment

We mainly consider buildings and no-fly zone the two kinds of obstacles in environment. They are modeled as cylinders and hemispheres respectively. The cylinder obstacle is described as shown in (4), where $z_1(x, y)$ is the obstacle height of position (x, y) ; (x_i, y_i) and h_i are the center coordinate and the height of the i th cylinder, respectively; R is the radius of the cylinder bottom.

$$z_1(x, y) = h_i, \quad r < R$$

$$r = \sqrt{(x - x_i)^2 + (y - y_i)^2} \quad (4)$$

The hemisphere obstacle is described as shown in (5), where where $z_2(x, y)$ is the obstacle height of position (x, y) ; (x_i, y_i) is the center coordinate of the i th hemisphere obstacle; R is the radius of hemisphere.

$$\begin{aligned} z_2(x, y) &= \sqrt{R^2 - r^2}, \quad r < R \\ \mathbf{r} &= \sqrt{(x - x_i)^2 + (y - y_i)^2} \end{aligned} \quad (5)$$

The combination of the two kinds of obstacles is shown in (6), where $z_1(x, y)$ and $z_2(x, y)$ represents the information of cylinder and hemisphere obstacle, respectively.

$$z(x, y) = \max(z_1(x, y), z_2(x, y), 0) \quad (6)$$

We build three different environment models A , B and C for path planning, which is shown in Fig. 3 in Sect. 4.3.

4.2 Implementation Details

We set the scale of the space to $5 \text{ km} \times 5 \text{ km} \times 1.5 \text{ km}$. Three environment models with different obstacles are created. The target position is $(0.3 \text{ km}, 0.3 \text{ km}, 0.3 \text{ km})$. The height of all the cylinder obstacles is set to 1500 m and the radius of the hemisphere obstacles is set to 500 m. The position of target and obstacles is fixed during the training process.

There are 3 obstacles in environment A and B respectively. In environment A and B , the initial position of UAV is randomly generated at the beginning of each episode around $(3 \text{ km}, 3 \text{ km}, 1 \text{ km})$. The two cylinder obstacles' center coordinates on $x - y$ plane are $(1.5 \text{ km}, 1 \text{ km})$ and $(2 \text{ km}, 2.5 \text{ km})$ in A , $(0.8 \text{ km}, 2 \text{ km})$ and $(2 \text{ km}, 2.5 \text{ km})$ in B . The center coordinate of the hemisphere obstacle is $(1 \text{ km}, 2 \text{ km}, 0 \text{ km})$ in A and $(1 \text{ km}, 1 \text{ km}, 0 \text{ km})$ in B .

Environment C is more complex than A and B . In environment C , the five cylinder obstacles' center coordinates on $x - y$ plane are $(0.8 \text{ km}, 2 \text{ km})$, $(2.5 \text{ km}, 1.5 \text{ km})$, $(2 \text{ km}, 4 \text{ km})$, $(3 \text{ km}, 3.5 \text{ km})$ and $(4 \text{ km}, 3 \text{ km})$. The two hemisphere obstacles' center coordinates are $(1 \text{ km}, 1 \text{ km}, 0 \text{ km})$ and $(2 \text{ km}, 2.5 \text{ km}, 0 \text{ km})$.

To train a model with excellent performance, we have made plenty of attempts to determine the the values of hyperparameters. The final values we use are shown in Table 1.

Table 1. The values of hyperparameters

Hyperparameter	Value
γ	0.99
Learning rate	0.0005
Batch size	64
Memory size	100,000,000
Update frequency	300

4.3 Results

We train the model for $30k$ episodes in A , $60k$ episodes in B and C with DQN, double DQN and improved double DQN respectively. The maximum steps of UAV is set to 200 during training. The total reward is shown in Fig. 2, where the red lines represent improved double DQN; the blue lines represent double DQN; the green lines represent DQN. In order to make the figure clearer, we average the reward every 100 episodes for A , 200 episodes for B and C . It is obvious that these three algorithms perform better and better with the increase of training episode. In A and C , double DQN converges ahead of DQN. DQN and double DQN converge almost simultaneously in B . But improved double DQN proposed by us converges ahead of DQN and double DQN in all environment models. What's more, the improvement of our method in environment C is greater than A and B , which indicates that the performance of improved double DQN performs better in a more complex environment.

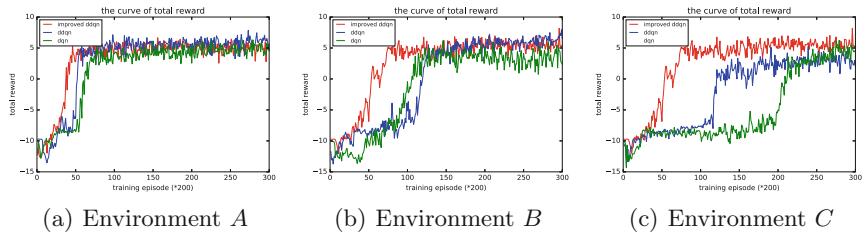


Fig. 2. Total reward in environment A , B and C

We test the trained models separately in environment A , B and C . The optimal results of path planning by the three algorithms are the same. Figure 3 shows the optimal path of UAV without collision from the initial position to the target position.

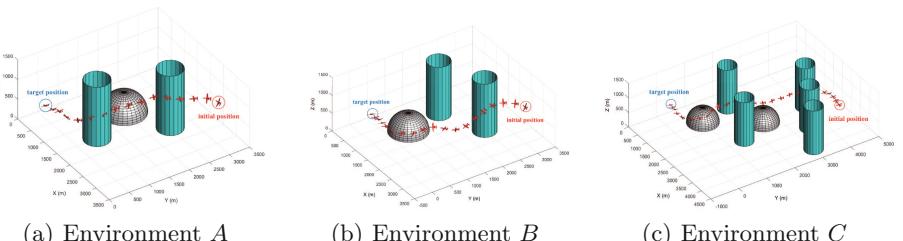


Fig. 3. Path planning for UAV with DRL in environment A , B and C

5 Conclusion

In this paper, a target-driven end-to-end planner based on DRL is proposed for UAV's collision-free path planning in 3D environment. The planner only takes target and obstacle information as input and outputs action command of UAV directly. Double DQN is used as the basic algorithm for its simple and effective improvement of DQN. In order to speed up the convergence of double DQN, the greedy exploration strategy is introduced to better balance exploration and exploitation. Comparative experiments are performed with DQN, double DQN and improved double DQN in three different environment models. The results prove that the trained planner is capable for producing feasible path and the method improved by us realizes training acceleration.

References

1. Zhang, H., Lin, W., Chen, A., et al.: Path planning for the mobile robot: a review. *Symmetry* **10**(10), 450 (2018). <https://doi.org/10.3390/sym10100450>
2. Li, Y., Wu, X., Zhang, Q., et al.: Study of the obstacle avoidance plan of UAV based on IRIS algorithm. In: International Conference Control Science and Systems Engineering, pp. 74–79 (2017). <https://doi.org/10.1109/CCSSE.2017.8087898>
3. Singh, Y., Sharma, S., Sutton, R., et al.: Feasibility study of a constrained Dijkstra approach for optimal path planning of an unmanned surface vehicle in a dynamic maritime environment. In: IEEE International Conference on Autonomous Robot Systems and Competitions, pp. 117–122 (2018). <https://doi.org/10.1109/ICARSC.2018.8374170>
4. Prabhu, S.G., Kyberd, P., Wetherall, J., et al.: Investigating an a-star algorithm-based fitness function for mobile robot evolution. In: International Conference on System Theory, Control and Computing (2018). <https://doi.org/10.1109/ICARSC.2018.8374170>
5. Yang, G., Vang, B., Serlin, Z., et al.: Sampling-based motion planning via control barrier functions. *arXiv: Robotics* (2019)
6. Huang, H., Zhuo, T.: Multi-model cooperative task assignment and path planning of multiple UCAV formation. *Multi. Tools Appl.* **78**(1), 415–436 (2019). <https://doi.org/10.1007/s11042-017-4956-7>
7. Silver, D., Schrittwieser, J., Simonyan, K., et al.: Mastering the game of go without human knowledge. *Nature* **550**(7676), 354–359 (2017). <https://doi.org/10.1038/nature24270>
8. Schulman, J., Wolski, F., Dhariwal, P., et al.: Proximal policy optimization algorithms. *arXiv: Learning* (2017)
9. Mnih, V., Kavukcuoglu, K., Silver, D., et al.: Playing Atari with deep reinforcement learning. *arXiv: Learning* (2013)
10. Mnih, V., Kavukcuoglu, K., Silver, D., et al.: Human-level control through deep reinforcement learning. *Nature* **518**(7540), 529–533 (2015). <https://doi.org/10.1038/nature14236>
11. Van Hasselt, H., Guez, A., Silver, D., et al.: Deep reinforcement learning with double Q-Learning. In: National Conference on Artificial Intelligence, pp. 2094–2100 (2016)

12. Wang, Z., Schaul, T., Hessel, M., et al.: Dueling network architectures for deep reinforcement learning. In: International Conference on Machine Learning, pp. 1995–2003 (2016)
13. Osband, I., Blundell, C., Pritzel, A., et al.: Deep exploration via bootstrapped DQN. In: Neural Information Processing Systems, pp. 4033–4041 (2016)
14. Gu, S., Lillicrap, T., Sutskever, I., et al.: Continuous deep Q-learning with model-based acceleration. *arXiv: Learning* (2016)
15. Yan, C., Xiang, X., Wang, C., et al.: Towards real-time path planning through deep reinforcement learning for a UAV in dynamic environments. *J. Intell. Robot. Syst.* 1–13 (2019). <https://doi.org/10.1007/s10846-019-01073-3>
16. Wen, S., Zhao, Y., Yuan, X., et al.: Path planning for active SLAM based on deep reinforcement learning under unknown environments. *Intell. Serv. Robot.* 1–10 (2020). <https://doi.org/10.1007/s11370-019-00310-w>
17. Guo, S., Zhang, X., Zheng, Y., et al.: An autonomous path planning model for unmanned ships based on deep reinforcement learning. *Sensors*, **20**(2) (2020). <https://doi.org/10.3390/s20020426>
18. Mirowski, P., Grimes, M.K., Malinowski, M., et al.: Learning to navigate in cities without a map. *arXiv: Artificial Intelligence* (2018)
19. Mobarez, E.N., Ouda, A.N., Zekry, A., et al.: Mathematical representation, modeling and linearization for fixed wing UAV. *Int. J. Comput. Appl.* **147**(2), 24–31 (2016)



Fuzzy Adaptive Triangular Formation Control for a Class of Nonlinear System

Hao Wang, Qin Wang, Hong Shen^(✉), and Tianping Zhang

Yangzhou University, Yangzhou 225100, China
shenhong@yzu.edu.cn

Abstract. An adaptive triangular formation control with uncertain nonlinear dynamics and external disturbances is studied in this paper. The potential energy function method is introduced to design the adaptive formation control law, nonlinear dynamics are approximated by the first type fuzzy systems. Furthermore, the parameter adaptive estimation is used to eliminate the effects of approximation error and external disturbances. Finally, the nonlinear formation system can achieve a triangular or a line formation shape, and all agents' velocities converge to a common value without collision between any two agents. The effectiveness of the proposed control strategy is verified by simulation examples.

Keywords: Formation control · Bounded interference · Perturbation method · Globally asymptotic stability

1 Introduction

Due to the vigorous development and convergence of disciplines such as control science and communication engineering, multi-agent coordinated control has caught the attention of researchers in different fields at home and abroad recently. As shown in previous research, it has been found that compared with the traditional single agent, the multi-agent system formation control has shown great advantages [1] in pipeline production, formation patrol, cargo transportation, etc., therefore, the application of multi-agent formation control system is more and more extensive.

At present, coordination tracking control of multi-agent systems has attracted more and more attention, and many scholars have conducted a large number of research on it. Now, there are many practical methods of control such as position-based formation control [2–4], distance-based formation control [5–7] and bearing-only formation control [8,9]. Among them, formation control based on relative distance seems more practical because it can avoid collisions between adjacent agents, and is considered to be a more ideal distributed control method. In the existing literature [10], some scholars are devoted to the study of algebraic graph theory to illustrate the communication relationship between agents. Later, Rimon [11] and Koditschek [12] proposed another method, which

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 384–394, 2021.

https://doi.org/10.1007/978-981-15-8458-9_41

could construct potential energy functions more accurately. Inspired by these, Tanner et al. [13] considered the clustering problem of multi-agents, and designed a control algorithm by constructing a potential energy function based on the distance between communicating agents. The overall system is converged to the set where the negative gradient is equal to zero, and the velocity is converged to the same value, and there is no collision between adjacent agents. Jin et al. [14] put forward a robust cluster intelligent control strategy with nonlinear characteristics. The proposed cluster algorithm, which considers interference and modeling uncertainty, can only make the multi-agent system tend to a set of zero gradients without specific formation shape. So far, linear control has attracted the most attention of researchers, while few studies have focused on nonlinear control. For the formation control of a single agent, the plant of the system are nonlinear under most condition, while for the multi-agent system, the formation models are mostly single integrators or double integrators. In practice, many physical objects cannot be described by a linear model, and a number of limitations may exist. In view of this, the formation control problem for a class of nonlinear systems with external disturbances is considered in this paper, and the desired triangle nonlinear formation is researched by using fuzzy control and nonlinear control.

The control law should maintain the desired triangle or line formation at a common desired speed, and ensure collision avoidance between various communication bodies during the motion. Using the back-stepping technique and adaptive method the control scheme is designed. And the nonlinear dynamics terms are approximated by the first type fuzzy systems. Furthermore, the effects of approximation error and external disturbances are eliminated by use of the parameter adaptive estimation. Finally, the Barbalat's Lemma and non-smooth theory are used to demonstrate the global stability of the overall fuzzy formation system, and any two adjacent agents should never collide with each other during the motion.

2 Problem Statement

Consider a formation comprising three agents, each agent is described by a double integrator:

$$\begin{cases} \dot{r}_i = v_i \\ \dot{v}_i = f_i + u_i + d_{ui} \end{cases} \quad (1)$$

Where $r_i = [x_i, y_i]^T$ denotes the position of agent i , v_i represents its velocity, and u_i is a control input, $i = 1, 2, 3$. $f_i(r_i, v_i) = [f_{xi}(r_{xi}, v_{xi}), f_{yi}(r_{yi}, v_{yi})]^T$ is a continuously differentiable vector-valued function, which describes the nonlinear dynamics of agent i . Where d_{ui} , $i = 1, 2, 3$ is the disturbance, and we assume $|d_{ui}| \leq D_{ui}$, D_{ui} is an unknown constant.

In the cause of designing a stable adaptive formation control law, an I-type fuzzy system is used to approximate the uncertain dynamics of the system. Define $\chi_i = [r_i^T, v_i^T]^T$, $x_i = [r_{xi}, v_{xi}]^T$, $y_i = [r_{yi}, v_{yi}]^T$, compact set $\Omega_{\chi_i} = \{\chi_i | \|\chi_i\| \leq M_{\chi_i}\}$, where M_{χ_i} is an large positive constant, the determination

method is given in the theorem. Let $f_{xi}(r_{xi}, v_{xi}, \theta_{xi})$, $f_{yi}(r_{yi}, v_{yi}, \theta_{yi})$ is an I-type fuzzy logic system's approximation of $f_{xi}(r_{xi}, v_{xi})$ and $f_{yi}(r_{yi}, v_{yi})$ in the region Ω_{χ_i} , that is, $f_{xi}(r_{xi}, v_{xi}, \theta_{xi}) = \varphi_i^T(x_i)\theta_{xi}$, $f_{yi}(r_{yi}, v_{yi}, \theta_{yi}) = \varphi_i^T(y_i)\theta_{yi}$, where $\theta_{xi} = [\theta_{1i}, \theta_{2i}, \dots, \theta_{Mi}]^T$, $\theta_{yi} = [\theta_{1i}, \theta_{2i}, \dots, \theta_{Qi}]^T$, $\varphi_i(x_i) = [\phi_{1i}(x_i), \dots, \phi_{Mi}(x_i)]^T$, $\varphi_i(y_i) = [\varphi_{1i}(y_i), \dots, \varphi_{Qi}(y_i)]^T$; M_i and Q_i are the number of the rules, $\phi_{li}(x_i)$ and $\phi_{li}(y_i)$ are fuzzy basis functions, which are defined as follows:

$$\phi_{li}(x_i) = \frac{\mu_{F_I^{li}}(x_i)}{\sum_{l_i=1}^{M_i} \mu_{F_I^{li}}(x_i)} \quad (2)$$

θ_{li} is an adjustable parameter. $\mu_{F_I^{li}}(x_i)$ is a given membership function, and usually taken as $\mu_{F_I^{li}}(x_i) = \exp(-(\frac{x_i - b_i^{li}}{c_i^{li}})^2)$, where $b_i \in R$, $c_i > 0$.

Let

$$\Omega_{xi} = \{\theta_{xi} \mid \|\theta_{xi}\| \leq M_{xi}\} \quad (3)$$

$$\Omega_{yi} = \{\theta_{yi} \mid \|\theta_{yi}\| \leq M_{yi}\} \quad (4)$$

$$\theta_{xi} = \arg \min_{\theta_{xi} \in \Omega_{xi}} [\sup_{\chi_i \in \Omega_{\chi_i}} |f(r_{xi}, v_{xi}, \theta_{xi}) - f(r_{xi}, v_{xi})|] \quad (5)$$

$$\theta_{yi} = \arg \min_{\theta_{yi} \in \Omega_{yi}} [\sup_{\chi_i \in \Omega_{\chi_i}} |f(r_{yi}, v_{yi}, \theta_{yi}) - f(r_{yi}, v_{yi})|] \quad (6)$$

where, the positive constant M_{xi} is the design parameter. Let $\hat{\theta}_{xi}(t) \in \Omega_{xi}$ and $\hat{\theta}_{yi}(t) \in \Omega_{yi}$ be the estimated values of θ_{xi} and θ_{yi} at time t , respectively.

Define the optimal approximation error:

$$\delta_{xi} = f_{xi}(r_{xi}, v_{xi}) - f_{xi}(r_{xi}, v_{xi}, \theta_{xi}) \quad (7)$$

$$\delta_{yi} = f_{yi}(r_{yi}, v_{yi}) - f_{yi}(r_{yi}, v_{yi}, \theta_{yi}) \quad (8)$$

Define

$$\varepsilon_{xi} = \max_{\theta_{xi} \in \Omega_{xi}, \chi_i \in \Omega_{\chi_i}} |\delta_{xi}| \quad (9)$$

$$\varepsilon_{yi} = \max_{\theta_{yi} \in \Omega_{yi}, \chi_i \in \Omega_{\chi_i}} |\delta_{yi}| \quad (10)$$

where ε_{xi} and ε_{yi} are unknown bounded positive constant.

then

$$f_{xi}(r_{xi}, v_{xi}) = f_{xi}(r_{xi}, v_{xi}, \theta_{xi}) + \delta_{xi} = \varphi_i^T(x_i)\theta_{xi} + \delta_{xi} \quad (11)$$

$$f_{yi}(r_{yi}, v_{yi}) = f_{yi}(r_{yi}, v_{yi}, \theta_{yi}) + \delta_{yi} = \varphi_i^T(y_i)\theta_{yi} + \delta_{yi} \quad (12)$$

From formula (11)–(12)

$$f_i(r_i, v_i) = \phi_i(\chi_i)\theta_i + \delta_i \quad (13)$$

where $\varphi_i(\chi_i) = \begin{bmatrix} \varphi_i^T(x_i) & 0 \\ 0 & \varphi_i^T(y_i) \end{bmatrix}$, $\theta_i = \begin{bmatrix} \theta_{xi} \\ \theta_{yi} \end{bmatrix}$, $\delta_i = \begin{bmatrix} \delta_{xi} \\ \delta_{yi} \end{bmatrix}$.

Substituting Eq. (13) into Eq. (1), we get

$$\begin{cases} \dot{r}_i = v_i \\ \dot{v}_i = \phi_i(\chi_i)\theta_i + u_i + \delta_i + d_{ui} \end{cases} \quad (14)$$

We model the information architecture as an undirected graph $G = (V, E)$, where $V = \{1, 2, \dots, n\}$ is a set of vertices and $E \subseteq V \times V$ is a set of edges. Let N_i denote the set of neighboring vertices of vertex i . In this paper, the distance between each pair of adjacent vertices in the graph is equal to a given constant. Let each edge $e_{ij} \in E$ in graph G be assigned to a scalar parameter $d_{ij} = d_{ji}$, which represents the desired distance every agent should preserve. Let $D = \{d_{ij} : i \in V, j \in N_i\}$, then, the framework $\{G, D\}$ represents a desired formation. Denote by $\beta_{ij} = \|r_{ij}\|^2$ the Euclidean distance between agents i and j with $j \in N_i$. $r_{ij} = r_i - r_j$, $\|r_{ij}\|$ denotes the actual distance between agent i and j . Where $\hat{\theta}_i, \hat{\gamma}_i$ are the estimated value of θ_i, γ_i at time t , and $\gamma_i = [D_{ui} + \varepsilon_{xi}, D_{ui} + \varepsilon_{yi}]^T$, $i = 1, 2, 3$.

The formation control problem is to design the formation control law $u_i = h_i(r_{ij}, v_i, v^*, \hat{\theta}_i, \hat{\gamma}_i)$, $j \in N_i$ such that for any initial condition $r_i(0) \in R^2$, $i = 1, 2, 3$, the multi-agents system can achieve the globally asymptotically stable triangle or line formation, i.e. $\lim_{t \rightarrow \infty} (\|r_{ij}\| - d_{ij}) = 0$, with a common desired velocity v^* , i.e. $\lim_{t \rightarrow \infty} v_i = v^*$, and no collision occurs between any agents, that is to say, there does not exist $t = t_1 > 0$ so that $\|r_{ij}(t)\| = 0$, where $i = 1, 2, 3, j \in N_i$.

3 Design of Global Stabilizer

We give the potential function between agent i and its neighboring agent j whose definition is given in [15] as follows:

$$V_{ij} = \frac{(\beta_{ij} - d_{ij}^2)^2}{\beta_{ij}} \quad (15)$$

We also define

$$\rho_{ij} = \frac{\partial V_{ij}}{\partial \|r_{ij}\|^2} = \frac{\|r_{ij}\|^4 - d_{ij}^4}{\|r_{ij}\|^4} \quad (16)$$

$$\nabla_{r_i} V_i(\beta_{ij}) = \sum_{j \in N_i} 2\rho_{ij} r_{ij}$$

The total potential function of agent is given by

$$V_i = \sum_{j \in N_i} V_{ij}(\beta_{ij}) \quad (17)$$

Let $\tilde{r}_i = r_i - r^*$ where $\dot{r}^* = v^*$, $\tilde{v}_i = v_i - v^*$, $\tilde{u}_i = u_i - v^*$ then

$$\begin{cases} \dot{\tilde{r}}_i = \tilde{v}_i \\ \dot{\tilde{v}}_i = \phi_i(\chi_i)\theta_i + \tilde{u}_i + \delta_i + d_{ui} \end{cases} \quad i = 1, 2, 3 \quad (18)$$

Also, let $\tilde{r} = [\tilde{r}_1^T, \tilde{r}_2^T, \tilde{r}_3^T]^T$, $\tilde{v} = [\tilde{v}_1^T, \tilde{v}_2^T, \tilde{v}_3^T]^T$, $\tilde{u} = [\tilde{u}_1^T, \tilde{u}_2^T, \tilde{u}_3^T]^T$, then $\tilde{r}_{ij} = r_{ij}$, $\tilde{v}_{ij} = v_{ij}$ hence, the whole system can be written in the form as

$$\begin{cases} \dot{\tilde{r}} = \tilde{v} \\ \dot{\tilde{v}} = \phi(\chi)\theta + \tilde{u} + \delta + d_u \end{cases} \quad i = 1, 2, 3 \quad (19)$$

In the following, we will focus on the control law design by applying the backstepping design scheme.

Step 1: Starting with the \tilde{r} -subsystem of (18)

$$\dot{\tilde{r}} = \tilde{v} \quad (20)$$

We treat the variable \tilde{v} as a virtual control input. Then, design the feedback controller $\tilde{v} = g(\tilde{r}) = [g_1^T, g_2^T, g_3^T]^T$ to maintain the nominated distance between each pair of agents. Choose the following Lyapunov function:

$$W_1(t) = \sum_{i=1}^3 V_i = \sum_{i=1}^3 \sum_{j \in N_i} V_{ij}(\beta_{ij}) \quad (21)$$

Since $r_{ij} = -r_{ji}$, V_i is symmetric about r_{ij} , we have $\frac{\partial V_{ij}}{\partial r_{ij}} = \frac{\partial V_{ij}}{\partial r_i} = -\frac{\partial V_{ij}}{\partial r_j}$, $j \in N_i$.

Then

$$\dot{W}_1(t) = \frac{d}{dt} \sum_{i=1}^3 V_i = 2 \sum_{i=1}^3 \tilde{v}_i^T \nabla_{r_i} V_i$$

Let the virtual control $\tilde{v} = g(\tilde{r})$ be taken as

$$g_i(\tilde{r}) = - \sum_{j \in N_i} 2\rho_{ij} r_{ij} \quad (22)$$

$$\tilde{v} = g(\tilde{r}) \quad (23)$$

and it follows that

$$\dot{W}_1(t) = 2 \sum_{i=1}^3 \tilde{v}_i^T \nabla_{r_i} V_i = -2 \sum_{i=1}^3 \left\| \sum_{j \in N_i} 2\rho_{ij} r_{ij} \right\|^2 \leq 0 \quad (24)$$

Step 2: Introducing the error variable

$$\bar{v} = [\bar{v}_1^T, \bar{v}_2^T, \bar{v}_3^T]^T = \tilde{v} - g(\tilde{r}) \quad (25)$$

and differentiating it with respect to time, so system (18) can be rewired as

$$\begin{cases} \dot{\tilde{r}} = \bar{v} + g(\tilde{r}) \\ \dot{\bar{v}} = \phi(\chi)\theta + \tilde{u} + \delta + d_u - \dot{g}(\tilde{r}) \end{cases} \quad (26)$$

Now, consider the Lyapunov function

$$W_2(\tilde{r}, \bar{v}) = W_1(t) + \sum_{i=1}^3 (\bar{v}_i^T \bar{v}_i + \frac{1}{\eta_i} (\theta_i - \hat{\theta}_i)^2 + \frac{1}{\lambda_i} (\gamma_i - \hat{\gamma}_i)^2) \quad (27)$$

differentiating the function $W_2(\tilde{r}, \bar{v})$ with respect to time, we have

$$\begin{aligned}\dot{W}_2(\tilde{r}, \bar{v}) = & 2 \sum_{i=1}^3 \left(\sum_{j \in N_i} 2\rho_{ij} r_{ij} \right)^T \bar{v}_i + 2 \sum_{i=1}^3 \left(\sum_{j \in N_i} 2\rho_{ij} r_{ij} \right)^T g_i(\tilde{r}) \\ & + 2 \sum_{i=1}^3 \bar{v}_i^T \phi_i(\chi_i) \theta_i + 2 \sum_{i=1}^3 \bar{v}_i^T \tilde{u}_i + 2 \sum_{i=1}^3 \bar{v}_i^T (\delta_i + d_{ui}) \\ & - 2 \sum_{i=1}^3 \bar{v}_i^T \dot{g}_i(\tilde{r}) - \frac{2}{\eta_i} \sum_{i=1}^3 (\dot{\hat{\theta}}_i(\theta_i - \hat{\theta}_i)) - \frac{2}{\lambda_i} \sum_{i=1}^3 (\dot{\hat{\gamma}}_i(\gamma_i - \hat{\gamma}_i))\end{aligned}\quad (28)$$

Motivated by [16, 17], the formation control law is designed as follows

$$u_1 = -\bar{v}_1 + \dot{g}_1(\tilde{r}) + g_1(\tilde{r}) - k_1 \sqrt{\sum_{j=2,3} \rho_{1j}^2} (a_1 + \text{sgn}(\bar{v}_1)) - \phi_1(\chi_1) \hat{\theta}_1 - B_1 \hat{\gamma}_1 + \dot{v}^* \quad (29)$$

$$u_i = -\bar{v}_i + \dot{g}_i(\tilde{r}) + g_i(\tilde{r}) - \phi_i(\chi_i) \hat{\theta}_i - B_i \hat{\gamma}_i + \dot{v}^*, \quad i = 2, 3 \quad (30)$$

Adaptive law:

$$\dot{\hat{\theta}}_i = \eta_i \phi_i^T(\chi_i) \bar{v}_i \quad (31)$$

$$\dot{\hat{\gamma}}_i = \lambda_i B_i \bar{v}_i \quad (32)$$

Where $a_1 = [\sin(w_1 t) \cos(w_1 t)]^T$ is a unit vector, k_1 is a positive constant. Here, we define the symbolic function as: when $x < 0$, $\text{sgn}(x) = -1$; when $x > 0$, $\text{sgn}(x) = 1$; when $x = 0$, $\text{sgn}(x) = 0$; $\text{sgn}(\bar{v}_i) = [\text{sgn}(\bar{v}_{xi}) \text{ sgn}(\bar{v}_{yi})]^T$; $B_i = \begin{bmatrix} \text{sgn} \bar{v}_{xi} & 0 \\ 0 & \text{sgn} \bar{v}_{yi} \end{bmatrix}$, $\bar{v}_i = [\bar{v}_{xi}, \bar{v}_{yi}]^T$, $i = 1, 2, 3$.

4 Stability Analysis

Consider the system (1), the control law is determined by (29)–(32), and the potential energy function $W_2(\tilde{r}, \bar{v})$ is determined by Eq.(27), then the control strategy can ensure that the multi-robot system forms a certain formation and no collision between any adjacent robots. Note that W_2 is smooth and hence regular, its generalized gradient [18] is a singleton which is equal to its usual gradient everywhere in the state space: $\partial W_2 = \{\nabla W_2\}$ [19]. We can obtain

$$\begin{aligned}\dot{W}_2(\tilde{r}, \bar{v}) = & -2 \sum_{i=1}^3 \left\| \sum_{j \in N_i} 2\rho_{ij} r_{ij} \right\|^2 + 2(-k_1 \sqrt{\sum_{j=2,3} \rho_{1j}^2} (\bar{v}_1^T a_1 + \bar{v}_1^T K[\text{sgn}(\bar{v}_1)]) \\ & + 2 \sum_{i=1}^3 \bar{v}_i^T (\delta_i + d_{ui}) + 2 \sum_{i=1}^3 \bar{v}_i^T (\phi_i(\chi_i) \theta_i - \phi_i(\chi_i) \hat{\theta}_i) - 2 \sum_{i=1}^3 \bar{v}_i^T \bar{v}_i \\ & - 2 \sum_{i=1}^3 \bar{v}_i^T B_i \hat{\gamma}_i - \frac{2}{\eta_i} \sum_{i=1}^3 (\dot{\hat{\theta}}_i(\theta_i - \hat{\theta}_i)) - \frac{2}{\lambda_i} \sum_{i=1}^3 (\dot{\hat{\gamma}}_i(\gamma_i - \hat{\gamma}_i)))\end{aligned}\quad (33)$$

Where $K[f](x)$ is called *Filippov* set-valued mapping [20]. In the above consequence (33) we have used Theorem 1 in [21] to figure the inclusions of the *Filippov* set. On account of $K[\text{sgn}(x)](x) = |x|$, we have

$$\begin{aligned} \dot{W}_2(\tilde{r}, \bar{v}) &\leq -2 \sum_{i=1}^3 \left\| \sum_{j \in N_i} 2\rho_{ij} r_{ij} \right\|^2 - 2 \sum_{i=1}^3 \|\bar{v}_i\|^2 \\ &\leq 0 \end{aligned} \quad (34)$$

When $\dot{\theta}_i = \eta_i \phi_i^T(\chi_i) \bar{v}_i$, $\dot{\gamma}_i = \lambda_i B_i \bar{v}_i$, the above formula is established.

It can be seen from Eq. (34) that the function is non-increasing, so the following lemma is obtained:

Lemma 1. Consider system (1) driven by the controller (29)–(32), and starting from any initial state in the invariant set $S = \{r_{ij}, \bar{v}_i \mid W_2(t) \leq W_2(0)\}$. Then it holds that no collision occurs between any two adjacent agents.

Proof. When $W_2(0) < \infty$, the continuity of $W_2(t)$ shows that the set $\{r_{ij}, \bar{v}_i\}$ is closed, then the set S is a compact set. While $\dot{W}_2(t) \leq 0$, then $W_2(t) \leq W_2(0)$, $\forall t > 0$; On the basis of the definition of positively invariant set [22], then $S = \{r_{ij}, \bar{v}_i \mid W_2(t) \leq W_2(0)\}$ is the invariant set. For any $r_{ij} \in S$, the time derivative of $W_2(t)$ remains non-positive for $\forall t \geq 0$ from (34). In consequence, $W_2(t) \leq W_2(0) < \infty$ for all $\forall t \geq 0$. So that

$$\frac{-\sqrt{W_2(0)} + \sqrt{W_2(0) + 4d_{ij}^2}}{2} \leq \|r_{ij}\| \leq \frac{\sqrt{W_2(0)} + \sqrt{W_2(0) + 4d_{ij}^2}}{2}$$

It is easily seen that $\frac{-\sqrt{W_2(0)} + \sqrt{W_2(0) + 4d_{ij}^2}}{2}$ is strictly positive. Therefore $\|r_{ij}\| > 0, \forall t \geq 0$, no collision occurs between any two adjacent agents.

Lemma 2. Consider the system (1), the control law is determined by (29)–(32), the system converges to a unique set of equilibrium points

$$M = \{r_{ij}, \bar{v}_i \mid \rho_{ij} = 0, \bar{v}_i = 0, j \in N_i, i = 1, 2, 3\}$$

Proof. Define $\tilde{\gamma}_i = \gamma_i - \hat{\gamma}_i$, $\tilde{\theta}_i = \theta_i - \hat{\theta}_i$, because $W_2(t) \in L_\infty$, that is $W_2(+\infty)$ exists, then $\bar{v}_i, \tilde{v}_i, \tilde{\theta}_i, \tilde{\gamma}_i, r_i, i = 1, 2, 3$ is bounded, so $\tilde{v}_i, \dot{\tilde{v}}_i \in L_\infty$; further, there are $\int_0^{+\infty} \left\| \sum_{j \in N_i} \rho_{ij} r_{ij} \right\|^2 dt \leq W_2(0) - W_2(+\infty) < +\infty$ and $\int_0^{+\infty} \tilde{v}_i^2 dt \leq W_2(0) - W_2(+\infty) < +\infty$; and $\bar{v}_i, \dot{\bar{v}}_i \in L_\infty$, $\|r_{ij}(t)\| \neq 0$, so we can know from Barbalat's lemma [23] that $\lim_{t \rightarrow \infty} \sum_{j \in N_i} \rho_{ij} r_{ij} = 0$, $\lim_{t \rightarrow \infty} \bar{v}_i(t) = 0$, $i = 1, 2, 3$

Hence, $\lim_{t \rightarrow \infty} g_i(\tilde{r}) = -\lim_{t \rightarrow \infty} \sum_{j \in N_i} 2\rho_{ij} r_{ij} = 0$. It can be seen from Eq. (25), $\lim_{t \rightarrow \infty} \tilde{v}_i(t) = 0$. That is $\dot{\tilde{v}}_i = \dot{v}_i - v^* = \dot{v}_i$

$$\lim_{t \rightarrow \infty} \tilde{v}(t) = 0 \Rightarrow \lim_{t \rightarrow \infty} (v_i - v^*) = 0 \Rightarrow \lim_{t \rightarrow \infty} v_i = v^*$$

$$\begin{aligned}\dot{\tilde{v}}_1 &= -\bar{v}_1 + \dot{g}_1(\tilde{r}) + g_1(\tilde{r}) \\ &\quad - k_1 \sqrt{\sum_{j=2,3} \rho_{1j}^2 (a_1 + \text{sgn}(\bar{v}_1))} + \phi_1(\chi_1)\theta_1 + \delta_1 + d_{u1} - \phi_1(\chi_1)\hat{\theta}_1 - B_1\hat{\gamma}_1\end{aligned}$$

Because the speed is a constant, the acceleration is zero, i.e. $\lim_{t \rightarrow \infty} \dot{v}_i = \lim_{t \rightarrow \infty} (\dot{v}_i - \dot{v}^*) = \lim_{t \rightarrow \infty} \dot{v}_i = 0$. Since a_1 is a periodic function changing with t , we have that $\sqrt{\sum_{j=2,3} \rho_{1j}^2 (a_1 + \text{sgn}(\bar{v}_1))} = 0 \Rightarrow \sqrt{\sum_{j=2,3} \rho_{1j}^2} = 0 \Rightarrow \rho_{12} = 0, \rho_{13} = 0$.

Moreover, since $\lim_{t \rightarrow \infty} \sum_{j \in N_i} \rho_{ij} r_{ij} = 0 \Rightarrow \rho_{23} = 0$, which means that $\rho_{ij} = \rho_{ji}, j \in N_i \Rightarrow \rho_{ij} = 0$.

In summary, the set of balance points of the system is the set of expected balance points: $M = \{r_{ij}, \bar{v}_i \mid \rho_{ij} = 0, \bar{v}_i = 0, j \in N_i, i = 1, 2, 3\}$. That is, all agents reach the desired distance, and the velocity of each agent reaches the desired value. Thus, multiple agents can achieve the desired formation.

5 Simulations

Consider the following class of multi-agents system with uncertain nonlinear dynamics and external disturbances, whose dynamic equation is:

$$\begin{cases} \dot{r}_i = v_i \\ \dot{v}_i = r_i \frac{1-e^{-|v_i|}}{1+e^{-|v_i|}} + u_i + 1.2 \cos r_i \end{cases}$$

We choose $D_{ui} = 1.2$, the number of the fuzzy system rules $M = N = 10$, $\hat{\gamma}(0) = 0.5$, initial adjustable parameter $\hat{\theta}_{xi}(0) = 0.5$, $\hat{\theta}_{yi}(0) = 0.5$, $\lambda_1 = \eta_1 = 2$, $\lambda_2 = \eta_2 = 2$, $\lambda_3 = \eta_3 = 2$.

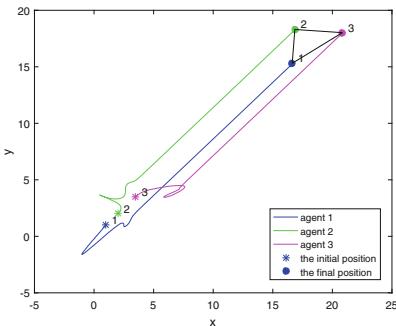


Fig. 1. The trajectories of agents

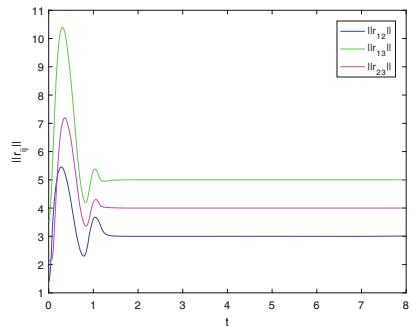
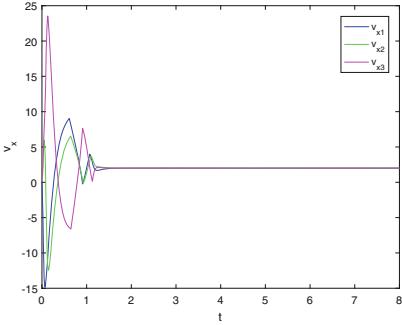
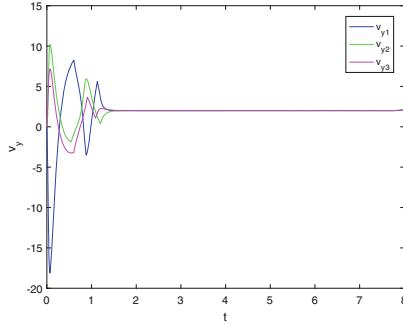


Fig. 2. Distance between any agents $\|r_{ij}\|$

If the initial positions of the agents are set to $r_1 = [1, 1]^T$, $r_2 = [2, 2]^T$ and $r_3 = [3.5, 3.5]^T$, which means that the three agents are collinear and located

**Fig. 3.** Agents' velocity v_{xi} **Fig. 4.** Agents' velocity v_{yi}

so close to each other, we choose $w_1 = 1$, $k_1 = 1.4$. $D = [d_{12}, d_{23}, d_{13}]^T$ is the desired distance vector, where $d_{12} = 3$, $d_{23} = 4$, $d_{13} = 5$, and the desired velocity is $v^* = 2$. The results show that the three agents have reached the desired globally stable formation, and all agents are moving at the same desired speed.

Due to space constraints, we present only one simulation result, which also proves the effectiveness of the proposed algorithm in another line formation shape (Figs. 1, 2, 3 and 4).

6 Conclusions

This paper is devoted to studying the fuzzy adaptive formation control problem of a class of nonlinear systems with uncertain dynamics and external disturbances. First, nonlinear dynamics are approximated by the first type fuzzy systems. Furthermore, the parameter adaptive estimation method is introduced to achieve the purpose of eliminating the effects of approximation error and external disturbances. According to the potential energy function, the back-stepping technique and the adaptive perturbation method are used to resolve the global stability of the overall fuzzy formation system, and all agents' velocities converge to a common value with no collision between any agents. In the end, simulations are given to demonstrated the effectiveness of the proposed control algorithm.

Acknowledgment. This work was supported in part by the National Natural Science Foundation of China under grants 61803331, 61873229, 61806175 and 61873346, in part by the Jiangsu Planned Projects for Postdoctoral Research Funds 1601024B and the Natural Science Foundation of Jiangsu Province BK20170515.

References

1. Shen, Q.K., Shi, P., Zhu, J.W.: Adaptive consensus control of leader-following systems with transmission nonlinearities. *Int. J. Control.* **9**(2), 317–328 (2019)
2. Verginis, C.K., Nikou, A., Dimarogonas, D.V.: Robust formation control in SE (3) for tree-graph structures with prescribed transient and steady state performance. *Automatica* **103**, 538–548 (2019)
3. Han, D.K., Dimitra, P.: Robust multitask formation control via parametric lyapunov-like barrier functions. *IEEE Trans. Autom. Control* **64**(11), 4439–4453 (2019)
4. Sun, Z.Y., Marina, H.G., Seyboth, G.S., Anderson, B.D.O., Yu, C.B.: Circular formation control of multiple unicycle-type agents with nonidentical constant speeds. *IEEE Trans. Control Syst. Technol.* **27**(1), 192–205 (2018)
5. Tian, Y.P., Wang, Q.: Global stabilization of rigid formation in the plane. *Automatica* **49**(5), 1436–1441 (2013)
6. Marina, H.G., Jayawardhana, B., Cao, M.: Distributed scaling control of rigid formations. In: IEEE 55th Conference Decision and Control (CDC), pp. 5140–5145 (2016)
7. Huang, H., Yu, C.B., Wang, X.K.: Control of triangular formations with a time-varying scale function. In: Proceedings of the American control conference, pp. 4828–4833 (2011)
8. Trinh, M.H., Zhao, S., Sun, Z., Zelazo, D., Anderson, B.D.O., Ahn, H.S.: Bearing-based formation control of a group of agents with leader-first follower structure. *IEEE Trans. Autom. Control* **64**(2), 598–613 (2019)
9. Zhao, S.Y., Li, Z.H., Ding, Z.T.: Bearing-only formation tracking control of multi-agent systems. *IEEE Trans. Autom. Control* **64**(11), 4541–4554 (2019)
10. Liu, G.P., Zhang, S.J.: A survey on formation control of small satellites. *Proc. IEEE* **106**(3), 440–457 (2018)
11. Rimon, E., Koditschek, D.E.: Exact robot navigation using artificial potential functions. *IEEE Trans. Robot. Autom.* **8**(5), 501–518 (1992)
12. Koditschek, D.E., Rimon, E.: Robot navigation functions on manifolds with boundary. *Adv. Appl. Math.* **11**(4), 412–442 (1990)
13. Tanner, H.G., Jadbabaie, A., Pappas, G.T.: Stable flocking of mobile agents part II: dynamic topology. In: Proceedings of the 42nd IEEE Conference of Decision and Control, pp. 2010–2015 (2003)
14. Jin, T., Liu, Z.W.: Cluster formation for multi-agent system under disturbances and unmodelled uncertainties. *IET Control Theory Appl.* **11**(15), 1630–2635 (2017)
15. Dimarogonas, D.V., Johansson, K.H.: Further results on the stability of distance-based multi-robot formations. In: American Control Conference, pp. 2972–2977 (2009)
16. Wang, Q., Li, Y., et al.: Globally stable rigid formation control for multi-robot systems. In: Proceedings of the Thirty-Fourth Chinese Control Conference, pp. 7505–7510 (2015)
17. Krstic, M.: Nonlinear and Adaptive Control Design. Wiley, New York (1995)
18. Clarke, F.: Optimization and Nonsmooth Analysis. Addison-Wesley, Reading (1983)
19. Shevitz, D., Paden, B.: Lyapunov stability theory of nonsmooth systems. *IEEE Trans. Autom. Control* **49**(9), 1910–1914 (1994)
20. Filippov, A.: Differential Equations with Discontinuous Right-Hand Sides. Kluwer, Norwell (1988)

21. Paden, B., Sastry, S.S.: A calculus for computing Filippov's differential inclusion with application to the variable structure control of robot manipulators. *IEEE Trans. Circuits Syst.* **34**(1), 73–82 (1987)
22. Krstic, M., Kanellakopoulos, I., Kokotovic, P.V.: *Nonlinear and Adaptive Control Design*. Wiley, New York (1995)
23. Ge, S.S., Hang, C.C., Lee, T.H., Zhang, T.: *Stable Adaptive Neural Network Control*. Kluwer Academic Publishers, New York (2002)



Deep Learning Based Pathologic Images Recognition Upon Invasive Bladder Cancer

Tong Li^{1(✉)} and Qi Zhang^{1,2}

¹ School of Information Technology and Management, University of International Business and Economics, Beijing 100029, China
uibebitong@163.com

² Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China

Abstract. Cystoscope image is an important reference for bladder cancer diagnosis. Considering that intelligent analysis of the cystoscopic helps to improve the diagnose efficiency of the doctors, we propose a deep learning based prediction model in this paper to classify invasive and noninvasive bladder cancers. Our model adopts four feature extraction layers based on CNN, and achieves good classification performance through continuous optimization of model parameters. The experimental results show that the model recognition accuracy can reach up to 88.24%, and it can effectively distinguish invasive bladder cancer from non-invasive bladder cancer.

Keywords: Deep learning · Cystoscopic image analysis · CNN

1 Introduction

Nowadays, cancer has become a major killer of human health and life [1]. Bladder cancer is one of the top 10 most common tumors, and the most common malignancy of the urinary system. In China, the incidence of bladder cancer ranks as the most common of genitourinary tumors, and in western countries, it is only inferior to prostate cancer [2]. To make a definite diagnosis of bladder cancer, in clinical practice, cystoscopy is the most important method.

However, in most hospitals, the recognition and analysis of cystoscopic images data are performed by the naked eyes of doctors. This method is not only time-consuming but also encountering the misdiagnosis problem. Different doctors may give completely different results for the same picture. For young doctors, it is more likely to make misdiagnosis upon the bladder tumors [3]. Both the aforementioned factors lead to the incidence of misdiagnosis and missed-diagnosis. Due to these conditions, the classification of bladder cancer pathological images by means of artificial intelligence is becoming an important issue that required to be addressed.

The treatment of bladder cancer varies greatly according to the invasion of bladder tumor [4]. Non-muscle-invasive bladder cancer is usually treated by electric resection, while muscle-invasive bladder cancer is often treated by total bladder resection [5]. At present, cystoscopy can only observe the inside of the bladder tumor, but it is not able to determine the depth of infiltration. Therefore, the patients must undergo another diagnostic electrotomy after cystoscopy to complete the resection of the tumor and the basal tissue of the tumor [6]. This severely increases the burden on the patients.

In recent years, imaging diagnosis utilizing Deep Learning (DL) has been developed in the field of medicine [7]. Endoscopic imaging-based diagnosis using DL has been clinically applied in the field of gastroenterology, but its application to the field of urology is initiated until recently [8]. Atsushi Ikeda et al. put forward an artificial intelligence based cystoscopic diagnosis support system for bladder cancer [9]. In addition, Okyaz Eminaga et al. used deep convolution neural network to classify cystoscopic images [10]. These results prove the effectiveness of introducing deep learning, but they have not considered the issue of infiltration. Due to these considerations, in this paper, the deep learning technology will be applied to classify invasive and non-invasive bladder cancer.

In face of the practical problems encountered in the diagnosis of bladder cancer, we propose a convolutional neural network (CNN) based recognition model in this paper to classify the invasive tumors from noninvasive ones. In addition, the SVM model has been widely used in image recognition due to its high efficiency and applicability [11]. In order to test the performance of the model, we use SVM to conduct a comparison experiment on the data set in this paper. The experimental results verify that applying the deep learning techniques in clinical practice for bladder tumor recognition helps to realize the morphological diagnosis of bladder cancer with different degrees of invasion under cystoscope.

2 Preliminaries

The images data selected in this paper are the actual cystoscope images data from the department of urology, Peking University third hospital from June 2018 to January 2020.

2.1 Classify the Pictures

Based on the data from the third hospital of Peking University, cystoscope pathological images are divided into non-invasive bladder cancer and invasive bladder cancer. After screening, 254 invasive images and 545 non-invasive images are obtained, totaling 799.

2.2 Data Augmentation

Since deep learning requires a large number of images, but the size of the original data set is relatively small, it is necessary to enlarge the original data set. In this

paper, the methods of image flipping, rotating random angle, adding random colors, enhancing contrast and enhancing brightness are adopted to obtain more images. Finally, 1778 invasive and 2180 non-invasive images are obtained.

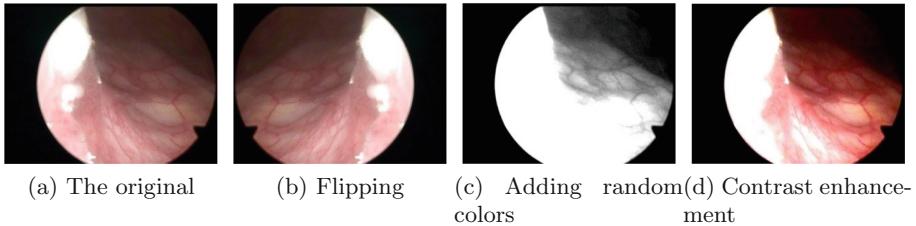


Fig. 1. (a) is the original cystoscopy image. (b)–(d) are the images obtained after flipping, adding random colors and enhancing contrast, respectively.

2.3 Experimental Environment

In terms of hardware, this paper adopts Intel(R)Core(TM) i7-8565u cpu 1.80 GHz server, 8 GB running memory, Windows 10 operating system, among which the graphics card of the data research and development machine is NVIDIA GeForce MX250. Using the Tensorflow framework.

3 Methods

CNN and SVM are widely used image recognition methods, and this paper will study on the basis of these two methods [12]. CNN is a deep learning network framework that imitates biological natural visual cognitive mechanism and adopts image convolution computation [13]. It is currently the most popular image recognition method. At the same time, SVM is a new approach to machine learning based on optimization theory and is the linear classifier with the largest spacing defined in the feature space [14].

3.1 CNN

CNN is mainly composed of convolution layer, activation layer, pooling layer and fully connected layer, and its structure diagram is shown in the following figure [15]. Through the combination of a series of convolutional layers and pooling layers, CNN enables the network to have a certain degree of displacement, scaling and nonlinear deformation stability. Finally through the full connection layer output results [16]. Internally, CNN uses the gradient descent method to minimize the loss function to reverse adjust the weight in the network, and at the same time improve the accuracy of the neural network through continuous iterative training (Fig. 2).

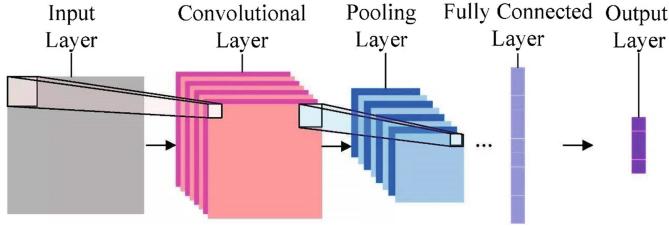


Fig. 2. Schematic diagram of CNN

Convolution Layer. The convolution layer consists of a number of convolution kernels. Convolution kernel operation is the operation between kernel function and local part of image, which is essentially the discrete convolution operation between two two-dimensional matrices. The operation principle is as follows:

$$\mathbf{c}(p, q) = \mathbf{a}(p, q)\mathbf{b}(p, q) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \mathbf{a}(p-i, q-j) \mathbf{b}(i, j), \quad (1)$$

where \mathbf{a} is a two-dimensional matrix of gh , and \mathbf{b} is a two-dimensional matrix of rs convolved with each other. $M = \text{Max}(g, r)$, $N = \text{Max}(h, s)$.

CNN obtains the characteristic image by convolution layer operation [18]. Each convolution layer has multiple convolution kernel functions. Through the way of weight sharing, the same convolution kernel shares the same group of weights when it performs operations with different local images in a certain step length, so as to reduce the calculation parameters and the calculation amount.

Pooling Layer. The pooling layer, also known as the sampling layer, is an operation that aggregates features of different locations within a specific range of an image. After convolution operation, the image has local features, and each pixel region generally has a high similarity. The average or maximum value of features in the image region can be used to represent the overall features of the region, which are respectively called mean pooling and maximum pooling. In this paper, the maximum pooling is adopted, and its mathematical formula is expressed as follows:

$$P = \max_{w \times w} \{\mathbf{A}^{l \times l}\}, \quad (2)$$

where P is the feature matrix obtained from the maximum pooling, l is the width of the feature graph, \mathbf{A} is the feature matrix activated by the convolution layer, and w is the width of the pooling region.

Full Connection Layer. Each node of the full connection layer is connected with all nodes of the previous layer, so the features extracted before can be integrated. In addition, through the full connection layer, the multi-layer convolution and pooling feature map in CNN is tiled into one-dimensional vector input, then the input of each layer is calculated by formula (3).

$$\mathbf{x}^l = f(\mathbf{w}^i \mathbf{x}^{l-1} + \mathbf{b}^l), \quad (3)$$

where \mathbf{x}^l is the weight coefficient of the fully connected network, \mathbf{x}^{l-1} is the feature graph, and \mathbf{b}^l is the bias term of the fully connected layer.

Activation Function. By introducing activation function, the nonlinearity of the neural network model can be increased, so as to enhance the ability to fit the nonlinear relationship of samples. Common activation functions include ReLU, Sigmoid and tanh:

$$ReLU = \max(0, x), \quad (4)$$

$$Sigmoid = \frac{1}{1 + e^{-x}}, \quad (5)$$

$$\tanh = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (6)$$

The CNN model designed in this paper increases the nonlinearity of the model by adding the ReLU activation function in the convolution layer and the fully connected layer, thus improving the learning ability of the model.

3.2 SVM

The basic idea of SVM learning is to solve the separation hyperplane which can divide the training data set correctly and has the largest geometric interval [17]. According to the Lagrange formula, the optimal separation hyperplane can be changed to the boundary of the decision, as shown in Eq. (7):

$$d(X_t) = \sum_{i=1}^n y_i p_i X_i X^t + q_0, \quad (7)$$

where X_t is the test tuple, y_i is the class label of support vector X_i , p_i and q_0 are the optimized numerical parameters, and n is the number of support vectors (Fig. 3).

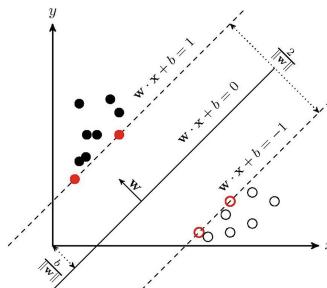


Fig. 3. SVM schematic diagram

4 Experiments

In this part, two series of experiments are conducted by CNN and SVM respectively upon the cystoscope pathological images.

4.1 The Data Set

The enhanced data are used for training the network model. Samples of each class are divided into training set, verification set and test set in a certain proportion. The experimental data are shown in Table 1.

Table 1. Sample size distribution of cystoscope image data set

Type	The training set	Validation set	The test set
Invasive	1000	400	378
Non-invasive	1100	500	500

4.2 Establishment of Network Model

In the data preprocessing part, the JPG files are decoded into RGB pixel grid, and then these pixel grids are transformed into floating-point tensor and normalized.

The above data sets are substituted into the designed convolutional neural network model. The convolutional neural network (CNN) consists of 1 input layer (400, 400, 3), 4 feature extraction layers, 1 fully connected layer and 1 output layer. Feature extraction layer is a connection structure composed of batch normalization layer, relu activation function, maximum pooling layer and convolution layer. In terms of parameter setting: output type is 2 (invasive and non-invasive), batch_size is 20, learning rate is 0.0001, loss function loss is binary_crossentropy, optimizer is RMSprop, epoch is 30, dropout is 0.5.

In order to evaluate the validity of the algorithm model for feature recognition of pathological image data, recording the number of epoch data per model and the number of correctly classified data are counted, and the accuracy A and loss L of the model are formed by the ratio of the two to verify the stability of the model. In formula (7), n represents the number of total data, and n_c represents the number of correct classification. The function adopted is the cross entropy function f, y is the category predicted by the model, and x is the correct label value.

$$\begin{cases} A = n_c/n, \\ L = f(x, y). \end{cases} \quad (8)$$

In Fig. 4 below, (a) shows the variation trend of accuracy of the training data set and the verification data set. It can be seen from the figure that with the increase of the number of iterations, the accuracy of the training set gradually

improves to 98%, and the accuracy of the verification set hovers around 60%. (b) shows the loss value and changing trend of the training data set and the verification data set. After 10 generations of training, the loss value of the test set gradually approaches 0, and the loss value of the verification fluctuates (Fig. 4).

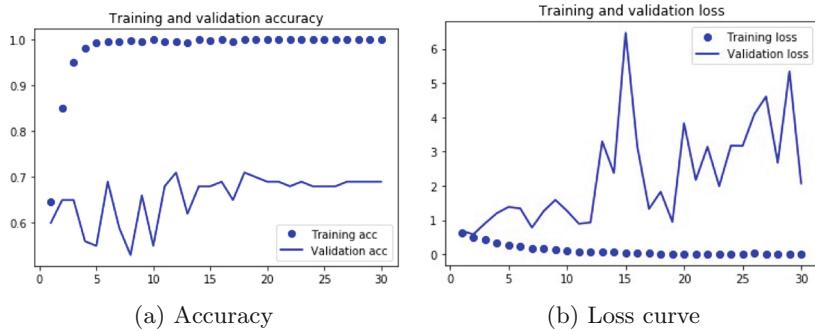


Fig. 4. Accuracy curve and loss curve.

According to the results in Table 2, the accuracy of test set prediction is 88.24%, which shows that the convolution neural network model used in this paper can distinguish the samples of invasive bladder cancer from non-invasive ones effectively. Due to the small number of original data sets, the lack of detailed image segmentation and manual marking of training sets, the deviation values within a certain range are in the acceptable range.

Table 2. Comparison of the two models

Model	Accuracy/%	Time consuming
CNN	88.24	4 h and 15 min
SVM	61.7	4 h and 15 min

4.3 Comparative Analysis

For the SVM model, the numbers of invasive and non-invasive training samples are 1399 and 1649 respectively, and the testing samples are 494 and 549 respectively. Through experiments, the obtained classification results of the two models are shown in Table 2.

It can be seen from Table 2 that the accuracy of CNN-based model is much higher than that of the SVM model, and the classification effect is obviously better than that of the SVM model. These results indicate that the CNN-based model can achieve better training accuracy and invasive cancer classification performance.

5 Conclusion

This study shows that a deep CNN can be used for diagnostic classification of cystoscopic images. After images enhancement and preprocessing, we establish the CNN-based classification model, and compare it with the SVM model. The experimental results show that the model presented in this paper achieves good accuracy, efficiently distinguishing the samples of invasive bladder cancer from non-invasive ones. This indicate that applying the deep learning model in bladder tumor recognition is of much help to assist doctors in the clinical medical diagnosis of bladder cancer. Further research will be carried out on improving the accuracy of discrimination and reducing the loss value.

Acknowledgement. This paper is funded by Beijing Natural Science Foundation (19L2037) and the Excellent Young Scholars Funding Project in UIBE (No. 19YQ10).

References

1. Takahashi, H., Tampo, H., Arai, Y., et al.: Applying artificial intelligence to disease staging: deep learning for improved staging of diabetic retinopathy. *PLoS ONE* **12**, e0179790 (2017)
2. National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov>
3. Li, H., et al.: Segmentation of brain tumors in multiparametric MR images via robust statistic information propagation. In: Proceeding of Asian Conference on Computer Vision (ACCV), vol. 6495, pp. 606–617 (2011)
4. Hoglund, M.: On the origin of syn- and metachronous urothelial carcinomas. *Eur. Urol.* **51**, 1185–1193 (2007)
5. Herr, H.W., Shipley, W.U., Bajorin, D.F.: Cancer of the bladder. In: Devita, V.T., Hellman, S., Rosenberg, S.A. (eds.) *Cancer: Principles and Practice of Oncology*, 6th edn. Lippincott Williams and Wilkins, Philadelphia, pp. 1005-1020 (2001)
6. Samplaski, M.K., Jones, J.S.: Two centuries of cystoscopy: the development of imaging, instrumentation and synergistic technologies. *BJU Int.* **103**, 154–158 (2009)
7. Hirasawa, T., Aoyama, K., Tanimoto, T., et al.: Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer* **21**, 653–660 (2018)
8. Nosato, H., Sakanishi, H., Takahashi, E., et al.: Image retrieval method for multiscale objects from optical colonoscopy images. *Int. J. Biomed. Imaging* **2017**, 7089213 (2017)
9. Ikeda, A., Nosato, H., Kochi, Y.: Support system of cystoscopic diagnosis for bladder cancer based on artificial intelligence. *J. Endourol.* **34**(3), 352–358 (2020). <https://doi.org/10.1089/end.2019.0509>
10. Eminaga, O., Eminaga, N., Semjonow, A.: Diagnostic classification of cystoscopic images using deep convolutional neural networks. *JCO Clin. Cancer Inf.* **2**, 1–8 (2018)
11. Xiaoji, W.: Research on image classification based on SVM. Nanjing University of Information Technology, Nanjing (2011)
12. Le, M.H., Chen, J., Wang, L., et al.: Automated diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks. *Phys. Med. Biol.* **62**, 6497 (2017)

13. Girshick, R., Donahue, J., Darrell, T., et al.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587. IEEE, Columbus (2014)
14. Lin, N.: Design and implementation of SVM-based public security intelligence automatic classification system. PLA University of Information Engineering, Zhengzhou (2007)
15. Le, C.Y., Boser, B.E., Denker, J.S., et al.: Handwritten digit recognition with a back-propagation network. In: Advances in Neural Information Processing Systems, pp. 396–404 (1990)
16. Krizhevsky, A., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105. IEEE, Lake Tahoe (2012)
17. Xue-gong, Z.: On statistical learning theory and support vector machine. *J. Autom.* **26**(1), 32–42 (2000)
18. Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**(1), 106–154 (1962)



An Autonomous Visual-Inertial-Based Navigation System for Quadrotor

Sansan Zhao, Rui Li^(✉), Yingjing Shi, and Haoran Li

School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, People's Republic of China
lirui@uestc.edu.cn

Abstract. In this paper, we present a practical autonomous navigation system based on the visual-inertial of a quadrotor. Due to the practical engineering requirement of improving the applicability of the advanced visual-inertial navigation fusion algorithm and 3D mapping algorithm, we realize the on-line 3D trajectory planning and tracking control algorithm with full consideration of UAV dynamics design, and finally complete the quadrotor autonomous navigation system consisting of UAV, upper computer and other software and hardware components. The feasibility is verified by actual flight experiments. The results show that the quadrotor autonomous navigation system can achieve high-precision positioning, online 3D reconstruction and dynamic autonomous navigation in a complex unknown environment without GPS. The system has good accuracy and robustness in real-time, which provides a strong technical support for the subsequent expansion of platform function.

Keywords: Quadrotor · Visual-inertial · TSDF · Trajectory planning · Autonomous navigation

1 Introduction

With the intelligent development and application of UAV, more and more attention has been paid to autonomous navigation, which is an important application of the intelligent realization of UAV. Accurate position and attitude estimation is the premises of the autonomous navigation of UAV. Common positioning methods include such as global positioning system (GPS), motion capture system, lidar, and vision, etc. Compared with other positioning methods, the visual method has the advantages of small size, low power consumption, low price, and abundant scene information, and gradually becomes the mainstream of mobile robot positioning methods. However, the stability and robustness of visual-only Simultaneous Localization And Mapping (SLAM) [14] system need to be improved in the environment of illumination variations and motion blur. To make up for the shortcomings of visual-only SLAM system, and considering the complementarity of Inertial Measurement Unit (IMU) and visual sensor, a

Visual-Inertial Odometry (VIO) with better stability, robustness and accuracy are obtained by fusing the two.

Motion planning is also an essential part of the UAV integrated navigation system. With the diversification of UAV application scenarios, the path planning scheme of common robots cannot meet the rapid navigation requirements of UAV in the 3D unknown environment. The complex movement and dynamics of the flight system, online real-time planning ability, safety feasibility, and integrity of planning results need to be considered in motion planning.

Scholarly works on UAV navigation are extensive. As early as 2011, Kumar et al. show autonomous micro quadrotor navigation integrated by multi-sensor such as monocular camera, IMU, and lidar [12]. In 2018, his team participated in the FLA project carried out by the U.S. Department of defense, using binocular MSCKF to provide UAV with high-precision state estimation, combined with the secondary planning method of Safe Flight Corridor (SFC) to generate flight trajectory, achieving 17.5m/s quadrotor high-speed autonomous navigation [13]. In 2016, a real-time method for local trajectory planning of the micro quadrotor was proposed by the Automation System Laboratory of the Swiss Federal Institute of Technology Zurich, and a complete obstacle avoidance navigation system was displayed on the quadrotor platform equipped with a visual-inertial navigation kit [10]. In 2017, the Technical University of Munich represents the local planning problem as a B-spline optimization problem to improve the success rate of trajectory generation and complete the real flight experiment in the airborne RGBD aircraft [16]. In 2017, ACL Laboratory of Massachusetts Institute of Technology adopt a new obstacle avoidance planning algorithm based on RGBD camera in the FLA project to realize fast obstacle avoidance navigation in an unknown environment with limited sensing range [6]. The Aerial Robotics Group of Hong Kong University of Science and Technology successively complete the loosely-coupled and tightly-coupled experiments of the monocular camera and IMU data, which is used for the pose estimation of quadrotor to achieve accurate self-positioning [5]. The Reliable Flight Control group of Beijing University of Aeronautics and Astronautics is committed to robust positioning and navigation technology. Heng et al. design the estimation system combining monocular camera, IMU, and ultrasonic information, and complete the precise hovering and straight-line stable flight experiment of UAV [1].

The contribution of this paper is to design and implement an autonomous navigation system platform of quadrotor based on visual-inertial fusion, which can realize high-precision positioning, online 3D reconstruction, and dynamic autonomous navigation in the non-cooperative complex unknown environment without GPS. This paper is organized as follows: the third chapter is concerned with the methodology employed for this study. It introduces the algorithm implementation of positioning, mapping and planning to build the autonomous navigation system of this paper; the fourth chapter uses the hardware platform of the autonomous navigation system of the quadrotor to carry out the actual flight experiments of indoor and outdoor navigation to verify the actual navigation flight effect of this system.

2 Visual-Inertial Odometry

The framework of visual-inertial odometry refers to the VINS-Mono [11] of HKUST, which is divided into six parts: vision processing front-end, IMU pre-integration, visual-Inertial initialization, back-end optimization, loop detection, and global pose optimization.

2.1 Vision Processing Front-End

To solve the problems of Harris corner feature threshold selection and clustering in VINS-Mono, this paper uses the sparse KLT (Kanade-Lucas-Tomasi) [15] optical flow method based on GFTT corner extraction, which ensures the number of corner extraction is moderate and evenly distributed in the image space. Firstly, the response value of each corner in the image is calculated, and the corner larger than the preset threshold is selected. Then, within a certain range of each corner as the center, the corner with the largest response value is retained. After feature extraction, the KLT optical flow method is used for tracking.

2.2 IMU Pre-integration

Usually, the frequency of IMU is much higher than that of the camera. If all the measured values of IMU are added to the constraints, it will waste a lot of computational resources, so it is very necessary for IMU Pre-integration. IMU provides the relative measurement of two keyframes, to construct the error function to iteratively optimize the pose of keyframes.

From the measurement model of IMU, the derivative of position, velocity, and rotation (PVQ) to time can be written as:

$$\dot{p}_{wb_t} = v_t^w \quad (1)$$

$$\dot{v}_t^w = a_t^w \quad (2)$$

$$\dot{q}_{wb_t} = q_{wb_t} \otimes \begin{bmatrix} 0 \\ \frac{1}{2}\omega^{b_t} \end{bmatrix} \quad (3)$$

Where the superscript b is in the IMU body frame, w is in the world frame, q_{bw} represents the transformation quaternion from the world frame to the IMU body frame.

Due to the state propagation of IMU is related to PVQ at the i^{th} moment, when the initial state changes at the i^{th} moment, it needs to be re-integrated and adjusted. To avoid this situation, the pre-integration model is adopted in this paper.

$$q_{wb_t} = q_{wb_i} \otimes q_{b_i b_t} \quad (4)$$

We transform the reference frame from the world frame to the local frame b_i . There are the following forms:

$$p_{wb_j} = p_{wb_i} + v_i^w \Delta t - \frac{1}{2} g^w \Delta t^2 + q_{wb_i} \iint_{t \in [i,j]} (q_{b_i b_t} a^{b_t}) \delta t^2 \quad (5)$$

$$v_j^w = v_i^w - g^w \Delta t + q_{wb_i} \int_{t \in [i,j]} (q_{b_i b_t} a^{b_t}) \delta t \quad (6)$$

$$q_{wb_j} = q_{wb_i} \int_{t \in [i,j]} q_{b_i b_t} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} \omega^{b_t} \end{bmatrix} \delta t \quad (7)$$

where

$$\alpha_{b_i b_j} = \iint_{t \in [i,j]} (q_{b_i b_t} a^{b_t}) \delta t^2 \quad (8)$$

$$\beta_{b_i b_j} = \int_{t \in [i,j]} (q_{b_i b_t} a^{b_t}) \delta t \quad (9)$$

$$\gamma_{b_i b_j} = \int_{t \in [i,j]} q_{b_i b_t} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} \omega^{b_t} \end{bmatrix} \delta t \quad (10)$$

The pre-integration component is only related to the IMU measurement.

In this paper, the system uses the mid-value integration method to pre-integrate in discrete time, that is to say, the pose of K to K+1 at two adjacent times is calculated by the average of the measured values a and w at two moments.

2.3 Visual-Inertial Initialization

Tightly-coupled visual-inertial fusion is a highly nonlinear system, which relies heavily on a robust initialization process. The joint initialization adopts a loosely-coupled method, including visual SFM and visual-inertial alignment.

Firstly, the camera scale pose and feature location map are estimated by using visual-only SFM, and generally, about 10 frames of images are selected for calculation. Then, the extrinsic parameters between the camera and the IMU can be calculated according to the relative rotation of the continuous images. By aligning the results of IMU pre-integration and visual-only SFM, the scale, gravity, velocity, and bias of IMU can be roughly recovered.

2.4 Sliding Window Back-End Optimization

Although the method based on nonlinear optimization in a tightly-coupled framework can bring higher accuracy to the system, it also increases the computation and complexity. To save the computation and ensure the real-time performance of the system, the optimization strategy of the sliding window is

adopted in this paper, which limits the keyframes in the window by marginalization. In the case of marginalization, state x_i of IMU and inverse depth λ_m of a feature are selectively marginalized from the window. At the same time, the measurement corresponding to the state of marginalization is transformed into prior information.

At i^{th} moment, the system state to be optimized in the sliding window is defined as follows:

$$\chi = [x_n, x_{n+1}, \dots, x_{n+N}, \lambda_m, \lambda_{m+1}, \dots, \lambda_{m+M}] \quad (11)$$

$$x_i = [p_{wb_i}, q_{wb_i}, v_i^w, b_a^{b_i}, b_g^{b_i}]^T, i \in [n, n+N] \quad (12)$$

x_i includes the position, orientation, and velocity of IMU in the world frame at i^{th} moment, as well as the bias estimation of acceleration and angular velocity in the IMU body frame. In the sliding window, n represents the body state, m represents the starting time of landmarks, N is the number of keyframes, and M is the number of road signs observed by all keyframes.

From the composition of the sliding window, we can see that there are three main cost functions in the tightly-coupled framework, namely, the prior information of marginalization, the measurement residuals of IMU and the visual observation residuals:

$$\min_{\chi} \left\{ \|r_p - J_p \chi\|_{\Sigma_p}^2 + \sum_{i \in B} \|r_b(z_{b_i b_{i+1}}, \chi)\|_{\Sigma_{b_i b_{i+1}}}^2 + \sum_{i \in B} \rho \left(\|r_f(z_{f_j}^{c_i}, \chi)\|_{\Sigma_{f_j}^{c_i}}^2 \right) \right\} \quad (13)$$

Where, r_p is the prior information in the sliding window, J_p is the Jacobian matrix obtained from the Hessian matrix of the last optimization result; $r_b(z_{b_i b_{i+1}}, \chi)$ is the measurement residual of IMU; $r_f(z_{f_j}^{c_i}, \chi)$ is the reprojection error of 3D feature points. ρ is a robust kernel function for eliminating outliers.

2.5 Loop Detection and Global Pose Graph Optimization

Generally, the visual state estimation only considers the relationship between frames of adjacent times, not the relationship between frames from the global perspective. At the same time, due to the limitation of computation by sliding window and marginalization scheme, the odometry will accumulate the error of the front part, resulting in large drift error and unreliable system. To solve this problem, an integrated relocalization module is added. The relocalization process starts from a loop detection and establishes the Feature-level connections between the candidate frame and the current frame. The corresponding relations of these features are closely integrated into the system, so the drift-free state estimation can be obtained at the minimum cost. Among them, loop detection plays a role in recognition, providing more accurate and effective data for pose graph optimization, to get a better global consistency estimation. In this paper, the DBOW2 [2] loop detection method based on the Bag-of-Words (BoW) model

is used to construct and query the keyframe database. After the loop closure candidates are successfully detected, the loop closure candidate with known pose and the matching frame in the sliding window are matched. The loop closure frame is added to the optimization of the sliding window, and then the global pose graph optimization is carried out to modify other poses to ensure global consistency. The specific process is shown in the Fig. 1.

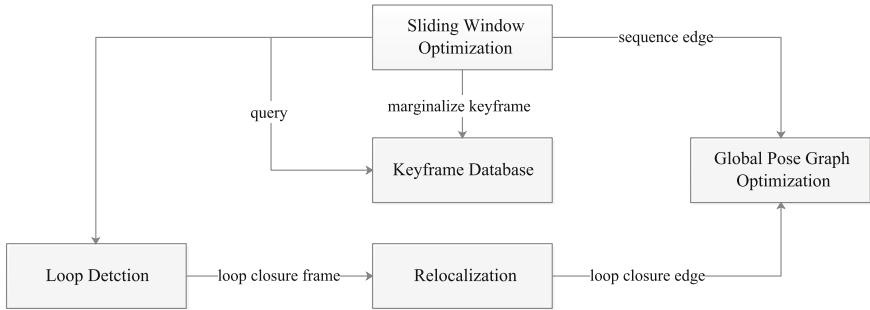


Fig. 1. Flow chart of loop detection and pose graph optimization

3 Real-Time 3D Dense Reconstruction

Real-time dense reconstruction is one of the most important parts of the autonomous navigation system. However, dense reconstruction often leads to huge computational costs. The basic requirement of the mapping system is to balance the accuracy of fusion depth measurement and the cost of accurate representation of the environment. For the quadrotor of autonomous navigation, what is useful is the free space information, not only the obstacles. An ideal planning map must have the ability to efficiently query the idle and occupied states. In this paper, a TSDF [9] (Truncated Signed Distance Field) fusion method is used to build a global dense map incrementally, which can work and update in an indoor and outdoor dynamic scene in real-time.

As shown in Fig. 2, the mapping process is mainly divided into two modules: TSDF construction and voxel map gridding. TSDF construction module is responsible for the 3D reconstruction of the surrounding environment. Firstly, it reads the pose and point cloud data of the current frame from the visual-inertial odometry module, and transfers these data to TSDF integration layer, which merges the point cloud data of the frame with the map of the previous moment to generate a new voxel map. Then the voxel map is transformed into OctoMap [4] (probabilistic 3D framework based on Octrees) for obstacle-avoidance navigation. The voxel map gridding module will transfer the generated voxel map into the grid layer, and voxel is converted to mesh representation through the Marching Cubes algorithm [7] to generate the map for visualization.

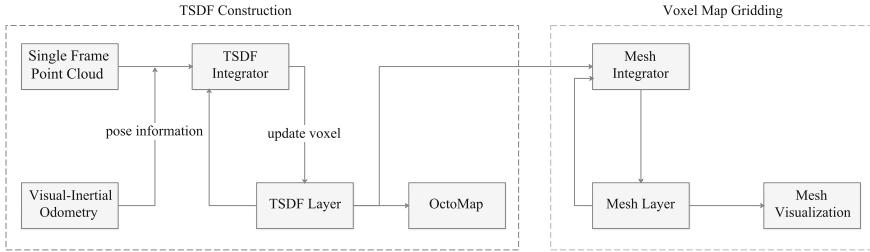


Fig. 2. Flow chart of dense mapping algorithm

4 Trajectory Planning

Based on the research and analysis of different UAV motion planning algorithms, this paper designs and implements a complete online trajectory planning and tracking control algorithm, in which the front-end of trajectory planning algorithm is based on the sampling global planning BIT*(Batch Informed Trees) [3] performs geometric path search, quickly finds an obstacle-free path. At the back-end, we consider UAV dynamics, then construct a trajectory optimization scheme of minimum-snap [8], and ensure the safety of the trajectory by inserting an intermediate point iteratively. Besides, after the planning process, the collision detection module uses real-time environmental information to detect whether the trajectory collides with the obstacle. At last, dynamic replanning is carried out for the trajectory, and the specific steps are shown in Fig. 3.

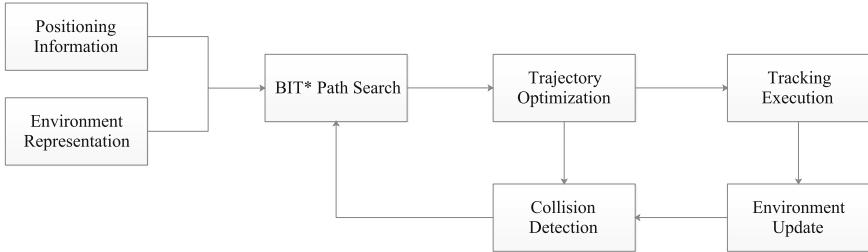


Fig. 3. Flow chart of trajectory planning navigation algorithm

5 Experimental Results

We perform indoor and outdoor navigation experiments to evaluate the reliability of the autonomous navigation system. In terms of hardware, the navigation module of the platform uses the ZED binocular stereo camera of Stereolabs and the NVIDIA Jetson TX2 modular embedded microcomputer; the flight control

module uses the Pixhawk 2 of the open-source Pixhawk series, with built-in three redundancy IMUs, including two groups of MPU9250, and L3GD20 and LS303D as standby. In terms of software, the Ubuntu 16.04 system is used in this paper, and the corresponding robot operating system framework of ROS (Robot Operating System) is established on this system.

5.1 Indoor Experiment

In the indoor navigation experiment, we choose the laboratory environment as the experimental area. Yaw angle control is added in the experiment to ensure the forward direction is always in the camera's angular field of view. The UAV completes the indoor turning navigation mission, and the upper computer of the ground station can observe the operation status of the platform in real-time to complete the navigation function. The real-time navigation flight is shown in Fig. 4. The camera's perspective and navigation display are from the upper computer of the ground station, and the yellow line represents the real-time planning trajectory. The feasibility of indoor autonomous navigation of the UAV platform is verified by the experiment.

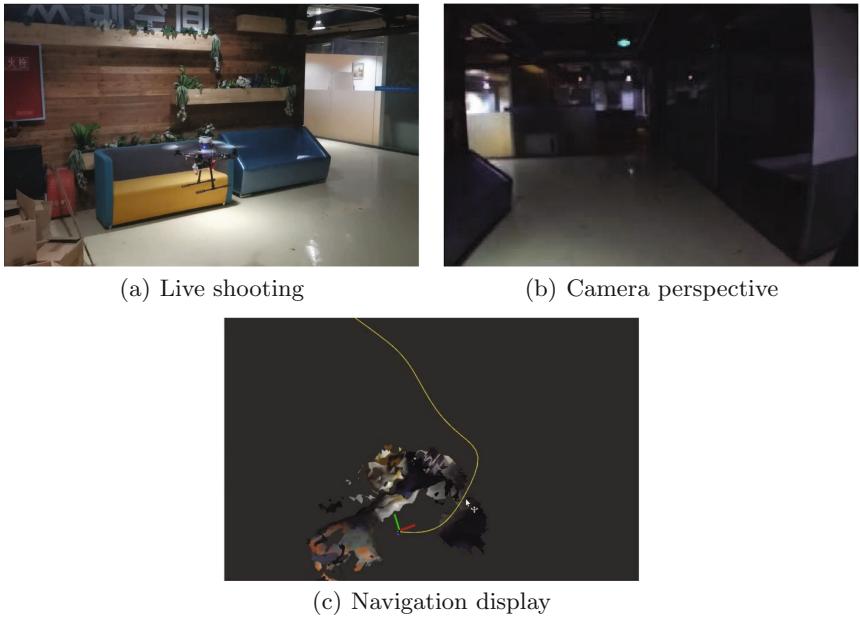


Fig. 4. Indoor navigation experiment

5.2 Outdoor Experiment

In the outdoor navigation experiment, as shown in Fig. 5, we choose a more complex forest environment as the experimental area. The whole navigation flight length is about 300 m, and the average flight speed is about 1 m/s. Due to the limited effective sensing range of the sensor, with the updating of the ground information, we carry out multiple trajectory replanning in the navigation process to avoid obstacles such as roots and branches. The experimental results show that the autonomous navigation system of UAV is feasible and has good real-time performance and robustness.

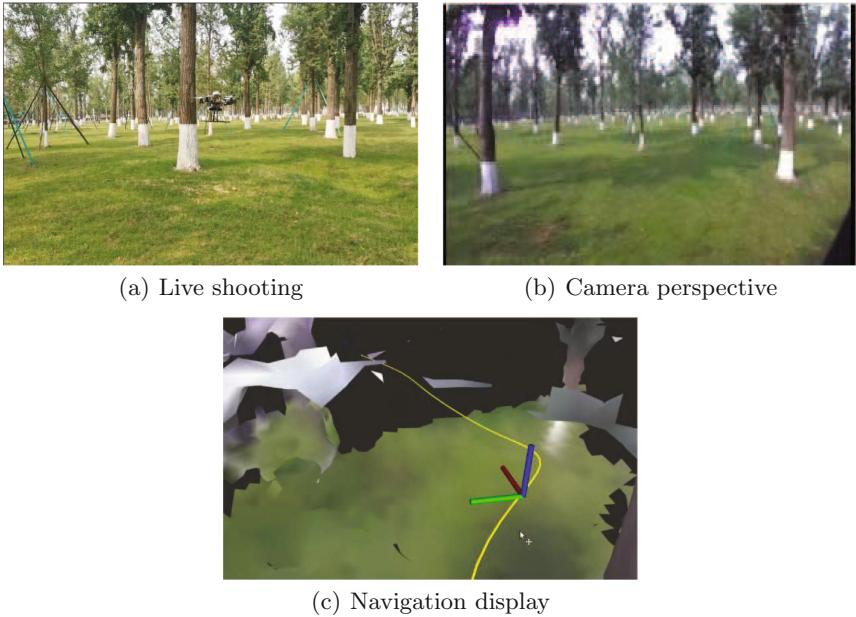


Fig. 5. Outdoor navigation experiment

6 Conclusions

Based on the research background of autonomous navigation of quadrotor visual-inertial navigation, this paper makes a comprehensive study on positioning, mapping, and planning control, and puts forward a complete platform design scheme. First of all, it improves the applicability of the tightly-coupled visual-inertial odometry algorithm on the actual flight platform. Secondly, it solves the two problems of map dynamic update and real-time construction by using TSDF construction and octree structure. Finally, it designs and implements a complete UAV online trajectory planning and tracking control algorithm. The

design is compatible with accuracy, real-time, and robustness, and has many practical scenarios.

This paper realizes the function of autonomous navigation in the unknown indoor and outdoor environment. However, in terms of the overall performance of the system, there is still some room for improvement. In future work, we will consider the fusion of visual-inertial navigation and GPS information, combined with global and local estimation. Besides, deep learning has been studied in visual SLAM, which can be used in feature extraction, loop detection, and other applications to improve the overall performance of the positioning algorithm.

Acknowledgement. This work is supported in part by the National Natural Science Foundation of China under grant (No. 61973055), the Fundamental Research Funds for the Central Universities (No. ZYGX2019J062), and a grant from the applied basic research programs of Sichuan province (No. 2019YJ0206).

References

1. Deng, H., Arif, U., Fu, Q., Xi, Z., Quan, Q., Cai, K.Y.: Visual-inertial estimation of velocity for multicopters based on vision motion constraint. *Robot. Autonom. Syst.* **107**, 262–279 (2018). <https://doi.org/10.1016/j.robot.2018.06.010>
2. Galvez-López, D., Tardos, J.D.: Bags of binary words for fast place recognition in image sequences. *IEEE Trans. Rob.* **28**(5), 1188–1197 (2012)
3. Gammell, J.D., Srinivasa, S.S., Barfoot, T.D.: Batch informed trees (BIT*): sampling-based optimal planning via the heuristically guided search of implicit random geometric graphs. In: Proceedings IEEE International Conference on Robotics & Automation (2014)
4. Hornung, A., Wurm, K.M., Bennewitz, M., Stachniss, C., Burgard, W.: OctoMap: an efficient probabilistic 3D mapping framework based on octrees (2012)
5. Lin, Y., Gao, F., Qin, T., Gao, W., Liu, T., Wu, W., Yang, Z., Shen, S.: Autonomous aerial navigation using monocular visual-inertial fusion. *J. Field Robot.* **35**, 23–51 (2017)
6. López, B.T., How, J.P.: Aggressive 3-D collision avoidance for high-speed navigation. In: 2017 IEEE International Conference on Robotics and Automation (ICRA) (2017)
7. Lorensen, W.E., Cline, H.E.: Marching cubes: a high resolution 3D surface construction algorithm. *Comput. Graph.* **21**(4), 163–169 (1987)
8. Mellinger, D., Kumar, V.: Minimum snap trajectory generation and control for quadrotors. In: 2011 IEEE International Conference on Robotics and Automation (ICRA) (2011)
9. Newcombe, R.A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A.J., Kohi, P., Shotton, J., Hodges, S., Fitzgibbon, A.: KinectFusion: real-time dense surface mapping and tracking. In: 2011 10th IEEE International Symposium on Mixed and Augmented Reality, pp. 127–136 (2011)
10. Oleynikova, H., Burri, M., Taylor, Z., Nieto, J., Galceran, E.: Continuous-time trajectory optimization for online UAV replanning. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2016)
11. Qin, T., Li, P., Shen, S.: Vins-mono: a robust and versatile monocular visual-inertial state estimator. *IEEE Trans. Rob.* **34**(4), 1004–1020 (2018)

12. Shen, S., Michael, N., Kumar, V.: Autonomous multi-floor indoor navigation with a computationally constrained MAV. In: 2011 IEEE International Conference on Robotics and Automation, pp. 20–25 (2011). <https://doi.org/10.1109/ICRA.2011.5980357>
13. Sun, K., Mohta, K., Pfrommer, B., Watterson, M., Liu, S., Mulgaonkar, Y., Taylor, C.J., Kumar, V.: Robust stereo visual inertial odometry for fast autonomous flight. *IEEE Robot. Autom. Lett.* **3**(2), 965–972 (2018)
14. Thrun, S.: Simultaneous Localization and Mapping, pp. 13–41. Springer, Berlin (2008). https://doi.org/10.1007/978-3-540-75388-9_3
15. Tomasi, C., Kanade, T.: Detection and tracking of point features. Technical report, International Journal of Computer Vision (1991)
16. Usenko, V.C., von Stumberg, L., Pangercic, A., Cremers, D.: Real-time trajectory replanning for MAVs using uniform B-splines and 3D circular buffer. CoRR abs/1703.01416 (2017)



Unstructured Road Segmentation Method Based on Super Pixel and Region Growing Algorithm

Tao Liu^{1,2(✉)}, Xi Zhong^{1,2}, and Lixin Zhang¹

¹ University of Science and Technology Beijing, Beijing 100083, China
liutao@ies.ustb.edu.cn

² Shunde Graduate School of University of Science and Technology Beijing,
Beijing 528300, Guangdong, China

Abstract. Given the large randomness and the low color information utilization rate, a region growing algorithm based on super pixel is proposed. It can also eliminate the uncertainty of the initial setting parameters of super pixel segmentation. The method in this paper was proposed by three steps. In the first step, set the most suitable number of super pixel by the lab feature histogram of the image. In the second step, segment the original image by the SLIC super pixel segmentation algorithm. In the third step, merge super pixel blocks by region growing algorithm and then obtain the target unstructured road. This paper uses Jaccard coefficients to evaluate the segmentation accuracy. In a warehouse environment, this method is more accurate than traditional region growing algorithm and normalized segmentation method based on SLIC super pixel.

Keywords: Color feature · Super pixel · Region growing algorithm · Unstructured road

1 Introduction

In the realization of intelligent logistics warehouses, the AGV technology is indispensable. Navigation problem of intelligent guided vehicles is divided into AGV car navigation problems on structured road and unstructured road based on whether there are clear lines on the road. The research on structured road navigation is mature at present, but the research on unstructured road navigation is not enough. Road segmentation is a key technology for intelligent guided vehicles to understand the environment. The quality of road segmentation determine the success of autonomous navigation [1].

Methods based on road region characteristics [2], template [3], and neural network [4] are used to segment unstructured road area. Based on road region characteristics, many scholars propose different methods. Wang, X. and others

use image texture features and color features to extract features of multiple dimensions. It uses PCA algorithm to obtain principal features, trains principal features to get a pixel SVM classifier and then gets the unstructured road area [5]. Li, Y. proposes an unstructured road segmentation method based on RGB entropy. This method selects the threshold value to do the initial segmentation based on RGB entropy histogram, and then use the side-to-side method to segment unstructured road areas accurately [6]. Qi, N. proposes a combination method based on model and feature. The method obtains a saliency map of road and divides the road area accurately by training a road model [7].

2 Region Growing Algorithm

Image segmentation is a technique of dividing an image into several interested region, which provides a basis for the subsequent processing of computer vision [8]. Region growth is a classic algorithm in image segmentation, which was first proposed by Levine et al. The basic idea of regional growth is simple. Firstly, the original image is mapped to a specific color space, which can be any space such as RGB, CMY, HSV, HIS, LAB, etc., or it can be converted to gray space. Then we sprinkle some seed points on the original image. Starting from the seed point, look for the pixel points that are similar to the seed point in the adjacent pixels, and mark it the same as the seed point. The criteria for evaluating similarity are not uniform, and there will be different algorithms according to the needs and different comparison factors, generally described as: $dist(p, q) < th$, $p, q \in I$.

In this formula, p is the seed pixel point, q is the neighboring pixel point which may be in an eight-neighbourhood or a four-neighborhood, I is the entire image pixel, and th is a threshold, $dist(p, q)$ represents the Euclidean distance of pixels p and q . The final result is dividing the pixels on the original image into two sets $\Omega(p_1)$, $\Omega(p_2)$:

$$\Omega(p_1) = \{p_1 \in N; dist(I(p_1) - I(q)) \leq th\} \quad (1)$$

$$\Omega(p_2) = \{p_2 \in N; dist(I(p_2) - I(q)) > th\} \quad (2)$$

Among them, $th > 0$, $dist(I(p) - I(q))$ is the calculated value of pixel similarity between pixels, $I(p)$ and $I(q)$ are the characteristic attributes of seed pixels and domain pixels Value, N is the neighborhood area around the seed point q .

If the gray scale image is (a), the position of p point is expressed by red box, and compare it with pixels in its eight neighborhoods (blue box). If th is 2, the areas where the gray value difference is less than 2 will be fused into the seed point area, and then get the figure (b); next, 7 is used as the seed point, 6 will also be fused, as shown in (c); then the next seed point will be 6, and point 6 and point 5 are fused, as shown in figure (d) (Fig. 1).

The fusion process is continued as this method, and the growth stops until there are no neighboring pixels in the 8 neighborhoods can be fused to a certain seed point. As shown in the above figure to figure (d), there is no pixel point that can continue to grow, then the region growth ends.

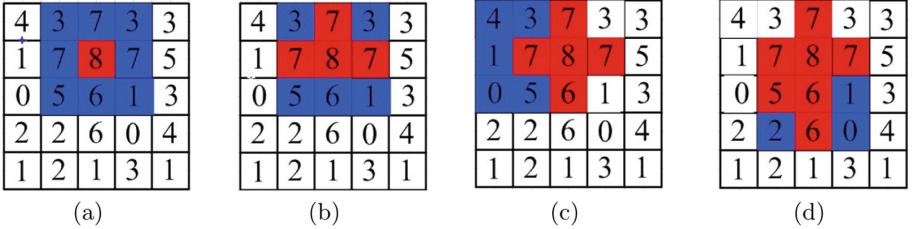


Fig. 1. Schematic diagram of regional growth

3 SLIC Super Pixel Segmentation

The SLIC algorithm was firstly proposed by Achanta et al. [9, 10]. It is a super pixel segmentation algorithm based on k-means clustering. Firstly, select the initial super pixel clustering center, and distribute it on the original image evenly according to the setting number. If the setting number of seed points is K , the number of pixels in the original image is N , and the size of each super pixel block [11] is approximately $s = N/K$. Convert the picture to CIELAB space and form a five-dimensional feature vector $[l_g, a_g, b_g, x_g, y_g]$. $[x_g, y_g]$ is the coordinates of cluster centers [12], $[l_g, a_g, b_g]$ is the luminance component. Initialize the label of each pixel as $l(n) = -1$ and initialize the distance of each pixel as $d(n) = \infty$.

Change the cluster center to the minimum gradient position of an $n \times n$ neighborhood, generally n is 3, and then extend the label in the $2s$ neighborhood from each cluster center as shown in Fig. 2.

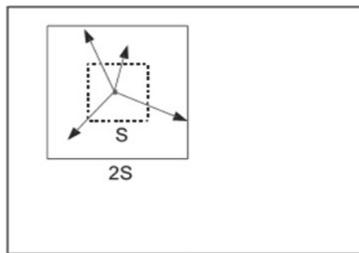


Fig. 2. SLIC search comparison range

The criterion for labeling labels is the distance from pixel in the neighborhood to the cluster center which is denoted as D_n . If $D_n < d(n)$, the pixel is merged into the area where the center is located, until all pixels are calculated and merged. And then recalculate D_n . After repeating it several times, the calculation ends until the coordinates of the clustering center are not changed.

$$D_n = D_{lab} + \frac{m}{s} d_{xy} \quad (3)$$

$$D_{lab} = \sqrt{(l_g - l_n)^2 + (a_g - a_n)^2 + (b_g - b_n)^2} \quad (4)$$

$$d_{xy} = \sqrt{(l_g - l_n)^2 + (a_g - a_n)^2} \quad (5)$$

Among them: m is a compact factor. As m getting larger, the influence of the distance term d_{xy} is getting greater; and as m getting smaller, the influence of the color distance term d_{lab} is getting greater.

4 Research Methods in This Paper

4.1 Initial Setting Value of Super Pixel Block

The initial setting value of the super pixel block is an important parameter of the SLIC algorithm, and the appropriate number of super pixel blocks will improve the accuracy of the subsequent segmentation. The three-dimensional component of lab space represents the brightness of the pixel, the component from green to red, and the component from blue to yellow. If the color is the same on the same area and the brightness is different, it will be visually regarded as a different area. In order to separate such subtle changes, the l component is studied to find out the appropriate initial number of the super pixel blocks.

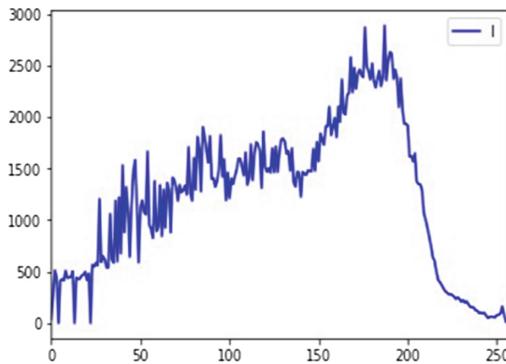


Fig. 3. Histogram of images L component

Turn the image into an image in lab space and calculate the histogram of the L component, as shown in Fig. 3. There will be some “small peaks” on the histogram of the L component. The peaks indicate that the brightness of the color area is clustered, that is, there is a region with uniform brightness values, so we should find the smallest peak. Record the step size as σ , and the difference

between the high and low values as θ , meeting these conditions is considered a peak.

When searching for the peak, it must meet condition (6) to (9):

$$l_hist[i] - l_hist[i - 1] > 0 \quad (6)$$

$$l_hist[i] - l_hist[i + 1] > 0 \quad (7)$$

$$l_hist[i] - l_hist[i - \sigma] > 0 \quad (8)$$

$$l_hist[i] - l_hist[i + \sigma] > 0 \quad (9)$$

$l_hist[i]$ represents the number of super pixels when l is i . Note the obtained minimum peak value is hpv , and the size of the entire image is $N = h \times w$, then the appropriate initial setting value of the super pixel [13] block should be:

$$n = \text{int}\left(\frac{N}{hpv}\right) \quad (10)$$

4.2 Super Pixel Segmentation and Merging

According to the most appropriate super-pixels setting value, select seed points, optimize seed points, and then fuse pixels in the S area to several pixel blocks, then a marked sheet is obtained as shown in Fig. 4. Count the pixel position of each marked area, find the centroid position of the pixel block, and calculate the average l , a , b value of a single super-pixel block as the l_i , a_i , b_i , so that the entire pixel block is represented by a pixel, expressed as $[x_i, y_i, l_i, a_i, b_i]$:

$$\bar{x}_i = \frac{1}{|c_i|} \sum_{n_j \in c_i} x_j \quad (11)$$

$$\bar{y}_i = \frac{1}{|c_i|} \sum_{n_j \in c_i} y_j \quad (12)$$

$$\bar{l}_i = \frac{1}{|c_i|} \sum_{n_j \in c_i} l_j \quad (13)$$

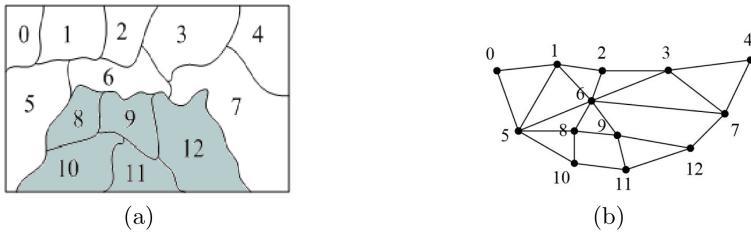
$$\bar{a}_i = \frac{1}{|c_i|} \sum_{n_j \in c_i} a_j \quad (14)$$

$$\bar{b}_i = \frac{1}{|c_i|} \sum_{n_j \in c_i} b_j \quad (15)$$

In these formulas, c_i represents the set of pixels in super pixel block number i , $|c_i|$ is the number of pixels in the super pixel block, and n_j represents single pixel in c_i .

The graph is a collection of fixed points and edges. The graph can be divided into an undirected graph and a directed graph. The super pixel block can be understood as an undirected graph after being pixelated, as shown in Fig. 5. In the undirected graph G , each vertex represents a representative pixel of a super pixel block, and each edge represents a connection between adjacent regions. Vertices and edges form a Region Adjacency Graphs (RAG) of super pixel segmentation.

0	0	0	1	2	2	2	2	3	3
0	0	1	1	1	2	2	2	3	3
4	0	1	1	2	2	2	3	7	3
4	4	4	1	2	6	2	6	7	7
4	4	5	6	6	6	6	6	7	7
4	4	5	6	6	6	6	6	8	7
4	5	5	5	6	6	6	11	8	8
9	9	5	5	10	10	10	11	8	12
9	5	5	9	10	10	10	11	11	11
9	9	9	9	10	10	11	11	11	12
9	9	10	10	10	10	10	11	11	12

Fig. 4. Marking table for superpixel segmentation**Fig. 5.** Superpixel segmentation map and RAG

Borrowing the idea of region growth, however, for the super pixel block, the position of each super pixel block is not fixed, so here we use the BFS algorithm to search for adjacent super pixel blocks. If the seed point is recorded as number 11 in the super pixel segmentation, then search the adjacent super pixel blocks and press them into a queue. And then find 10, 8, 9, 12. Press these four pixel blocks into the queue, search from block 10, block 5, 8, 9 are determined whether they are already in the queue. If they are in the queue, they will no longer be pushed into the queue. It can reduce a certain amount of calculation. Compare the two super-pixel blocks in turn. Similarity, if the two super pixel blocks are sufficiently similar, then merge the two super pixel blocks. Since we are only looking for the area of unstructured road, we only need to mark and merge super pixel blocks similar to the seed points.

4.3 Algorithm

Step 1: According to the histogram of the lab space of the image, set the initial number of super pixels of SLIC.

Step 2: Segment the image to get a label matrix using super pixel segmentation algorithm.

Step 3: Take out the coordinate values of each super pixel block of the SLIC segmentation result and store them in the *x_array* and *y_array* lists. And then

the centroid points of each super pixel block are obtained by using Eqs. (8) (9) and stored in value *centroi*.

Step 4: According to the location of the *centroi*, judge the adjacent relationship of the super pixel blocks, and obtain the RAG.

Step 5: According to the characteristics of the image, select a point below the middle of the image, and record its label as ‘*qishi*’.

Step 6: Initialize a queue. Press ‘*qishi*’ into the queue, pop the first label from the queue and record it as ‘seed’. Query the label of the adjacent super pixel block of the seed, and compare the average lab value of the adjacent super pixel block. T is the threshold of similarity, and the label is marked as 1. If the super pixel block has been marked as 1, the label will not be pushed into the queue array, otherwise it will be pushed into the queue.

Step 7: Pop up *array[0]* in turn, and repeat the Step 6 until all the elements in the array have been popped up.

5 Experimental Results

First of all, from the selection of the number of super pixels, the number of super pixels generated by the color information in the lab space is used to set the parameters of SLIC. If you manually set the number of super pixel divisions randomly and use the same number of super pixels for each image, the boundary fit degree of different images is different. If you want to increase the degree of boundary fitting, and increase the number of super pixels, it will increase the calculation amount. And it will also lead to the greater probability of misjudgment in the steps of merging super pixels.

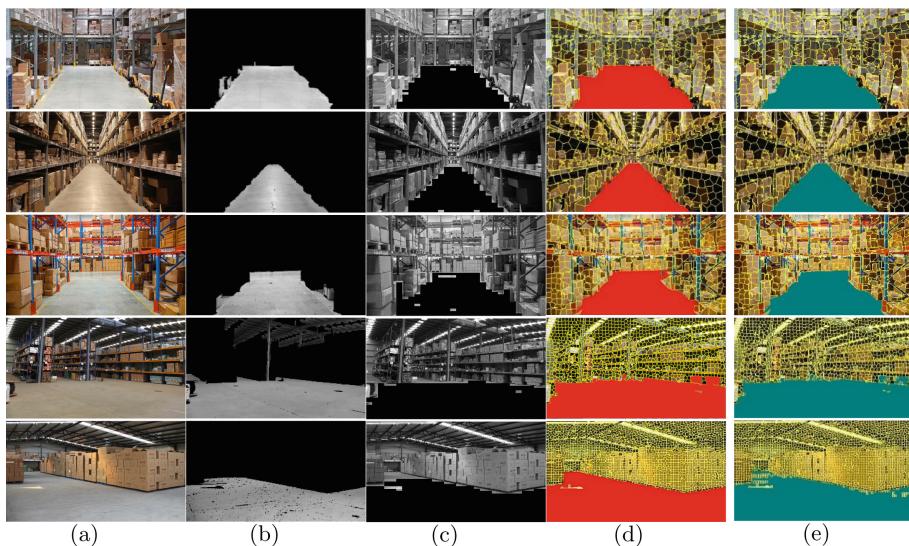
After experimental verification, it can indeed produce a super pixel segmentation map that is closer to the edge of the unstructured road. Comparing the super pixel blocks number of this method with a serials of other different number, the experimental results of the three images are as the Table 1:

Table 1. Results of different numbers of superpixel blocks

Pic	n = 100	n = 200	n = 300	n = 400	n = 500	n = 600	n = 700	n = 800	n = 900	n = 1000	This algorithm
cangku1	0.9523	0.9497	0.9477	0.9255	0.9377	0.9505	0.9232	0.9174	0.9509	0.9365	0.956
cangku2	0.98	0.9621	0.9832	0.9776	0.9769	0.9723	0.9617	0.9792	0.98	0.9803	0.971
cangku3	0.7157	0.7437	0.7386	0.6886	0.6515	0.7398	0.7187	0.6945	0.7129	0.7398	0.7397

A method for adaptively generating the number of super pixel segmentation here can ensure that the segmentation accuracy remains relatively high. Although it cannot be said to be the best, this method can avoid the randomness from manually setting.

After selecting the most suitable number of super pixels, the image is segmented by SLIC super pixel segmentation and merged by the RAG. Compared with the traditional area growth method, the algorithms in [13] and [14], the results are in Fig. 6. The results of Jaccard coefficient evaluation are shown in Table 2.

**Fig. 6.** Schematic diagram of regional growth**Table 2.** The results of Jaccard coefficient evaluation

pic	Area growth method	Literature [1]	Literature [2]	This algorithm
cangku1	0.9405383	0.93900143	0.67583152	0.9560632
cangku2	0.97579576	0.90922661	0.97780985	0.97780985
cangku3	0.68381972	0.67023201	0.724688249	0.731311523
cangku4	0.71571605999	0.9286867576	0.93571694	0.9418659205
cangku5	0.9175924224	0.92538820329	0.97788052986	0.93592008131

References

- Wang, J., Luo, Y., Qian, W.: Road segmentation recognition of color image region growing algorithm. *Autom. Instrum.* **166**(5), 158–159 (2014). <https://doi.org/10.3969/j.issn.1001-9227.2014.05.052>
- Hu, X., Sun, M., Su, X.: Detection method of completely unstructured road in pseudo color space. *J. Image Graph.* **17**(2), 203–208 (2012)
- Chen, Q., Jing, Y., Chen, J.: Unstructured road detection based on adaptive template. *J. Southeast Univ. (Nat. Sci. Edn.)* **37**(6), 1102–1107 (2007)
- Wang, H., Cai, Y., Jia, Y.: Scene adaptive road segmentation algorithm based on deep convolutional neural network. *J. Electron. Inf. Technol.* **39**(2), 263–269 (2017)
- Wang, X., Meng, F., Lv, G.: Unstructured road recognition based on PCA-SVM criterion to improve regional growth. *Comput. Appl.* **37**(6), 1782–1786 (2017). <https://doi.org/10.11772/j.issn.1001-9081.2017.06.1782>
- Li, Y., Fu, X., Xue, Q.: Unstructured road segmentation method based on RGB entropy. *Comput. Eng. Des.* **38**(06), 1570–1574 (2017). <https://doi.org/10.16208/j.issn1000-7024.2017.06.031>

7. Qi, N., Yang, X., Li, C.: Unstructured road detection via combining the model-based and feature-based methods. *IET Intell. Transp. Syst.* **13**(10), 1533–1544 (2019). <https://doi.org/10.1049/iet-its.2018.5576>
8. Khan, J., Adham, I.R., Bhu Iyan, S.: A customized gabor filter for unsupervised color image segmentation. *Image Vis. Comput.* **27**(4), 489–501 (2009). <https://doi.org/10.1016/j.imavis.2008.07.001>
9. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Ssstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2274–2282 (2012). <https://doi.org/10.1109/TPAMI.2012.120>
10. Achanta, R., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: Technical Report, EPFL, Technical report 149300: SLIC Superpixels (2010)
11. Yu, W., Wang, Y., Liu, H., He, J.: Superpixel-based CFAR target detection for high-resolution SAR images. *IEEE Geosci. Remote Sens. Lett.* **13**(5), 730–734 (2016). <https://doi.org/10.1109/LGRS.2016.2540809>
12. Liao, M., Li, Y., Zhao, Y.: A new image superpixel segmentation method. *J. Electron. Inf. Technol.* **42**(2), 364–370 (2020). <https://doi.org/10.11999/JEIT190111>
13. Gong, B., Lu, L., Cao, X.: Unstructured road segmentation based on gray features. *Comput. Knowl. Technol.* **11**(26), 147–149 (2015). <https://doi.org/10.14004/j.cnki.ckt.2015.2986>
14. Zhang, Y.: Research on normalized segmentation method based on SLIC Superpixels (2015)



Acoustical Field Modeling for Communication Through Steel Based on FDTD

Tao Liu^(✉), Jiajia Liu, Zongmei Bai, and Ouming Liu

Key Laboratory of Knowledge Automation for Industrial Processes
of Ministry of Education, School of Automation and Electrical Engineering,
University of Science and Technology Beijing, Beijing 100083, China
liutao@ies.ustb.edu.cn

Abstract. The acoustical field characteristics of the steel-through communication system are studied and the Finite-Difference Time-Domain is used to establish the acoustical field model of the ultrasonic propagated in the system using Yee-type mesh division. Further, the accuracy of acoustical field model is improved by the high-order difference method applied in the algorithm which the effectiveness is demonstrated by the dynamic simulation result of the acoustical field, and the determination coefficient for data verification is increased by about 2%.

Keywords: Steel-through system · Acoustical field modeling · Finite-difference time-domain · High-order difference · Numerical simulation

1 Introduction

The Radio Frequency (RF) signal is inapplicable in the communication of airtight metal structure for the Faraday shield, such as the airtight containers, pressure pipes, and the like. Ultrasonic wave has good directivity and strong penetrability as propagating in steel. It also guarantees safety and integrity of airtight construction so it becomes the best choice for communication in this field [1].

The steel-through communication system consists of transmitter, receiver and the communication channel include steel and piezoelectric transducers. A pair of sensors are symmetrically distributed. They can realize conversion between the ultrasonic signal and electrical signal. Steel is a special communication medium, so studying propagation characteristics of ultrasonic wave in steel is necessary and beneficial to optimize system performance.

Numerical simulation methods are used in ultrasonic field modeling, such as Finite Element Method (FEM) [2, 3], Finite Volume Method (FVM) [2], Boundary Element Method (BEM) [4] and Finite-Difference Time-Domain method (FDTD) [5]. The FEM is flexible to divide grids, but needs massive calculation for complex structures; The FVM can process complex networks, but improving

the accuracy is complicated; The BEM is limited to simple geometries with small scales; The FDTD can discretize differential equations directly and compute in parallel, but needs strict calculation conditions.

The FDTD is originally applied to the electromagnetic field for solving the Maxwell equation. Then the application has been extended to the simulation of elastic wave field and nondestructive testing field. Numerical analysis by FDTD can gain information about interior part of research materials [6, 7]. FDTD (2, 2) and FDTD (2, 4) are introduced for modeling scalar wave propagation in [8]. On this basis, acoustical field modeling for steel-through system is studied.

This paper applies FDTD to the steel-through communication system for the acoustical field modeling. The improvement is that it uses high-order difference method and transparent source technique, which improves accuracy and saves compute time, also it uses pseudo-color image to realize the visualization of acoustical field.

2 Acoustical Field Modeling Based on FDTD

The stress and velocity are researched in the elastic wave field, correspondingly, the intensity of electric field and magnetic field are studied in the electromagnetic field [9]. According to the structure of steel-through system, the acoustical field modeling is divided into two parts: the piezoelectric transducer and the steel.

2.1 Transducer Modeling

The FDTD algorithm is utilized to numerically calculate the thickness extension vibration mode of the piezoelectric transducer. The performance of piezoelectric material is expressed by acoustic and piezoelectricity equations, which construct to the discrete system of equations using FDTD method.

The thickness cutting wafer of the piezoelectric transducer is shown in Fig. 1 [10], where d represents the thickness, A represents the area of electrode surface, and the piezoelectric equations are shown in the formulas (1) and (2).

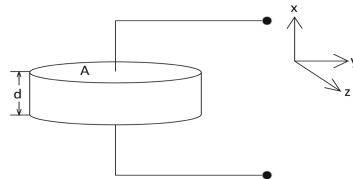


Fig. 1. Thickness cutting wafer of transducer

$$D = \varepsilon_{xx}^S E + e_{xx} S \quad (1)$$

$$T = -e_{xx} E + c_{xx}^E S \quad (2)$$

where E represents the electric field intensity; D represents the electrical displacement; S represents the strain; T represents the stress; c_{xx}^E represents the

elastic coefficient, which means that the increase of T caused by each unit of S added under constant E in the x direction; e_{xx} represents the piezoelectric stress constant, which means that the increase of D caused by each unit of S added; ε_{xx}^S represents the dielectric constant, which means that the increase of D caused by each unit of E added under constant S in the x direction.

The transducer vibrates in thickness mode and there is no free charge in the crystal, therefore $\partial D/\partial x = 0$. According to formulas (3) and (4), formulas (1) and (2) are further deduced to obtain the first-order partial differential equations of electric field governing equation and the wave propagation equation in the piezoelectric material shown in the formulas (5) and (6), respectively.

$$\partial v/\partial x = \partial S/\partial t \quad (3)$$

$$\partial v/\partial t = (1/\rho) \cdot \partial T/\partial x \quad (4)$$

$$\partial T/\partial t = (e_{xx}^2/\varepsilon_{xx}^S + c_{xx}^E) \partial v/\partial x - (e_{xx}/\varepsilon_{xx}^S) \partial D/\partial t \quad (5)$$

$$\partial E/\partial t = - (e_{xx}/\varepsilon_{xx}^S) \partial v/\partial x + (1/\varepsilon_{xx}^S) \partial D/\partial t \quad (6)$$

where v represents the velocity of ultrasonic in the transducer, and ρ represents the density of piezoelectric material.

Adopted one-dimensional discrete Yee-cell model where the time step Δt is assumed to satisfy the Courant stability condition, formulas (3)–(6) are discretized into FDTD equations as (7)–(9).

$$v_{k+1/2}^{n+1} = v_{k+1/2}^n + (\Delta t/\rho\Delta x) \left(T_{k+1}^{n+1/2} - T_k^{n+1/2} \right) \quad (7)$$

$$T_k^{n+1/2} = T_k^{n-1/2} + \left(\frac{e_{xx}^2}{\varepsilon_{xx}^S} + c_{xx}^E \right) \frac{\Delta t}{\Delta x} (v_{k+1/2}^n - v_{k-1/2}^n) - \frac{e_{xx}}{\varepsilon_{xx}^S} (D^{n+1/2} - D^{n-1/2}) \quad (8)$$

$$E_k^{n+1/2} = E_k^{n-1/2} - \frac{e_{xx}\Delta t}{\varepsilon_{xx}^S\Delta x} (v_{k+1/2}^n - v_{k-1/2}^n) + \frac{1}{\varepsilon_{xx}^S} (D^{n+1/2} - D^{n-1/2}) \quad (9)$$

where Δx is the spatial step sizes in the x direction; N represents the total time; the meaning of $v_l^t = v_{k+1/2}^{n+1}$, $l \in [0, 0.5, 1, 1.5, \dots, d]$, $t \in [0, 0.5, 1, 1.5, \dots, N]$ is the velocity value of the ultrasonic at the l^{th} position in the transducer at time t , and the rest of the parameters have similar meanings to $v_{k+1/2}^{n+1}$.

2.2 Steel Modeling

The variables involved in the acoustical field of the steel are velocity $v = (v_x, v_y)$ and stress $\sigma = (\sigma_{xx}, \sigma_{yy}, \sigma_{xy})$, in which v_x represents horizontal component of velocity, v_y represents vertical component of velocity, σ_{xx} and σ_{yy} represent normal stresses in x and y direction, respectively, σ_{xy} represents shear stress. In view of isotropic properties of steel in [11], the first-order differential equations of the ultrasonic waves in the steel are composed of the formulas (10)–(14).

$$\rho \cdot (\partial v_x/\partial t) = (\partial \sigma_{xx}/\partial x) + (\partial \sigma_{xy}/\partial y) \quad (10)$$

$$\rho \cdot (\partial v_y / \partial t) = (\partial \sigma_{xy} / \partial x) + (\partial \sigma_{yy} / \partial y) \quad (11)$$

$$\partial \sigma_{xx} / \partial t = (\lambda + 2\mu) \cdot (\partial v_x / \partial x) + \lambda \cdot (\partial v_y / \partial y) \quad (12)$$

$$\partial \sigma_{yy} / \partial t = (\lambda + 2\mu) \cdot (\partial v_y / \partial y) + \lambda \cdot (\partial v_x / \partial x) \quad (13)$$

$$\partial \sigma_{xy} / \partial t = \mu ((\partial v_y / \partial x) + (\partial v_x / \partial y)) \quad (14)$$

where λ and μ are lame elastic coefficient. Compared with the transducer piezoelectric wafer, the area and thickness of the steel are much larger. Therefore, the two-dimensional Yee-cell mesh is applied when the formulas (10)–(14) are discrete by the FDTD method.

The formulas (10)–(14) are discretized to obtain discrete difference equations (15)–(19) by the FDTD method. The discrete points of the variables v and σ are cross-configured in two-dimensional sound fields.

$$v_x|_{i,j}^{n+1/2} = v_x|_{i,j}^{n-1/2} + \frac{\Delta t}{\rho} \left(\frac{\sigma_{xx}|_{i+1/2,j}^n - \sigma_{xx}|_{i-1/2,j}^n}{\Delta x} + \frac{\sigma_{xy}|_{i,j+1/2}^n - \sigma_{xy}|_{i,j-1/2}^n}{\Delta y} \right) \quad (15)$$

$$v_y|_{i+1/2,j+1/2}^{n+1/2} = v_y|_{i+1/2,j+1/2}^{n-1/2} + \frac{\Delta t}{\rho} \left\{ \frac{\sigma_{xy}|_{i+1,j+1/2}^n - \sigma_{xy}|_{i,j+1/2}^n}{\Delta x} + \frac{\sigma_{yy}|_{i+1/2,j+1}^n - \sigma_{yy}|_{i+1/2,j}^n}{\Delta y} \right\} \quad (16)$$

$$\sigma_{xx}|_{i+1/2,j}^{n+1} = \sigma_{xx}|_{i+1/2,j}^n + \Delta t \left(\frac{(\lambda + 2\mu)(v_x|_{i+1,j}^{n+1/2} - v_x|_{i,j}^{n+1/2})}{\Delta x} + \frac{\lambda(v_y|_{i+1/2,j+1/2}^{n+1/2} - v_y|_{i+1/2,j-1/2}^{n+1/2})}{\Delta y} \right) \quad (17)$$

$$\sigma_{yy}|_{i+1/2,j}^{n+1} = \sigma_{yy}|_{i+1/2,j}^n + \Delta t \left(\frac{\lambda(v_x|_{i+1,j}^{n+1/2} - v_x|_{i,j}^{n+1/2})}{\Delta x} + \frac{(\lambda + 2\mu)(v_y|_{i+1/2,j+1/2}^{n+1/2} - v_y|_{i+1/2,j-1/2}^{n+1/2})}{\Delta y} \right) \quad (18)$$

$$\sigma_{xy}|_{i,j+1/2}^{n+1} = \sigma_{xy}|_{i,j+1/2}^n + \Delta t \left(\frac{\mu(v_x|_{i,j+1}^{n+1/2} - v_x|_{i,j}^{n+1/2})}{\Delta y} + \frac{\mu(v_y|_{i+1/2,j+1/2}^{n+1/2} - v_y|_{i-1/2,j+1/2}^{n+1/2})}{\Delta x} \right) \quad (19)$$

In which $v_x|_{i,j}^{n+1/2}$ represents the v_x value at which the space position is (i, j) and the time is $n + 1/2$. The representation of other variables is the same.

2.3 Boundary Conditions

Since the ultrasonic wave is reflected at the boundary of the steel plate, the Perfectly Matched Layer(PML) boundary is set at the artificial cutoffs on both sides of the steel plate without the transducer. The distribution of the PML layer is shown in Fig. 2.

The PML layer includes two regions of edges and corners. The media parameters of the PML layers in the four top corner regions are (d_x, d_y) , the media parameters perpendicular to the x-axis are $(d_x, 0)$, and the media parameters perpendicular to the y-axis are $(d_y, 0)$. In this way, the acoustic wave reduces its reflection phenomenon as it passes through the boundary of the medium.

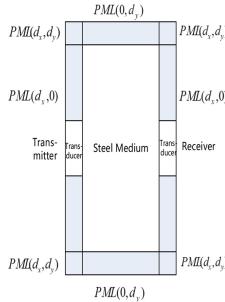


Fig. 2. Distribution of PML layer settings

In the PML layer, the propagation of ultrasonic waves is different and needs to be processed. Suppose that each component of v and σ is split into two sub-components. Combined with formulas (10)–(14), the equation of the PML layer is obtained, and then the FDTD method is also used for the difference.

3 Acoustical Field Modeling Based on High-Order FDTD

The FDTD algorithm does not fully utilize the value of the Yee grid point, and its calculation accuracy can be further improved to meet the requirements of industrial precision. For FDTD (2T, 2M) in [12], 2T and 2M represent orders on time and space, the order improvement can effectively improve the calculation accuracy. Specifically, each increment of T increases the variable on the entire grid at the corresponding time; each increment of M only increases the variable of the adjacent space. Considering the two factors of calculation accuracy and calculation speed, FDTD (2, 4) algorithm is applied to model the acoustical field of steel. And the high-order acoustical field model includes the transducer model and the high-order steel model mentioned in Sect. 2 in this section. The PML setting remains unchanged.

Based on the differential derivation formula in [8], the FDTD (2, 4) equations with the second-order in time and the fourth-order in space are shown as formulas (20)–(24).

$$\begin{aligned} v_x|_{i,j}^{n+1/2} &= \frac{\Delta t}{\rho} \left(\frac{9(\sigma_{xx}|_{i+1/2,j}^n - \sigma_{xx}|_{i-1/2,j}^n)}{8\Delta x} - \frac{\sigma_{xx}|_{i+3/2,j}^n - \sigma_{xx}|_{i-3/2,j}^n}{24\Delta x} \right) \\ &+ \frac{\Delta t}{\rho} \left(\frac{9(\sigma_{xy}|_{i,j+1/2}^n - \sigma_{xy}|_{i,j-1/2}^n)}{8\Delta y} - \frac{\sigma_{xy}|_{i,j+3/2}^n - \sigma_{xy}|_{i,j-3/2}^n}{24\Delta y} \right) + v_x|_{i,j}^{n-1/2} \end{aligned} \quad (20)$$

$$v_y|_{i+1/2,j+1/2}^{n+1/2} = \frac{\Delta t}{\rho} \left(\frac{9(\sigma_{xy}|_{i+1,j+1/2}^{n+1/2} - \sigma_{xy}|_{i,j+1/2}^n)}{8\Delta x} - \frac{\sigma_{xy}|_{i+2,j+1/2}^{n+1/2} - \sigma_{xy}|_{i-1,j+1/2}^n}{24\Delta x} \right) \\ + \frac{\Delta t}{\rho} \left(\frac{9(\sigma_{yy}|_{i+1/2,j+1}^n - \sigma_{yy}|_{i+1/2,j}^n)}{8\Delta y} - \frac{\sigma_{yy}|_{i+1/2,j+2}^n - \sigma_{yy}|_{i+1/2,j-1}^n}{24\Delta y} \right) + v_y|_{i+1/2,j+1/2}^{n-1/2} \quad (21)$$

$$\sigma_{xx}|_{i+1/2,j}^{n+1} = \Delta t (\lambda + 2\mu) \left(\frac{9(v_x|_{i+1,j}^{n+1/2} - v_x|_{i,j}^{n+1/2})}{8\Delta x} - \frac{v_x|_{i+2,j}^{n+1/2} - v_x|_{i-1,j}^{n+1/2}}{24\Delta x} \right) \\ + \Delta t \lambda \left(\frac{9(v_y|_{i+1/2,j+1/2}^{n+1/2} - v_y|_{i+1/2,j-1/2}^{n+1/2})}{8\Delta y} - \frac{v_y|_{i+1/2,j+3/2}^{n+1/2} - v_y|_{i+1/2,j-3/2}^{n+1/2}}{24\Delta y} \right) + \sigma_{xx}|_{i+1/2,j}^n \quad (22)$$

$$\sigma_{yy}|_{i+1/2,j}^{n+1} = \sigma_{yy}|_{i+1/2,j}^n + \Delta t \lambda \left(\frac{9(v_x|_{i+1,j}^{n+1/2} - v_x|_{i,j}^{n+1/2})}{8\Delta x} - \frac{v_x|_{i+2,j}^{n+1/2} - v_x|_{i-1,j}^{n+1/2}}{24\Delta x} \right) \\ + \Delta t (\lambda + 2\mu) \left(\frac{9(v_y|_{i+1/2,j+1/2}^{n+1/2} - v_y|_{i+1/2,j-1/2}^{n+1/2})}{8\Delta y} - \frac{v_y|_{i+1/2,j+3/2}^{n+1/2} - v_y|_{i+1/2,j-3/2}^{n+1/2}}{24\Delta y} \right) \quad (23)$$

$$\sigma_{xy}|_{i+1/2,j}^{n+1} = +\sigma_{xy}|_{i,j+1/2}^n + \Delta t \mu \left(\frac{9(v_x|_{i,j+1}^{n+1/2} - v_x|_{i,j}^{n+1/2})}{8\Delta y} - \frac{v_x|_{i,j+2}^{n+1/2} - v_x|_{i,j-1}^{n+1/2}}{24\Delta y} \right) \\ + \Delta t \mu \left(\frac{9(v_y|_{i+1/2,j+1/2}^{n+1/2} - v_y|_{i-1/2,j+1/2}^{n+1/2})}{8\Delta x} - \frac{v_y|_{i+3/2,j+1/2}^{n+1/2} - v_y|_{i-3/2,j+1/2}^{n+1/2}}{24\Delta x} \right) \quad (24)$$

Compare FDTD (2, 4) and FDTD (2, 2), the number of grid, the distribution of spatial nodes, the interval of time and space are the same, and the difference is that FDTD (2, 4) adds one storage unit per dimension to each grid which ensures the correctness of the algorithm and reduces the numerical dispersion error. In addition, the stability condition of FDTD (2, 2) and FDTD (2, 4) is $\lambda\Delta t \leq h/\sqrt{2}$, $\lambda\Delta t \leq 6h/7\sqrt{2}$ to satisfy the Courant stability condition and ensure the stability of the model [13]. The amount of calculation increase only in space, it is found that the increase of the calculation amount is meaningful with respect to the accuracy of the model.

4 Simulation

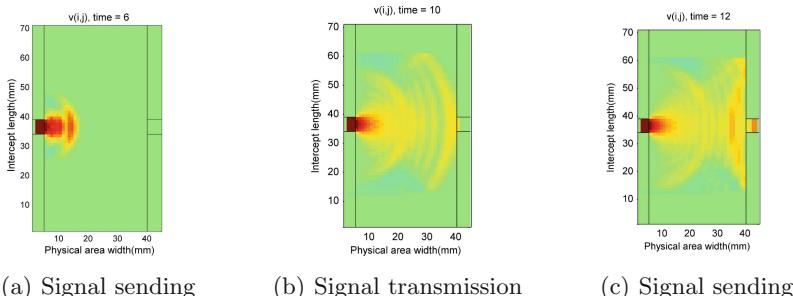
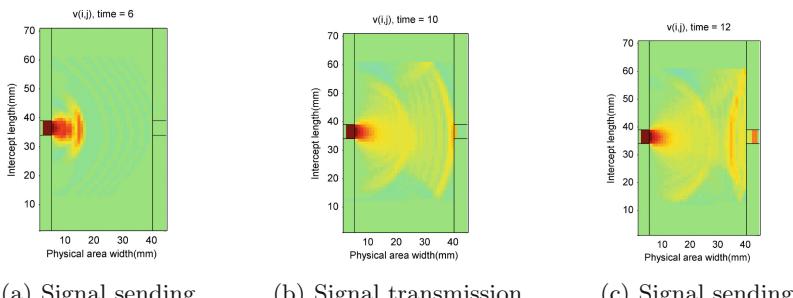
The PML boundary in [14] is applied as the absorption boundary model in the simulation of steel-through system. To set reasonable parameters of the PML layer can simulate the echo effect of ultrasonic waves at the steel boundary.

The thickness of the steel plate is 28 mm. The transducer is selected with a thickness 2.25 mm and a radius of 30 mm, generating resonance per 0.104 MHz. To compare the FDTD (2, 2) with FDTD (2, 4) algorithms, the parameter settings of the physical properties of the materials and the ultrasonic performance parameters are the same as in Table 1.

Both v and σ all describe the propagation characteristics of ultrasonic waves in the steel. Only one of them can be selected during the simulation, and v is sim-

Table 1. Parameter settings

Parameter	Symbol	Value/Unit
Center frequency of transducers	f_0	1 MHZ
Lame constant1	λ	5.7×10^{10}
Lame constant2	μ	1.05×10^{11}
The density of the steel	ρ	7800 kg/m^3
The shear wave velocity	v_s	3230 m/s
The longitudinal wave velocity	v_p	5850 m/s
The size of Yee-type grid	—	70×45
The size of each grid	$\Delta x = \Delta y = h$	$600 \mu\text{m}$

**Fig. 3.** The simulation results of FDTD(2, 2)**Fig. 4.** The simulation results of FDTD(2, 4)

ulated in this paper. The results of simulation are shown in Fig. 3–4, where a and b represent the FDTD (2, 2) and FDTD (2, 4) simulation results, respectively.

Figure 3–4 are the acoustical field distribution of the ultrasonic waves propagating at different times in the steel, Fig. 3–Fig. 4 (a) is the scene that transmitting transducer emits a pulse and ultrasonic waves just enter the steel, Fig. 3–Fig. 4 (b) is the ultrasonic waves propagating in the steel. At the artificial truncation boundary without transducers on both sides of steel, the ultrasonic waves

have weak reflections, but most are absorbed. Figure 3–Fig. 4 (c) is receiving transducer receiving ultrasonic signal. As passing through the steel boundary, the ultrasonic waves can partially enter the PML absorption layer and partly continue to propagate in the steel plate in the form of echo.

Thus, the acoustical field in steel-through system is clearly understood and the feasibility of simulating the ultrasonic acoustical field by the FDTD method is demonstrated. For the effectiveness of the FDTD algorithm, the stability condition of the FDTD (2, 4) have a coefficient of, that is, the maximum time interval of the FDTD (2, 4) is smaller than that of the FDTD (2, 2), so the stability of the FDTD (2, 4) acoustical field model is stricter. Figure 4 is relatively stable, the clutter is less, and the beam is more concentrated to a, which proves that high-order FDTD algorithm can reduce the numerical dispersion error.

The actual output data of the receiving transducer of the oscilloscope is derived and the main loop data is extracted as shown by the “-” in Fig. 5. In addition, there is the data waveform for the FDTD (2, 2), FDTD (2, 4) model in Fig. 5. The output data of the models established by both the FDTD (2, 2) and FDTD (2, 4) method is consistent with the actual waveform. Further, the expression of determination coefficient R^2 that can measure the validity of the model is determined as formula (25) [15].

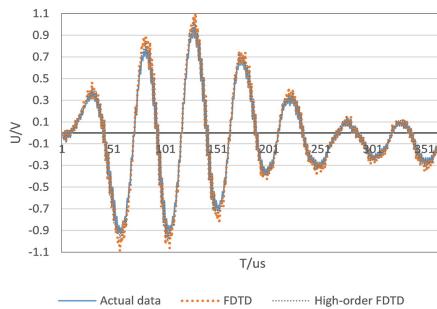


Fig. 5. Receiving transducer simulation result

$$R^2 = \left[\sum_{i=0}^n (y_i^* - \bar{y})^2 \right] / \left[\sum_{i=0}^n (y_i - \bar{y})^2 \right] \quad (25)$$

where y_i represents the actual data, y_i^* represents the model data and \bar{y} represents the average data of y_i .

The determinable coefficient of the acoustical field model obtained by the FDTD (2, 4) algorithm is represented by R_4^2 and its value is 0.9451, and by the FDTD (2, 2) is $R_2^2 = 0.9275$. That is, the determinable coefficient obtained by the fourth-order FDTD algorithm increased by about 2% relative to the second-order, indicating that the calculation accuracy of the FDTD (2, 4) algorithm is higher.

5 Conclusion

The field of transmission channel of steel-through communication system is analyzed and modeled by FDTD method, including the modeling of transducer and steel plate. And the propagation process of ultrasonic transmission in steel-through communication system is simulated based on MATLAB platform. In order to improve the accuracy of the FDTD sound field model, this paper considers the influence of longitudinal and shear waves, and uses higher-order method to improve the FDTD model. In addition, it is demonstrated that the fourth-order FDTD is more stable than second-order method, and the former model is more accurate that its accuracy is improved by 2%.

The ultrasonic acoustical field is modeled for the transducer and steel respectively by FDTD method in the steel-through system. It is of great significance to analysis communication channel. The propagation process of the ultrasonic wave is observed visually by the simulation of the acoustical field, which provides a basis for analyzing the characteristics of the ultrasonic wave in the steel. By visualizing the sound field in the steel-through system, it can observe the propagation process of ultrasonic wave intuitively, and provide a theoretical basis for the subsequent research.

References

1. Roa-Prada, S., Scarton, H.A., Saulnier, G.J., et al.: An ultrasonic through-wall communication (UTWC) system model. *J. Vib. Acoust.* **135**(1), 011004 (2013). <https://doi.org/10.1115/1.4007565>
2. Xie, Y., Yin, W., Liu, Z., et al.: Simulation of ultrasonic and EMAT arrays using FEM and FDTD. *Ultrasonics* **66**, 154–165 (2016). <https://doi.org/10.1016/j.ultras.2015.10.020>
3. Wei, F.X., Yu, G.X., Wang, Z.L.: Ultrasonic testing numerical simulation of austenitic stainless steel pipeline welding cladding. *Appl. Mech. Mater.* **687**, 878–881 (2014). <https://doi.org/10.4028/www.scientific.net/AMM.687-691.878>
4. Lei, J., Zhang, C., Bui, T.Q., et al.: Transient dynamic interface crack analysis in magnetoelastic bi-materials by a time-domain BEM. *Eur. J. Mech.* **49**, 146–157 (2015). <https://doi.org/10.1016/j.euromechsol.2014.07.010>
5. Yee, K.: Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Trans. Antennas Propag.* **14**(3), 302–307 (1966). <https://doi.org/10.1109/TAP.1966.1138693>
6. Naito H, Bolander J.E.: Damage detection method for RC members using local vibration testing. *Eng. Struct.* **178**(JAN.1), 361–374 (2019). <https://doi.org/10.1016/j.engstruct.2018.10.031>
7. Chady, T., Lopato, P.: Testing of glass-fiber reinforced composite materials using terahertz technique. *Int. J. Appl. Electromag. Mech.* **3**(3), 1599–1605 (2010). <https://doi.org/10.3233/JAE-2010-1290>
8. Chen, H., Zhou, H., Sheng, S.: General rectangular grid-based time-space domain high-order finite-difference methods for modeling scalar wave propagation. *J. Appl. Geophys.* **133**, 141–156 (2016). <https://doi.org/10.1016/j.jappgeo.2016.07.021>

9. Niu, Q.X., Liu, Y.J., Song, D.J., et al.: Research of anti-ultraviolet nano-film structure based on the FDTD method. *Optik - Int. J. Light Electron Opt.* **127**(2), 539–543 (2015). <https://doi.org/10.1016/j.ijleo.2015.10.042>
10. Tian, H., Fu, Z.Q.: Study on the sandwich piezoelectric ceramic ultrasonic transducer in thickness vibration. *J. Shaanxi Norm. Univ.* **40**(04), 34–38 (2012). <https://doi.org/10.15983/j.cnki.jsnu.2012.04.004>
11. Li, T.B.: Computational Acoustics: Equations and Calculation Methods for Sound Fields. Science Press, Beijing (2005)
12. Bui, N., Guiffaut, C., Reineix, A., et al.: A new conservative high-order modified FDTD (2,4) Scheme. *IEEE Trans. Antennas Propag.* **65**(1), 269–277 (2017). <https://doi.org/10.1109/TAP.2016.2627568>
13. Sullivan, D.M.: Electromagnetic Simulation Using the FDTD Method. IEEE Press, USA (2000). <https://doi.org/10.1109/9780470544518.index>
14. Bao, W., Teixeira, F.L.: Performance analysis of perfectly matched layers applied to spherical FDTD grids. *IEEE Trans. Antennas Propag.* **66**(2), 1035–1039 (2018). <https://doi.org/10.1109/TAP.2017.2781221>
15. Liu, H., Zheng, Y., Shen, J.: Goodness-of-fit measures of R^2 for repeated measures mixed effect models. *J. Appl. Stat.* **35**(10), 1081–1092 (2008). <https://doi.org/10.1080/02664760802124422>



Defect Sample Generation System Based on DCGAN for Glass Package Electrical Connectors

Qunpo Liu^{1,2(✉)}, Mengke Wang¹, and Naohiko Hanajima³

¹ Department of Robotics Engineering, Henan Polytechnic University, 2001 Century Avenue, Jiaozuo 454003, Henan, People's Republic of China
1qpny@hpu.edu.cn

² Henan International Joint Laboratory of Direct Drive and Control of Intelligent Equipment, 2001 Century Avenue, Jiaozuo 454003, Henan, People's Republic of China

³ College of Information and Systems, Muroran Institute of Technology, 27-1 Mizumoto-cho, Muroran-shi, Hokkaido 050-8585, Japan
<http://www.hpu.edu.cn/>

Abstract. Precision glass packaged electrical connectors are widely used to connect precision modular electrical appliances. Manual detection is currently used due to its small size. In order to improve the detection efficiency and accuracy, deep learning is applied to the automatic detection of the defect. In this paper, in order to solve the problem of lack of defect samples, a defect sample generation system for micro-precision glass package electrical connectors based on deep convolutional generation confrontation network (DCGAN) is constructed. A network model of defect samples based on DCGAN is constructed, and algorithm design, parameter settings, network training and experiments are performed. More than 1,500 samples were generated. Finally, the generated samples were tested for defects (including missing blocks and bubbles). The results show that the samples generated by this method can be used as real samples in defect detection training.

Keywords: Micro-precision glass terminals · DCGAN · Sample generation

1 Introduction

Micro-precision glass encapsulated electrical connectors (hereinafter referred to as glass terminals) are widely used in connecting precision modular electrical appliances, mainly composed of low-temperature glass layer, high-temperature glass layer and Kovar alloy column. The white column in the middle is the terminal of the glass terminal, which is usually composed of Kovar alloy. The outermost layer represents the Kovar ring area of the glass terminal. composition. The composite glass column after sintering is shown in Fig. 1.

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 434–441, 2021.
https://doi.org/10.1007/978-981-15-8458-9_46



Fig. 1. Sintered composite glass column

Due to the factors of processes and materials, some glass layers have defects such as bubbles, cracks, and missing blocks. Due to the small size and high accuracy requirements, the current inspection is mainly done manually, which is prone to eye fatigue [1], and it is easy to lead to produce false detection or even leakage Inspection and other issues. The best way to solve this problem is to design an automatic detection system with high precision and accuracy. However, the number of defective samples is too small at present. Therefore, the research on the sample generation technology of glass terminal defect images is also an urgent issue that needs attention. The most popular generative models in recent years include autoencoders, autoregressive models, and generative confrontation networks [2]. Generative adversarial networks (GAN) [3] is a generative model. Based on game theory, it alternately trains a set of generators and discriminators to learn the feature distribution of data. It is possible to learn feature representation without labeling samples. This article uses this method to generate samples.

2 A Network Model of Glass Terminal Defect Sample Generation Based on a Deep Convolutional Generation Confrontation Network

The GAN model has the disadvantages of easy collapse and difficult to achieve dynamic balance during the training process. Therefore, a convolutional neural network with strong feature extraction capability is introduced into the generation model for unsupervised training to improve the learning effect of GAN [4], Forming a DCGAN [5] neural network model that combines a fully convolutional neural network and a generational confrontation network to make its training process more stable and generate higher quality glass terminal samples.

The network structure of the glass terminal defect sample image generation based on DCGAN is shown in Fig. 2. First, the end surface image of the glass terminal is collected to obtain the original data set, and a Gaussian noise z is input into the generation model. The Gaussian noise z passes through a layer

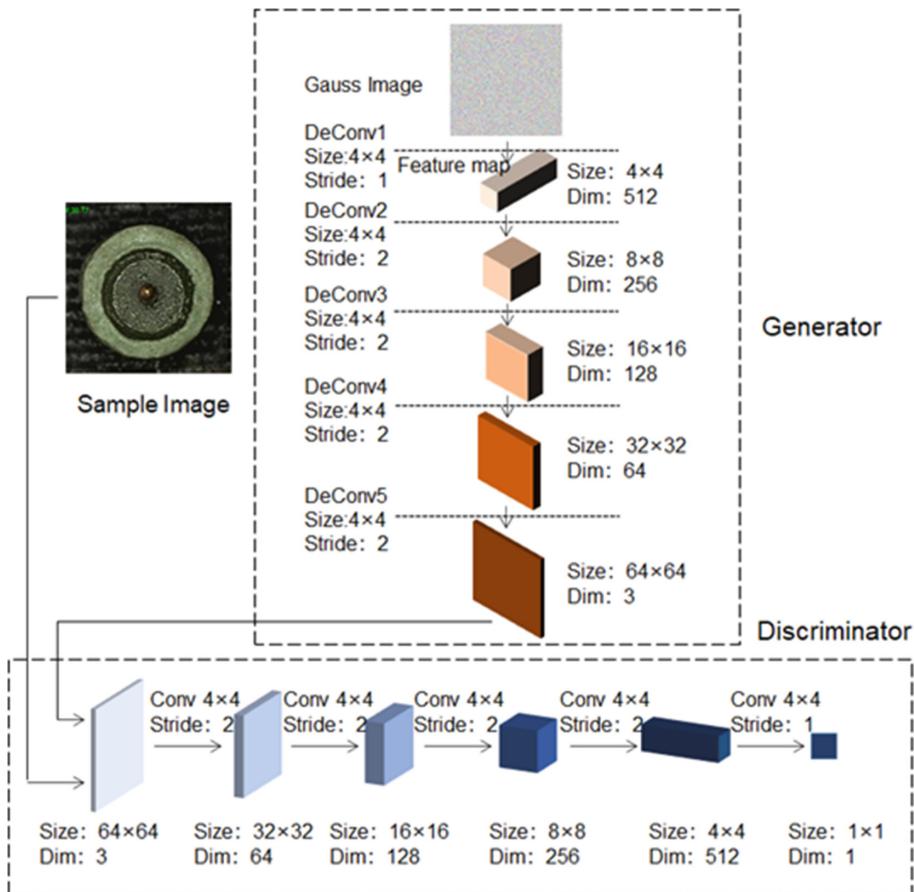


Fig. 2. The network structure of the glass terminal sample image generation

of convolution kernels in the generation model. The deconvolution layer with a layer convolution kernel size of 4×4 and a step size of 2 outputs a feature map with a size of 64×64 and a dimension of 3. This feature map and the collected original data are passed into the DCGAN discriminant model. After a series of convolution operations with a convolution kernel size of 4×4 in the generation model, the final size is 1×1 and the dimension is 1. Fake glass terminal sample image, at this time DCGAN uses the original glass terminal data set and the glass terminal image generated by the generator to train and continuously optimize its own parameters until the discriminator cannot distinguish the source of the glass terminal image, then the training process ends.

3 Model Training Mechanism

The idea of generating an adversarial network (GAN) is to learn from the competition, and to complete the glass terminal defect sample generation through the adversarial training of the two networks of generative model G and discriminant model D. The generative model tries to make the defect sample data generated by itself deceive the discriminant model as much as possible. The discriminant model correctly discriminates whether the data is real or generated by the generative model as much as possible. Such alternating training is equivalent to a two-player game [6] process. As the training continues, both the generative model and the discriminant model are constantly confronting and evolving, and finally the two models reach a dynamic balance (Nash equilibrium). At this time, the sample data generated by the generative model is very close to the real glass terminal image data Distribution, and the discriminant model cannot identify the correct source of the data.

During the training, one side is fixed to update the other side's weight, and both sides strive to optimize themselves so that they can compete against each other and play games with each other, and finally reach a dynamic balance. The training process can be divided into two parts, the first part is to update the parameters of the discriminator, and the second part is to update the parameters of the generator. The mathematical representation of this process is as follows:

$$V(G, D) = E_{x \sim P_{data}} [\log D(x)] + E_{x \sim P_G} [\log(1 - D(X))] \min_G \max_D V(G, D) \quad (1)$$

During training, we first fix the generator G, take an appropriate discriminator D so that V (G, D) gets the maximum value, and the integral is:

$$\begin{aligned} V &= E_{x \sim P_{data}} [\log D(x)] + E_{x \sim P_G} [\log(1 - D(X))] \\ &= \int_x [P_{data}(x) \log D(x) + P_G(x)(1 - \log D(x))] dx \end{aligned} \quad (2)$$

To maximize the integral, that is, to obtain an optimal D * to maximize the following formula:

$$P_{data} \log D(x) + P_G(x) \log(1 - D(x)) \quad (3)$$

Under the premise that the data set is given and the generator G is fixed, Pdata (x) and PG (x) can be regarded as constants, so that V (D) can be maximized:

$$D^*(x) = \frac{P_{data}(x)}{P_{data}(x) + P_G(x)} \quad (4)$$

We bring the optimal discriminator D back to V (G, D), then the optimal function of the generator is obtained:

$$\begin{aligned} \max V(G, D) &= V(G, D^*) \\ &= E_{x \sim P_{data}} [\log \frac{P_{data}(x)}{P_{data}(x) + P_G(x)}] + E_{x \sim P_G} [\log \frac{P_G(x)}{P_{data}(x) + P_G(x)}] \end{aligned} \quad (5)$$

4 Algorithm Core Parameter Configuration

The generator nz is set to 100, the size ngf of the generator output feature map is set to 64, and the number of channels nc of the output image is set to 3 (RGB image). The pooling layer is not used in this article, because the use of convolution is equivalent to letting the network learn the way of pooling at the same time, and using the Leaky Relu activation function [7] can speed up the gradient propagation and help training.

The loss function used is a binary cross entropy loss (BCELoss) [8–10]. The label of the real glass terminal picture is 1, the label of the generated glass terminal picture is 0, and then two optimizers are defined respectively. The Adam optimizer with a learning rate of 0.0002 and a beta1 of 0.5 is used. The z vector comes from a Gaussian distribution. During the training process, a fixed noise is periodically added to the generator to see how the image slowly deforms from the noise.

After the parameter configuration is completed, the end-face picture data of the glass terminal is first acquired and processed uniformly, including reading the glass terminal image, center-cutting the image, regularizing, etc. to obtain a total of 200 original samples. We then perform network training and glass terminal sample image generation according to the process shown in Fig. 2.

5 Experimental Results and Analysis

This design is based on the idea of DCGAN. Through the game process between the discriminator and the generator, the accuracy of the generator and the authenticity of the generated samples are continuously improved, so as to complete the sample generation of glass terminals. As shown in Fig. 3, there are actually more than 1,500 samples generated, and an example of the output glass terminal image of the generator is shown in Fig. 4.

The purpose of generating the image of the glass terminal sample is to facilitate the use of the computer to detect defects and improve the detection accuracy and efficiency. Therefore, the generated sample should be tested, considering whether the glass terminal sample generated by this design is usable, and the sample detection of missing blocks and bubbles. The results are shown in Fig. 5 and Fig. 6, respectively.

It can be seen from the results of the missing block detection in Fig. 4 and the bubble detection result in Fig. 5 that the glass terminal sample based on the DCGAN algorithm can be used as a real picture when it is used for the missing block and bubble detection test.

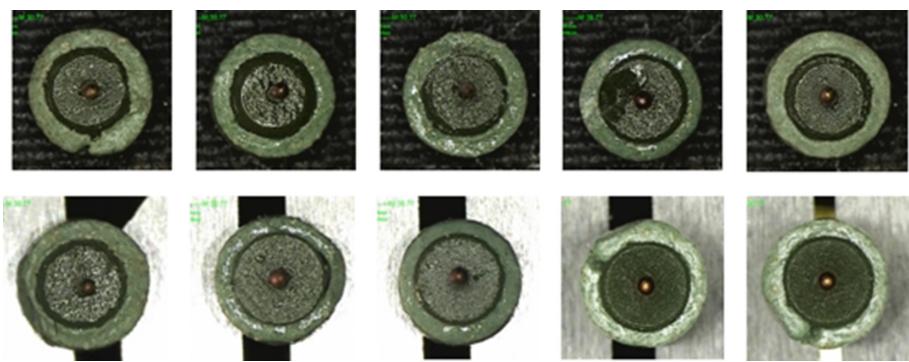


Fig. 3. Examples of original data sets

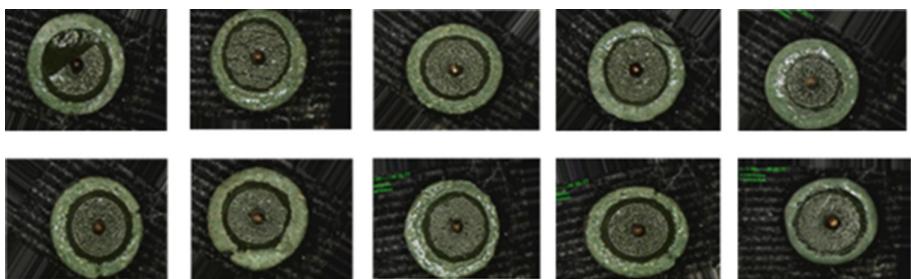


Fig. 4. Example of sample generation

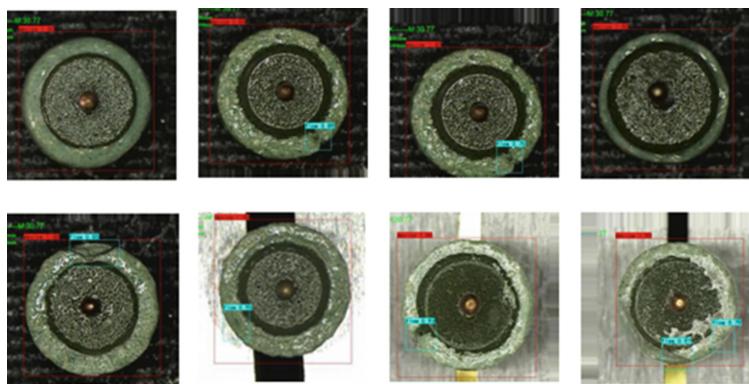


Fig. 5. Examples of missing block sampling test results

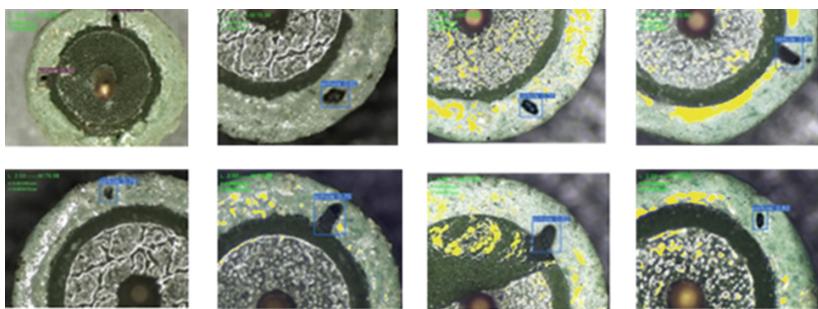


Fig. 6. Example of bubble sampling test results

6 Conclusion

In this paper, from the perspective of generating glass terminal defect samples, the convolutional neural network and the generated confrontation network are fused to generate a realistic glass terminal image. Based on the built DCGAN glass package electrical connector defect sample generation model, after hyper-parameter setting and training, finally 125 defective samples and 75 non-defective samples generate more than 1,500 defective samples. After sampling the generated samples, the results show that the samples generated based on the model constructed in this paper meet the needs of deep learning training and can be used as real samples.

References

1. Zhenteng, H., Jianping, C., Houdong, D., Rui, L.: Research and application of RTV coating quality inspection method based on machine vision. *Insul. Surge Arrest.* **05**, 233–240 (2012). <https://doi.org/10.16188/j.isa.1003-8337.2018.05.041>
2. Mi, Y.: The research of GAN-BASED scheme for radar spectrogram augmentation. Beijing University of Posts and Telecommunications (2019)
3. Goodfellow, I.J., et al.: Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, Canada, vol. 2, pp. 2672–2680 (2014)
4. Yinan, Y., Linhai, Q., Hong, W., Linping, S.: Research on generation technology of small sample data based on generative adversarial network. *J. Electr. Power Constr.* **40**(05), 71–77 (2019). <https://doi.org/10.3969/j.issn.1000-7229.2019.05.009>
5. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks [DB/OL], 07 February 2016
6. Qiang, G., Zhonghao, J.: Amplification of small sample library based on GAN equivalent model. *Electr. Measur. Instrum.* **56**(06), 76–81 (2019). <https://doi.org/10.19753/j.issn1001-1390.2019.06.013>
7. Ye, C., Guan, W.: A review of application of generative adversarial networks. *J Tongji Univ. (Nat. Sci.)* **48**(04), 591–601 (2020). <https://doi.org/10.11908/j.issn.0253-374x.19204>

8. Xiaoshun, W., Dan, C., Haibin, Q.: Least squares transfer generative adversarial networks. *Comput. Eng. Appl.* **55**(14), 24–31 (2019). <https://doi.org/10.3778/j.issn.1002-8331.1903-0430>
9. Yingmin, J.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000). <https://doi.org/10.1109/87.845885>
10. Yingmin, J.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003). <https://doi.org/10.1109/TAC.2003.815033>



A Review of Fault Features for Inter-turn Short Circuit Fault of PM Motor

Liuyang Shen^(✉), Ziyi Fu, and Sumin Han

School of Electrical Engineering and Automation, Henan Polytechnic University,
Jiaozuo 454003, China
1098108971@qq.com

Abstract. With the development of technology in various fields, permanent magnet (PM) motors have been widely used for their simple, efficient, light and small characteristics, and its fault problem has always been concerned by people. Among the faults of PM motors, the common and serious fault is the inter-turn short circuit (ITSC) fault. If it is not diagnosed in time, the fault will expand rapidly and cause other faults. Because of its strong destructiveness, the diagnosis of ITSC fault is crucial. In the process of fault diagnosis, fault feature extraction is a very significant link. According to the different fault feature extraction methods, the fault features in the literatures in recent years are summarized into 5 types in this paper, and fault location, fault severity detection and other evaluation indexes are also used to evaluate all fault features.

Keywords: PM motor · ITSC fault · Fault feature

1 Introduction

The gradual maturity of the control technology of PM motors has made the application of PM motors greatly developed in various fields. Since PM motors often work in complex environments, the motors are prone to failure during long-term operation. In many kinds of faults of PM motor, the ITSC fault is the most common type [1], and it has the extremely negative influence. When the fault occurs, if the fault is incapable of being detected in time and measures are taken, it will cause other types of faults to occur and eventually lead to serious consequences. Therefore, in order to make sure the motor works normally, and make sure the safe and effective production work is carried out smoothly, it is so urgent to discuss the methods of diagnosis about this ITSC fault.

The steps in the whole process of fault diagnosis commonly include: first of all, monitor the fault semaphore, then process signals and extract features of the measured signals to obtain the fault features, and finally, the fault features can be directly compared with the threshold set by the fault diagnosis, if the threshold is exceeded, the fault will be alarmed. Certainly, the fault features can also be identified and classified by intelligent algorithms to achieve the purpose of fault diagnosis.

Sometimes, the fault semaphore that directly measured is inconspicuous, or the fault information and the fault type are not one-to-one correspondence, so the extraction of fault features is indispensable. Above all the steps, it is significant to extract fault features [2]. In recent years, many literatures have proposed many descriptions of fault features of ITSC in PM motor. In order to make these fault features better referred and utilized by scholars, and to bring some convenience for scholars in this research direction, what follows of this paper is the summary and evaluation of fault features.

2 Fault Features

2.1 Fault Features Based on Frequency Analysis

(1) *Harmonic of Stator Current.* Literature [3] shows that when the fault occurs, the stator current will generate harmonics in the motor. The harmonic frequency is:

$$f_{fault} = \left(1 \pm \frac{2k+1}{p}\right)f_s, k = 1, 2, 3, \dots \quad (1)$$

Therefore, the harmonic of stator current is regarded as the fault feature, and LDA classifier is applied to gauging the ITSC fault, then fault classification and fault severity judgment are carried out.

(2) *Second Harmonic of q-axis Current.* In literature [4], FFT analysis method of *q*-axis current is carried out, and the second-order harmonic component is taken as the monitoring object, which is constantly compared with the corresponding value under non-fault conditions. Finally, the online fault diagnosis is achieved. The method of extracting the harmonic component of stator current by FFT analysis method is good at simple calculation and realization, but for nonlinear and non-stationary signals, FFT analysis method has limitations.

(3) *Energy of WT Coefficient.* In literature [5], the stator current is disintegrated through coif5 wavelet function, then the energy of d2 decomposition coefficient about phase current is extracted as a feature. The PNN is established for fault diagnosis. Although the wavelet function has a good application in non-stationary domain, the accuracy with time and frequency is difficult to achieve both, and it has a strong dependence on the choice of wavelet basis function. It can not adapt to the dynamic change of motor operation, so its practical application is limited.

(4) *Standard Deviation of IMF Based on HHT.* In literature [6], the HHT analysis method is applied to stator current and the standard deviation of IF of the first IMF is regarded as the fault feature, then it compares with the threshold value. This method is able to realize the diagnosis with different levels of fault severity degree. It is proved that the HHT analysis method is capable of detecting the early ITSC fault. Although HHT analysis method can make up for the limitations of wavelet analysis method, it is only suitable for time-frequency analysis of narrow-band signals, which limits its practical application.

(5) *Third Harmonic Amplitude of Stator Current Based on PSD*. In literature [7], the power spectral density (PSD) method is applied to solve the frequency spectrum of stator current. The result indicates that the third harmonic amplitude of stator phase current can be regarded as the fault feature to test the fault, and it is influenced with the severity of the fault. This method can detect the fault location and the fault severity.

(6) *Second Harmonic of DC Current*. In literature [8], based on the new PMA model proposed in this paper, the second harmonic of the DC output current in the model is deemed as the fault feature to detect ITSC fault. In order to predict the temperature between the fault turns, the PMA model is coupled with the model based on LP and Fe to realize the estimation of winding life.

(7) *Slope and Correlation Coefficient of Normalized Current Third Harmonic*. Literature [9] considers that the function of fundamental frequency is usually greater than harmonic amplitude, so it is necessary to filter out the fundamental frequency. The general time-frequency analysis method calculates at least a whole bandwidth of the spectrum when filtering the fundamental frequency, which requires a large amount of calculation. Therefore, the article uses the order tracking algorithm VFK-OT of vold-kalman filtering, which uses an adaptive band-pass filter to extract useful harmonic frequencies, which greatly reduces the amount of calculation. On this basis, the article defines two fault features: the normalized current third harmonic slope m and the correlation coefficient r . The experimental results prove that the method is trustworthy and sensitive when operating at non-stationary speeds, but it cannot achieve fault location.

(8) *Voltage Frequency Feature Based on Space Vector*. In literature [10], a combined space vector D is constructed by two different coordinate systems. The information in the two different coordinate systems is integrated into the associative space vector $\vec{D} = \vec{v}_{\alpha\beta} \cdot \vec{v}_{\alpha 2\beta 2}$. The article performed a full spectrum analysis of the voltage and current in space vector D, and the frequency component is taken as the fault feature for fault detection. This method can realize the early inter-turn fault diagnosis at different working points and different fault degrees, and it has good robustness.

(9) *Mixed Signal Feature*. Literature [11] separately analyzed the vibration signal and current signal of the permanent magnet motor by FFT and wavelet packet to obtain their frequency spectrum and used the method which called correlation analysis to fuse the vibration video features and current time-frequency features to obtain the fault features. This method can get more obvious fault features and improve the sensitivity of fault detection.

(10) *Second Harmonic of Value Function Based on Model Predictive*. Literature [12] shows that the DC component and the second harmonic component produced by the Value function based model predictive is the most obvious at the time with the fault, so the DC component and the second harmonic component

of the Value function is applied to detect the fault. This method does not require additional equipment and sampling work, and it effectively suppress the spread of fault, but it can not achieve fault location.

(11) *Higher Harmonic Amplitude of Air Gap Flux Density.* When the ITSC fault emerges in PM motor, the air gap flux density of PM motor will be affected. In literature [13], the various degrees of fault are analyzed. The higher harmonic amplitude of air gap flux density is regarded as the fault feature to identify the fault winding, and the fundamental amplitude is regarded as the fault feature to determine the severity of fault.

(12) *Back EMF Fundamental Frequency.* In literature [14], the air gap flux density is measured by the search coil, and the back EMF generated by it is divided into the magnetic field component and the armature component. Take amplitude of fundamental component as the fault feature. Because the fault feature is the amplitude of the fundamental wave, it will not be affected by harmonic and is effective under any load conditions. But this method is an invasive fault diagnosis method, which must be added with search coil when the permanent magnet motor is manufactured, so the flexibility of this method is poor.

(13) *Higher Harmonic of Flux Generated by Inverter.* In literature [15], the fault feature is obtained by calculating the voltage on the search coil and the high-order harmonic of flux which is generated by the inverter power supply in the motor. The fault feature α_i is contrasted with the threshold value, and the result can detect whether the fault occurs, and the subscript of the fault feature can determine the fault location concurrently. The fault feature F_S is able to identify the fault severity. This method is capable of realizing the fault location and fault severity degree detection accurately and intuitively, but the installation of the search coil increases the detection cost.

2.2 Fault Features Based on Parameter Estimation

(1) *Estimation of Short Circuit Turns.* In literature [16], the fault features are constructed by extended Kalman filter (EFK) method, and the calculation method of adaptive threshold is given:

$$\text{Fault_indicator} = \left\langle \left(\sum_{i=A,B,C} |n_{s/c}|_i - n_{s/c_ref} \right) \right\rangle_{T/2} \quad (2)$$

$$n_{s/c_ref} = \left\langle \sum_{i=A,B,C} |n_{s/c}|_i \right\rangle_{5T} \quad (3)$$

The fault feature can realize online monitoring of short-circuit turns ratio, and the speed and robustness of fault diagnosis are better.

(2) *Parameter Estimation of Permanent Magnet Motor.* In literature [17], the least square method is employed in estimating the parameters of PM motor. The fault feature is defined as the value of structural distance among the extended park model and the PM synchronous motor. This feature can achieve the stable and sensitive fault detection.

(3) *Parameter Estimation of Fault Phase and Severity.* In literature [18], two open-loop observers are applied to estimate the fault phase parameter P and the fault severity parameter G , and PSO algorithm is applied to optimize the observer parameters. This method can detect fault phase and fault severity, and has strong robustness to parameter uncertainty, but it can not detect minor fault.

(4) *Parameter Estimation of Inductance Value.* In literature [19], based on the consideration of the distribution of stator teeth, rotor yoke and stator winding, the self-induction value of fault phase is estimated by using the model of PM motor with magnetic equivalent circuit, and it is considered as the fault feature. This method can realize fault detection, and the fault severity can also be determined.

2.3 Fault Features Based on Symmetric Component Analysis

(1) *Zero Sequence Component.* Literature [20] proposes a diagnosis strategy using zero sequence component fault features. It obtains the zero-sequence voltage and current through calculation, and two new fault features FI and d_j are defined using the amplitude of the zero sequence component and the initial phase of the zero sequence component. The fault features are defined as:

$$FI = \frac{\mu i_f}{\omega_r} \quad (4)$$

$$d_j = \begin{cases} |\theta_{i1} - \theta_j|, |\theta_{i1} - \theta_j| \leq \pi \\ 2\pi - |\theta_{i1} - \theta_j|, |\theta_{i1} - \theta_j| > \pi \end{cases} \quad (5)$$

This method is stable with variation of speed and load, and it not only estimate the fault severity but also locate the fault phase [21].

(2) *Negative Sequence Component.* Literature [22] proposes a diagnosis strategy using negative sequence component fault features. After the conversion of negative sequence current and voltage, the proposed fault features are not affected by the rotation speed, but only related to the fault parameters μ and R_f .

(3) *Ratio of Negative Sequence Voltage to Positive Sequence Voltage.* Literature [23] performs $\alpha\beta$ coordinate transformation on the reference voltage of the PI controller. Then it calculate the positive and negative sequence components of

the transformed reference voltage as fault feature. The fault features is defined as follows:

$$R_{NP} = \sqrt{\frac{(V_\alpha^-)^2 + (V_\beta^-)^2}{(V_\alpha^+)^2 + (V_\beta^+)^2}} \quad (6)$$

This method is simple and fast, but it can not realize fault location.

2.4 Fault Features Based on Difference

(1) *Average Value of d-axis Voltage Difference.* Literature [24] considers the operation mode about PM motor, then samples the voltage of the fault winding and the healthy winding, then calculates the d-axis voltage difference between them, and the average value in 20 sampling periods is calculated. The sign of the average value is used to determine the faulty winding.

(2) *Current Residual.* Literature [25] the estimated current value of the health model is different from the actual detected fault current value, then take the obtained current residual value as the fault feature and contrasts with the set threshold. The method proposed in the literature can get the better effect under the condition of steady-state operation and needs to be adaptive in the transient mode. This method combines the advantages of two methods that are negative sequence analysis and current residual, which improves the sensitiveness and steadiness of fault detection.

(3) *Instantaneous Reactive Power Difference.* In literature [26], the instantaneous power of the fault motor is contrasted with the healthy motor, and the D-value between the two motors is computed. Then the average value of that D-value obtained in K cycles is considered as the fault feature, which is Related to the number of short circuit turns. So this fault feature is able to realize online fault diagnosis.

2.5 Fault Features Based on Intelligent Algorithm

(1) *Fault Features Based on Deep Learning.* Literature [27] uses a deep learning model to extract the fault feature about negative sequence current. The fault sample is expanded by using the generative countermeasure network. Then, the feature of fault sample is learned by using sparse self coding network. This method improves the deficiency that numerous fault samples are needed for deep learning, and accomplishes the fault diagnosis and does not require much sampling information, and has good integration and robustness.

(2) *Fault Features Based on Grey Correlation Theory*. In Literature [28], several typical fault symptoms are considered as the assessment index of grey correlation analysis. The fault signals that are Waiting for processing and the fault samples are analyzed by grey correlation degree. Then, the results are weighted by entropy weight theory, and based on this theory, the most prominent fault features are enhanced and the unimportant fault features are attenuated. This method can reduce the insufficiency of the single feature with obvious error and improve the efficiency Fault detection accuracy. At the same time, this method is straightforward, and it is not required additional economic cost.

(3) *Fault Features Based on ANN*. In Literature [29], the frequency spectrum of stator current with different fault severity is analyzed under different load and speed. The results show that the amplitude of third harmonic of stator current and the severity of the fault are in the direct ratio, so the amplitude of the third harmonic is trained as the input of ANN in this paper. This method can realize fault detection and fault severity identification.

2.6 Evaluation of Fault Features

In this section, the fault features summarized above are compared based on several evaluation indexes. Table 1 lists the matching status between each fault feature and evaluation index. In Table 1:

- A means Able to realize fault location.
- B means Able to detect fault severity.
- C means No additional sensors required.
- D means Able to realize online fault diagnosis.
- E means Fault features are non-invasive.
- F means Verified by experiments.
- G means Considering saturation.
- H means Satisfy robustness.
- I means Suitable for inverter power supply.
- ✓ means satisfy the evaluation indexes.
- ✗ means dissatisfy the evaluation indexes.
- * means uncertain.

Table 1. Compare of fault features.

2.1														2.2				2.3				2.4				2.5			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
A	*	*	*	*	✓	*	✗	✗	*	*	✓	✓	✓	✓	✓	✓	*	✗	✓	*	*	✗	*	*	*	*	*	*	
B	✓	✗	✗	✗	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	*	*	*	*	*	*	*	*		
C	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗	✗	✗	✓	✗	✗	✓	✓	✓	✓	✓	✗	✗	✗	✗	✓			
D	*	✓	✓	✓	*	✓	✓	✓	✓	✓	*	✓	✓	✓	✓	*	*	✓	*	✓	✓	*	✓	✓	✓	✓	*		
E	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	*	✓	✓			
F	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
G	*	*	*	*	*	*	✓	*	*	*	*	*	*	*	*	*	✓	✓	*	✓	*	*	*	*	*	*	*		
H	*	*	*	✓	*	✓	✓	*	*	*	✓	✓	✓	*	✓	*	✓	*	✓	✓	✓	*	✓	✓	✓	✓	✓		
I	*	*	*	*	*	*	✓	✓	*	*	*	*	✓	✓	*	*	✓	*	✓	*	✓	*	*	*	*	*	*		

3 Conclusions

This paper summarizes the fault features of ITSC fault in PM motor in recent years based on 5 kinds different fault extraction methods. Because the fault diagnosis capabilities of each fault features are different, as well as the applicable working conditions are distinct, finally, the paper analyses 26 fault features with 9 evaluation indexes. This paper can provide some reference for research of ITSC fault diagnosis of PM motor.

Acknowledgments. This work was supported by the National Key Research and Development Program. Grant no. 2016YFC0600906.

References

1. Hu, H.Q., Shi, W.F., Venditti, B.: Research on modeling and fault diagnosis of inter-turn short fault of permanent magnet synchronous motor. *J. Comput. Methods Sci. Eng.* 1–15 (2020). <https://doi.org/10.3233/JCM-194127>
2. Faiz, J., Nejadi-Koti, H., Valipour, Z.: Comprehensive review on inter-turn fault indexes in permanent magnet motors. *IET Electr. Power Appl.* **11**, 142–156 (2017). <https://doi.org/10.1049/iet-epa.2016.0196>
3. Haddad, R.Z., Strangas, E.G.: On the accuracy of fault detection and separation in permanent magnet synchronous machines using MCSA/MVSA and LDA. *IEEE Trans. Energy Convers.* **31**, 924–934 (2016). <https://doi.org/10.1109/TEC.2016.2558183>
4. Peng, W., Zhao, F., Wang, Y.X., Guan, T.Y.: On line detection method for inter turn short circuit fault of PMSM. *Adv. Technol. Electr. Eng. Energy* **37**, 41–48 (2018). <https://doi.org/10.12067/ATEEE1703103>
5. Fu, Z.Y., Liu, J.L., Zhang, X.X.: Research on inter-turn short circuit fault diagnosis of dual redundancy permanent magnet brushless DC motor. *Trans. China Electrotech. Soc.* **29**, 104–109 (2014). <https://doi.org/10.19595/j.cnki.1000-6753.tces.2014.01.015>
6. Alvarez-Gonzalez, F., Griffó, A., Wang, B.: Permanent magnet synchronous machine stator windings fault detection by HilbertCHuang transform. *J. Eng.* **2019**, 3505–3509 (2019). <https://doi.org/10.1049/joe.2018.8173>
7. Yassa, N., Racheck, M.: Modeling and detecting the stator winding inter turn fault of permanent magnet synchronous motors using stator current signature analysis. *Math. Comput. Simul.* **167**, 325–339 (2020). <https://doi.org/10.1016/j.matcom.2018.04.012>
8. Chen, L., Wang, J., Sun, Z.: Electromagnetic-thermal coupled modelling and analysis of inter-turn short-circuit faults of a permanent magnet alternator. *J. Eng.* **2019**, 4426–4431 (2019). <https://doi.org/10.1049/joe.2018.8053>
9. Urresty, J., Riba, J., Romeral, L.: Diagnosis of interturn faults in PMSMs operating under nonstationary conditions by applying order tracking filtering. *IEEE Trans. Power Electron.* **28**, 507–515 (2013). <https://doi.org/10.1109/TPEL.2012.2198077>
10. Immovilli, F., Bianchini, C., Lorenzani, E., Bellini, A., Fornasiero, E.: Evaluation of combined reference frame transformation for interturn fault detection in permanent-magnet multiphase machines. *IEEE Trans. Ind. Electron.* **62**, 1912–1920 (2015). <https://doi.org/10.1109/TIE.2014.2348945>

11. Chen, Y., Liang, H., Wang, C.D., Liang, S.Y., Zhong, R.Q.: Detection of stator inter-turn short-circuit fault in PMSM based on improved wavelet packet transform and signal fusion. *Trans. China Electrotech. Soc.* **35**, 228–234 (2020). <https://doi.org/10.19595/j.cnki.1000-6753.tces.180354>
12. Ding, S.C., Wang, Q.M., Hang, J., Hua, W., Wang, Q.J.: Inter-turn fault diagnosis of permanent magnet synchronous machine considering model predictive control. *Proc. CSEE* **39**, 3697–3708 (2019). [https://doi.org/10.13334/j.0258-8013.pcsee](https://doi.org/10.13334/j.0258-8013.pcsee0.13334/j.0258-8013.pcsee)
13. Qiu, H., Zhao, X., Yang, C., et al.: Influence of inter-turn short-circuit fault considering loop current on electromagnetic field of high-speed permanent magnet generator with Gramme ring windings. *J. Electr. Eng. Technol.* **14**, 701–710 (2019). <https://doi.org/10.1007/s42835-019-00122-z>
14. Da, Y., Shi, X., Krishnamurthy, M.: A new approach to fault diagnostics for permanent magnet synchronous machines using electromagnetic signature analysis. *IEEE Trans. Power Electron.* **28**, 4104–4112 (2013). <https://doi.org/10.1109/TPEL.2012.2227808>
15. Zeng, C., Huang, S., Yang, Y., Wu, D.: Inter-turn fault diagnosis of permanent magnet synchronous machine based on tooth magnetic flux analysis. *IET Electr. Power Appl.* **12**, 837–844 (2018). <https://doi.org/10.1049/iet-epa.2017.0865>
16. Aubert, B., Rgnier, J., Caux, S., Alejo, D.: Stator winding fault diagnosis in permanent magnet synchronous generators based on short-circuited turns identification using extended Kalman filter. *Acta IMEKO* **3**, 4–9 (2014). <https://doi.org/10.21014/acta.imeko.v3i4.146>
17. Liu, Z.: Simulation for fault diagnosis of permanent magnet motor based on mathematics model. *Micromotors* **48**, 60–64+87 (2015). <https://doi.org/10.15934/j.cnki.micromotors.2015.12.014>
18. Lee, J., Moon, S., Jeong, H., Kim, S.W.: Robust diagnosis method based on parameter estimation for an interturn short-circuit fault in multipole PMSM under high-speed operation. *Sensors* **15**, 29452–29466 (2015). <https://doi.org/10.3390/s151129452>
19. Faiz, J., Nejadi-Koti, H., Exiri, A.H.: Inductance-based inter-turn fault detection in permanent magnet synchronous machine using magnetic equivalent circuit model. *Electr. Power Compon. Syst.* **45**, 1016–1030 (2017). <https://doi.org/10.1080/15325008.2017.1293196>
20. Hang, J., Zhang, J., Cheng, M., Huang, J.: Online interturn fault diagnosis of permanent magnet synchronous machine using zero-sequence components. *IEEE Trans. Power Electron.* **30**, 6731–6741 (2015). <https://doi.org/10.1109/TPEL.2015.2388493>
21. Qi, Y., Bostancı, E., Gurusamy, V., Akin, B.: A comprehensive analysis of short-circuit current behavior in PMSM interturn short-circuit faults. *IEEE Trans. Electron.* **33**, 10784–10793 (2018). <https://doi.org/10.1109/TPEL.2018.2809668>
22. Jeong, H., Moon, S., Lee, J., Kim, S.W.: Inter-turn short fault diagnosis of permanent magnet synchronous machines using negative sequence components. In: 2016 IEEE International Conference on Industrial Technology (ICIT), pp. 170–174. IEEE Press, Taipei (2016). <https://doi.org/10.1109/ICIT.2016.7474745>
23. Meinguet, F., Semail, E., Kestelyn, X., Mollet, Y., Gyselinck, J.: Change-detection algorithm for short-circuit fault detection in closed-loop AC drives. *IET Electr. Power Appl.* **8**, 165–177 (2014). <https://doi.org/10.1049/iet-epa.2012.0316>
24. Chen, Y., Chen, X., Shen, Y.: On-line detection of coil inter-turn short circuit faults in dual-redundancy permanent magnet synchronous motors. *Energies* **11**, 662 (2018). <https://doi.org/10.3390/en11030662>

25. Hu, R., Wang, J.B., Mills, A., Chong, E., Sun, Z.: Current residual based stator inter-turn fault detection in permanent magnet machines. *IEEE Trans. Ind. Electron.* (2020). <https://doi.org/10.1109/TIE.2020.2965500>
26. Chen, Y.G., Zhao, X.B., Yang, Y.K., Shi, Y.C.: Online diagnosis of inter-turn short circuit for dual-redundancy permanent magnet synchronous motor based on reactive power difference. *Energies* **12**, 510 (2019). <https://doi.org/10.3390/en12030510>
27. Li, Y.J., Zhang, Z.L., Li, M.H., Wei, H.F., Zhang, Y.: Fault diagnosis of interturn short circuit of permanent magnet synchronous motor based on deep learning. *Electr. Mach. Control.* 1–8 (2018)
28. Zhang, Y., Wang, A., Guo, B., Li, H.: Diagnosis methods for inter-turn short-circuit fault degree of permanent magnet synchronous motor stator winding. In: 2019 6th International Conference on Systems and Informatics (ICSAI), pp. 927–931. IEEE Press, Shanghai (2019). <https://doi.org/10.1109/ICSAI48974.2019.9010126>
29. Cira, F., Arkan, M., Gumus, B.: Detection of stator winding inter-turn short circuit faults in permanent magnet synchronous motors and automatic classification of fault severity via a pattern recognition system. *J. Electr. Eng. Technol.* **11**, 416–424 (2016). <https://doi.org/10.5370/JEET.2016.11.2.416>



Development of On-Line Monitoring System for Cremation Machine Based on PXI Bus

Lin Tian^{1,2(✉)}, Fengguang Huang^{1,3}, Feng Shi¹, and Bingjie Li^{1,2}

¹ 101 Research Institute of Ministry of Civil Affairs, Beijing 100070, China
tianlin23@outlook.com

² Key Laboratory of Cremation, Ministry of Civil Affairs, Beijing 100070, China
³ Key Laboratory of Pollution Control, Ministry of Civil Affairs,
Beijing 100070, China

Abstract. This system is based on the characteristics of cremation equipment operating conditions, automatic detection technology, and environmental protection standards. It uses intelligent sensors to measure the operating temperature of the cremator and its flue gas post-processing equipment, furnace negative pressure, fuel consumption, pollutant emissions and other parameters, Comprehensive and real-time detection of cremation process, and management of sampling data. The system is researched and developed from the following four aspects: using virtual instruments and fieldbus technology to develop testing platforms; upgrading on-site testing data acquisition system hardware; developing testing system software; establishing a multi-channel data acquisition system.

Keywords: PXI bus · LabVIEW · Online monitoring of cremation equipment · Data acquisition

1 Introduction

1.1 Current Status of Cremator Testing Technology

The cremation machine detection systems at home and abroad are all based on the traditional instrument measurement technology. With the development of computer technology, modern detection platforms use computers as the core of software operation and data calculation, plus various sensors, response components and detection objects. The general measurement and control platform is only developed for specific detection objects, and can only complete the general detection function under limited use environment conditions, and the detection data cannot realize inter-platform interoperability. With the development of the industry and technological progress, in the fields of technological innovation such as design and RandD of manufacturing equipment, the amount of data that

needs to be acquired and calculated is exploding, the real-time requirements of the detection results are increasingly higher, and the detection objects are distributed in a spatial range. The scale is getting wider and the scale is getting bigger and bigger, which puts forward new requirements for modern detection systems.

1.2 Main Technical Difficulties

In response to the research goals and content of the industry, in the process of building the structure of the system and implementing the design function, the following key technical problems and difficulties are gradually solved: From the system architecture, software platform, hardware interface, communication method and sensor selection of the detection platform, system architecture complete, reliable system and stable operation are achieved. Choose sensors, electronic components and standard circuits that have been tested and proven to meet the environmental requirements of the inspection site, and reduce the degree of response to the on-site environment during signal transmission. Try to select standard components to ensure the reliability of the sensor during operation. Electromagnetic compatibility processing is performed on the hardware of the entire detection system to improve the anti-interference ability of the equipment. The software structure design takes into account the fault tolerance, to avoid unrecoverable errors when the system is used in the field. For signal transmission, the communication method of cable connection is reasonable and reliable, and the wireless connection ensures that data transmission is not lost.

2 Overview of the Cremator Online Monitoring System

2.1 Data Acquisition Module

This system adopts multiple types of signals, and selects real-time sampling to collect input signals. According to the different types of test data, the signal to be tested is divided into digital signals and analog signals. Different signals require different hardware to process. The signals are sorted and sorted during data collection, and suitable sensors and detection boards are selected to meet the detection accuracy Requirements. Using virtual instrument technology, the test system based on the PXI platform has the characteristics of high test efficiency, fast fault location, high level of automation, high reliability, and flexible operation. This technology has been greatly improved in the design, production, testing and inspection of electronic products. Range of applications [1]. The virtual instrument hardware platform of the detection system mainly includes four categories of PXI bus chassis, controller, signal board and interface equipment. Connect the sensor and the bus chassis through the field cable, and send the signal to each signal board through the interface device developed on the bus chassis, and complete the functions of signal acquisition, amplification, A/D conversion, etc. The controller reads the data in the board according to the sampling needs. According to the response speed and sampling frequency of the

board, the signal acquisition of the board can reach the order of 1000/s. If all the data collected by the board is displayed by the detection software, especially the waveform signal, it will occupy a lot of PXI controller CPU resources bringing great difficulties to the graphics processing of the core computing unit. Therefore, we use LabVIEW's "producer-consumer" model to develop programs. Send the detected data to the buffer of the board first, then set the sampling frequency according to the detection needs, extract the signals in the buffer according to the sampling requirements, and finally display the waveform of the linear signal through the front panel. The principle of data acquisition between the PXI board and the LabVIEW program is shown in Fig. 1.

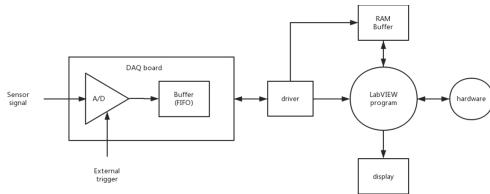


Fig. 1. Schematic diagram of data acquisition

2.2 Data Analysis Module

Within the range of the PXI bus controller's calculation speed of CPU resources and the PXI bus data transmission bandwidth, a reasonable software program can be designed to enable the virtual instrument platform of this detection system, achieve the effect superior to the original single-item detection method corresponding to the single-item. LabVIEW provides DAQ such as user interface development, data acquisition, and database call. The detection system can directly use the drivers on the PXI board. Therefore, the developer of this system analysis module will follow the principles of modules, universality, and portability. The design of the data analysis module meets the technical requirements of the detection platform and is conducive to secondary development and system upgrade.

3 System Hardware Composition

3.1 On-Site Testing Machine Hardware Configuration

The on-site inspection machine of this project adopts PS PXI-9106 portable test platform, which is compatible with PXI/CompactPCI module and has general PXI function. The various sensors configured in the field inspection machine adopt the two-wire or three-wire mode, and the output is a standard signal of

0–10 V or 0–20 mA, which can be directly connected to the 9106 field inspection machine. After the data collection and display is completed, the data is transmitted to the external terminal through the wireless transmitter, and the database is saved at the same time, which is convenient for future query and other operations. The 9106 field test machine includes the following main components: PS PXI-3051 is a dual-core PXI system controller, which can provide an ideal platform for PXI-based measurement and control system applications; PS PXI-3364 is a multi-function based on PXI bus Data acquisition card, providing 32 analog inputs and 16 digital input and output ports, built-in 18-bit ADC, and 4 simultaneous analog outputs, 2 counters, suitable for automated testing of multi-channel analog signals in this system; PS PXI-3281 is A high-precision thermocouple signal acquisition card provides 32-channel thermocouple signal input and 8 built-in cold junction compensation channels. The cold junction compensation thermistor is located on the MDU-3281 near the terminal, which can provide cold junction compensation more accurately Temperature; PS PXI-3305 is a PXI-based digital IO quantity control acquisition card, which provides 64 bidirectional digital IO channels. Each IO can be individually programmed to control input or output, which is convenient for controlling multiple different jobs at the same time on a single card. Used in the environment of voltage equipment.

3.2 Sensor Hardware Configuration

Measurement of Temperature. The temperature of the main combustion chamber during the operation of the cremator is one of the most important performance indicators to measure the combustion state of the cremator. In the field inspection, the data obtained by thermocouple measurement is only the thermo-electric potential difference caused by the temperature difference between the two junctions of the thermocouple. For the cold junction temperature of the thermocouple, correction and compensation are required to refer to the thermocouple index table. Detect temperature. Therefore, in combination with the PS PXI-3281 thermocouple board used in this system, the detection thermocouple is connected to the thermostat junction box MDU-3281 connected to the board, and the thermistor is compensated by the cold end in the junction box near the terminal position. The cold junction compensation temperature can be obtained more accurately. Finally, the electric potential signal is sent to the 9106 on-site testing machine, combined with the cold junction compensation temperature signal, and the detected temperature data is obtained through the program calculation in the background.

Measurement of Furnace Pressure. The negative pressure value of the main combustion chamber of the cremator is also one of the important performance indicators to measure the quality of the combustion state of the cremator, and it is also the main indicator to judge the sealing performance of the furnace body. This project uses the yokogawa EJA120A differential pressure transmitter to measure the negative pressure value of the main combustion chamber.

The pressure difference at both ends of the differential pressure transmitter is automatically converted into a 4–20 mA current signal in the table and is connected to the MDU The 3841-I type IV signal conversion box is converted into a voltage signal after being connected, and then input into the PS PXI-3364 multi-function data acquisition card through the cable and handed over to the PS PXI-3051 controller for calculation the pressure difference between the inside of the main combustion chamber and the environment is finally obtained.

Measurement of Cremation Time. The cremation time of a single body is an important indicator to measure the efficiency of the cremation machine. In this system, by detecting the stroke switch signal of the furnace door of the cremator, input the PS PXI-3305 switch board as the starting point, and use the clock of the controller to measure the time of cremation.

Start and Stop of the Burner. This parameter is an auxiliary parameter describing the working state of the cremation machine during the cremator detection. It can indirectly reflect the effect of the temperature increase in the main combustion chamber by comparing the coupling relationship between the temperature and pressure of the main combustion chamber when the burner is working And the response time of the induced draft fan inverter to the negative pressure adjustment after the negative pressure changes. In this project, by detecting the conduction signal of the burner photocell of the main combustion chamber of the cremator, enter the PS PXI-3305 digital switch card and import it into the main program of the virtual instrument.

Measurement of Fuel Consumption. The fuel consumption of the cremator directly determines the long-term economic cost of the equipment. The elliptic gear flowmeter we use has its own ModBUS-rtu protocol interface, and is connected to the PS PDC-3854-2 signal conversion box of the field inspection machine through a shielded cable, and the data is sent to the PS PXI-3051 main controller for processing.

Measurement of Pollutant Concentration. For single flue gas sensor, the sampling signal processing is relatively cumbersome, so we choose to use testo 350 flue gas analyzer to detect the concentration of pollutants. The device comes with PROFIBUS-DP protocol interface, and it is connected to the PS PDC-3854 of the on-site testing machine through shielded cable -2 The signal conversion box is connected, and the data is sent to the PS PXI-3051 main controller for processing.

4 Data Analysis Software Development

Based on the integration of multi-parameter and multi-signal detection hardware in the quality testing of cremation machines, combined with the specific research

situation, we use LabVIEW as a development platform for the development of online testing system software for cremation machines.

4.1 Software Architecture

In the user test application interface design of the virtual instrument, the traditional instrument equipment operation interface is replaced, and the design and development personnel customize the design interface, and realize the control of the hardware equipment through the software interface.[2] The software architecture of the cremator online detection system is based on the logic of the detection activities and is designed according to the logic of the state transition of the software requirements. The architecture of this software system is based on the basic state structure, and enters different operation interfaces according to the use of inspectors. The sub-VI uses a general “start-main program-end” mode in the construction framework, and uses a fixed sequence in the program to force the sampling sequence to facilitate the system’s error recognition ability in complex detection. Using simple VI mode, only a single VI can complete the basic functions of detection, numerical calculation and display. At the same time, it does not respond to the operation of the interface to avoid the problem of system resource occupation caused by control. In the LabVIEW software platform, according to the sampling program flow of this system, set the While cycle as the main program. First, establish a conditional structure and create an enumeration variable, and connect the enumeration variable to the tunnel of the While loop, and connect the enumeration variable to the branch selector of the conditional structure. Second, use selector labels to add branches for each value. Then, add the elder brother function control on the front panel, and complete the initialization of the front panel boolean control, numerical control, graphic control and other functions. Finally, add and connect intermediate variables to complete each state jump code.

4.2 Sampling Module Implementation

The software flow of the signal acquisition module of this detection system is shown in Fig. 2. When the system enters the start state, the program enters the initialization state, the PXI boards for each data acquisition are started according to the port, the sampling system is in standby, and the state of the furnace door is cyclically detected. Taking the furnace door closing as the instruction to start the system detection, the system enters the sampling state. If the deer door closing instruction is not executed, the system continues to stand by. After sampling starts, the board detects the signal value of each port in turn and sends the value to the board buffer. The system calls the port value from the board buffer according to the set sampling frequency and displays it in the front panel.

4.3 Overall System Debugging

Virtual instrument technology is a major breakthrough to traditional instruments, and is a product of the combination of measurement and control

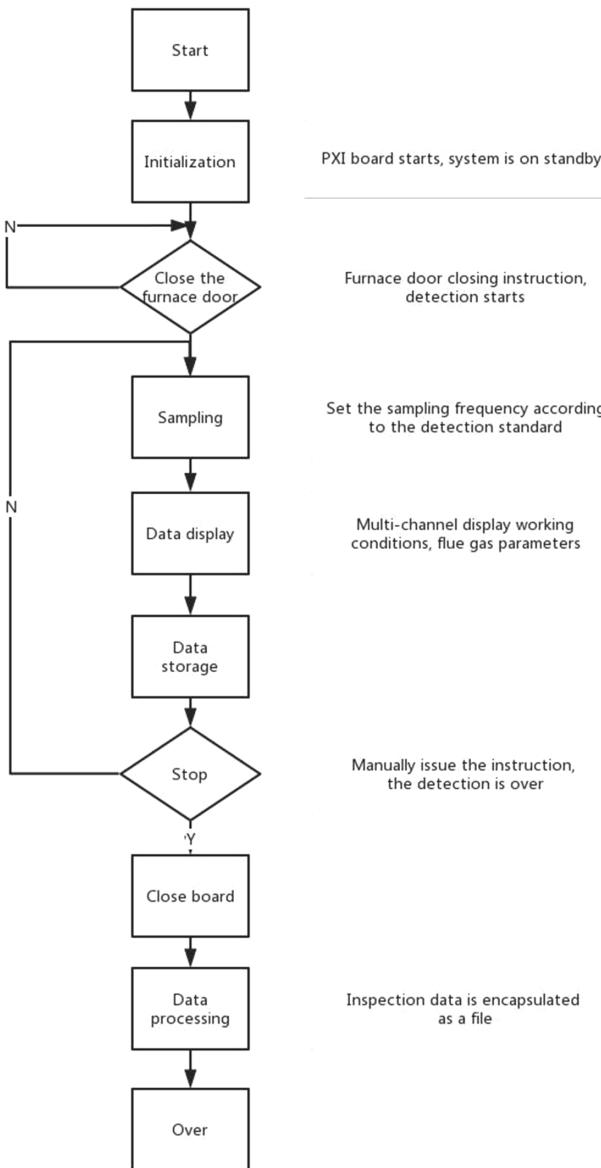


Fig. 2. Sampling module software flow chart

technology and computer technology, which fundamentally updates the concept of instruments [3]. The automatic test system with computer technology as the core is an important means to realize the automatic test of the function and performance of electronic equipment [4]. After the software architecture and functions of this detection system are realized, it is continuously debugged through

simulation and actual operation. The functions of highlight execution, probe, power off, and single-step execution of LabVIEW are used to find errors in the software and finally generate applications program. This system is implemented in accordance with the three parts of data acquisition, report generation and test module during debugging. When debugging the data collection function, the main test is to create a virtual instrument detection channel, check the integration of the digital collection function and software system that comes with the board, read and write data, timing clock, program trigger conditions and task function Etc. Completed the debugging from the creation of the channel to the clock setting, and then to the board sampling. When debugging the report generation function, we mainly design a cluster and add content to the report, focusing on testing the docking of the report with Word, HTML, Excel and other file formats.

References

1. Yuanhui, H., Yeshuang, T.: Automatic test system for a certain type of equipment based on PXI bus. *J. Sichuan Ordnance Eng.* **5**(30), 83–85 (2009)
2. Li, Z.: Design of a radar seeker test control system for a model. Southwest Jiaotong University, Chengdu (2008)
3. Sun, R., Ma, X.: Ship electronic equipment test system based on virtual instrument technology design. *Ship Sci. Technol.* **32**(4), 76–79 (2010)
4. Zhaoqing, L., Liyan, Q.: Research on distributed ATS architecture based on ATML. *Overall Technol. Aerosp.* **3**, 38–45 (2018)



Generating Synthesized CT from Cone-Beam Computed Tomography (CBCT) Using Artifact Disentanglement Network for Image-Guided Radiotherapy (IGRT)

Hanlin Cheng¹, Jiwei Liu^{1(✉)}, Jianfei Liu², Ronghu Mao³, and Pengjian Sun¹

¹ University of Science and Technology Beijing, Beijing 100083, China

liujiwei@ustb.edu.cn

² Anhui University, Anhui 230601, China

³ Henan Cancer Hospital, Henan 450008, China

Abstract. Recently, image generation, including generating synthesized CT (sCT) from CBCT, has become a research hotspot in medical image analysis. Many sCT generation architectures learn the mapping from CBCT images to corresponding CT images based on an unpaired image-to-image translation method, named cycle-consistent generation adversarial network (CycleGAN), but the anatomy accuracy of the sCT images generated by them needs to be improved. To address this problem, we propose a novel CBCT-based sCT generation architecture, which generates sCT images with CBCT anatomy and CT image quality and is trained with unpaired data. The sCT images generated by the proposed architecture have impressively high anatomy accuracy. Experiments show that the sCT images generated by the proposed method are more visually and quantitatively similar to real CT images than some recently-developed CycleGAN-based sCT generation architectures with the mean peak-signal-to-noise ratio (PSNR) improved from 26.15 dB to 29.66 dB.

Keywords: Image generation · Image translation · Unsupervised learning · Unpaired image-to-image translation

1 Introduction

CBCT is one of the imaging technologies currently widely used in IGRT [1,2] and has great clinical potential. However, the large artefacts and inaccurate Hounsfield unit (HU) values of CBCT makes many automated procedures failed to deal with clinical treatment [3–5,14], so it is necessary to improve the image quality of CBCT. Generating synthesized CT (sCT) images with CT image quality while retaining CBCT anatomy has become a popular topic in CBCT image quality improvement.

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 460–469, 2021.

https://doi.org/10.1007/978-981-15-8458-9_49

Many researchers have utilized image-to-image translation to generate sCT images [6–9], their architectures learned the mapping from CBCT images to corresponding CT images using a widely used unpaired image-to-image translation method named as CycleGAN [10]. CycleGAN utilizes the cycle consistency loss as a way of using transitivity to network training and realizes the translation from the source domain to the target domain without establishing a one-to-one mapping between training samples. But CycleGAN has a serious problem when it was used to generate sCT images based on CBCT, i.e., the anatomy accuracy of the sCT images generated by CycleGAN is unsatisfactory.

There are two main reasons for this problem: Firstly, CycleGAN focus on obtaining a good average effect on all datasets and pursue the similarity of major contents ignoring tiny textures. Secondly and the most importantly, CycleGAN maps images in the source image domain and target image domain to separate latent spaces in CBCT-to-CT image translation (Fig. 1a), but there should be a certain relation between these latent spaces.

To solve this problem, we propose a novel CBCT-based sCT generation architecture. It aims at learning a CBCT-to-CT image translation based on our assumptions on CBCT-to-CT translation and the disentangle representation [11, 12]. The network structure of the proposed architecture is based on a metal artifact reduction architecture which named as Artifact Disentanglement Network (ADN) [11]. The proposed architecture generates sCT images with CT image quality and CBCT anatomy without the need of paired training CT and CBCT, and its generated sCT images are similar to real CT images both visually and quantitatively. We compare the proposed architecture with two unsupervised unpaired image-to-image translation methods [7, 10], to prove its superiority in CBCT-based sCT generation.

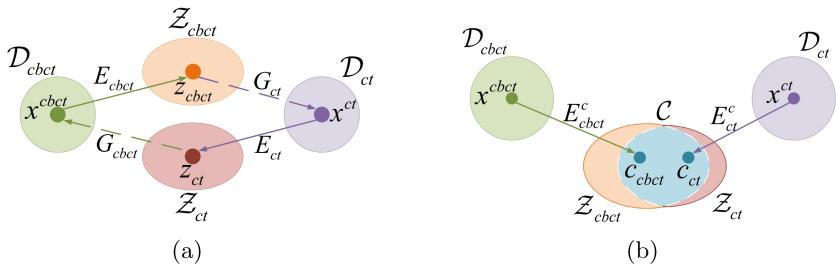


Fig. 1. Comparisons of CycleGAN and our method in CBCT-to-CT translation. Denote \mathcal{D}_{cbct} and \mathcal{D}_{ct} as the image domain of CBCT images and CT images, respectively, \mathcal{Z}_{cbct} and \mathcal{Z}_{ct} as the latent spaces that images in domains \mathcal{D}_{cbct} and \mathcal{D}_{ct} will be mapped to. (a) In CycleGAN, the latent spaces \mathcal{Z}_{cbct} and \mathcal{Z}_{ct} are separated. (b) We assume that spaces \mathcal{Z}_{cbct} and \mathcal{Z}_{ct} share a constant domain content space \mathcal{C} , which captures the anatomy information shared by the two image domains.

2 Methodology

2.1 Principle of Our CBCT-to-CT Translation

Denote \mathcal{D}_{cbct} and \mathcal{D}_{ct} as the image domains of CBCT and CT, respectively, \mathcal{Z}_{cbct} and \mathcal{Z}_{ct} as the latent spaces that the images in domains \mathcal{D}_{cbct} and \mathcal{D}_{ct} will be mapped to, we hold views differ from CycleGAN on the relation between these two latent spaces. Our assumptions about this relation is described as follows:

- i) Based on the observation that the anatomy information contained in domains \mathcal{D}_{cbct} and \mathcal{D}_{ct} should be consistent, we assume that spaces \mathcal{Z}_{cbct} and \mathcal{Z}_{ct} share a constant domain content space \mathcal{C} , which captures the anatomy information contained in the two domains (Fig. 1b).
- ii) Based on the preceding assumption and the fact that CBCT has large artifacts and inaccurate HU values, we assume that the latent spaces of domain \mathcal{D}_{cbct} could be disentangled into a shared content space \mathcal{C} and a specific attribute space \mathcal{A} . This disentanglement can be represented by mapping functions $E_{cbct}^c : \mathcal{D}_{cbct} \rightarrow \mathcal{C}$ and $E_{cbct}^a : \mathcal{D}_{cbct} \rightarrow \mathcal{A}$ (Fig. 2).
- iii) There are two mapping functions $E_{ct} : \mathcal{D}_{ct} \rightarrow \mathcal{C}$ and $G_{ct} : \mathcal{C} \rightarrow \mathcal{D}_{ct}$ which map the CT image y^{ct} from domain \mathcal{D}_{ct} to space \mathcal{C} and back again, i.e., $G_{ct}(E_{ct}(y^{ct})) = G_{ct}(c_y) \approx y^{ct}$, where $y^{ct} \in \mathcal{D}_{ct}$ (Fig. 2).

Denote $f : \mathcal{D}_{cbct} \rightarrow \mathcal{D}_{ct}$ as the function that maps CBCT image x^{cbct} from domain \mathcal{D}_{cbct} to domain \mathcal{D}_{ct} . We could obtain this important mapping function $f(x^{cbct}) = G_{ct}(E_{cbct}^c(x^{cbct})) = \hat{x}^{ct} \approx x^{ct}$ by combining the aforementioned mapping functions $E_{cbct}^c : \mathcal{D}_{cbct} \rightarrow \mathcal{C}$ and $G_{ct} : \mathcal{C} \rightarrow \mathcal{D}_{ct}$, where \hat{x}^{ct} is the synthesized CT (sCT) image generated from the CBCT image x^{cbct} (Fig. 2).

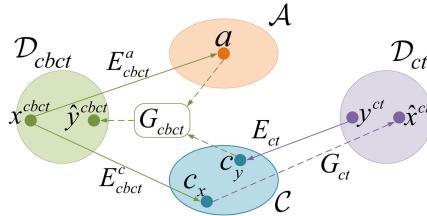


Fig. 2. The principle of our CBCT-to-CT translation. Denote $\{(x^{cbct}, x^{ct}) | x^{cbct} \in \mathcal{D}_{cbct}, x^{ct} \in \mathcal{D}_{ct}, f(x^{cbct}) \approx x^{ct}\}$ as a set of paired images, where $f : \mathcal{D}_{cbct} \rightarrow \mathcal{D}_{ct}$ is the core mapping function which maps image x^{cbct} from domain \mathcal{D}_{cbct} to domain \mathcal{D}_{ct} . We assume that: (i) the latent spaces of \mathcal{D}_{cbct} could be disentangled into a shared content space \mathcal{C} and a specific attribute space \mathcal{A} through mapping functions $E_{cbct}^c : \mathcal{D}_{cbct} \rightarrow \mathcal{C}$ and $E_{cbct}^a : \mathcal{D}_{cbct} \rightarrow \mathcal{A}$, (ii) there are two learnable mapping functions $E_{ct} : \mathcal{D}_{ct} \rightarrow \mathcal{C}$ and $G_{ct} : \mathcal{C} \rightarrow \mathcal{D}_{ct}$, make $G_{ct}(E_{ct}(y^{ct})) = G_{ct}(c_y) \approx y^{ct}$, where $y^{ct} \in \mathcal{D}_{ct}$. By combining the mapping functions $E_{cbct}^c : \mathcal{D}_{cbct} \rightarrow \mathcal{C}$ and $G_{ct} : \mathcal{C} \rightarrow \mathcal{D}_{ct}$, we could obtain the core mapping function $f(x^{cbct}) = G_{ct}(E_{cbct}^c(x^{cbct})) = \hat{x}^{ct} \approx x^{ct}$.

2.2 Overview of the Proposed Architecture

The overview of the proposed CBCT-based sCT generation architecture is shown in Fig. 3, where some components of this architecture are inspired by the recently-developed image translation methods [10–13]. The main network is based on the Artifact Disentanglement Network (ADN) developed by Liao et al to reduce metal artifact in CT [11], the network structures of our image encoders and generators are the same as the corresponding network structures in [11]. The proposed architecture consists of a pair of CT image encoder $E_{ct} : \mathcal{D}_{ct} \rightarrow \mathcal{C}$ and generator $G_{ct} : \mathcal{C} \rightarrow \mathcal{D}_{ct}$, a pair of CBCT image encoder $E_{cbct} = \{E_{cbct}^c : \mathcal{D}_{cbct} \rightarrow \mathcal{C}, E_{cbct}^a : \mathcal{D}_{cbct} \rightarrow \mathcal{A}\}$ and generator $G_{cbct} : \{\mathcal{C}, \mathcal{A}\} \rightarrow \mathcal{D}_{cbct}$, and domain discriminators $\{D_{ct}, D_{cbct}\}$ for both image domains.

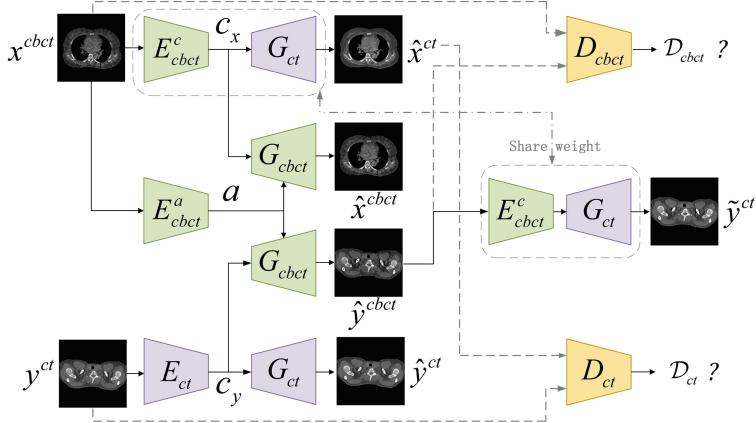


Fig. 3. Overview of the proposed CBCT-based sCT generation architecture. For any image x^{cbct} in domain \mathcal{D}_{cbct} , we could obtain sCT image \hat{x}^{ct} via mapping function $f(x^{cbct}) = \hat{x}^{ct} \approx x^{ct}$, which consists of the encoder $E_{cbct}^c(\cdot)$ and the generator $G_{ct}(\cdot)$, i.e., $f(x^{cbct}) = G_{ct}(E_{cbct}^c(x^{cbct})) = \hat{x}^{ct} \approx x^{ct}$, where $\{(x^{cbct}, x^{ct}) | x^{cbct} \in \mathcal{D}_{cbct}, x^{ct} \in \mathcal{D}_{ct}\}$ represents a set of paired images. All the encoders and generators of our architecture are trainable via unpaired images (x^{cbct}, y^{ct}) , where $x^{cbct} \in \mathcal{D}_{cbct}$ and $y^{ct} \in \mathcal{D}_{ct}$.

Given a set of unpaired images $\{(x^{cbct}, y^{ct}) | x^{cbct} \in \mathcal{D}_{cbct}, y^{ct} \in \mathcal{D}_{ct}\}$, the encoder E_{ct} maps images y^{ct} to content space \mathcal{C} , the encoder E_{cbct}^c maps the content component of images x^{cbct} to content space \mathcal{C} , and the encoder E_{cbct}^a maps the attribute component of x^{cbct} to attribute space \mathcal{A} . The above mapping process could be expressed as:

$$c_y = E_{ct}(y^{ct}), c_x = E_{cbct}^c(x^{cbct}), a = E_{cbct}^a(x^{cbct}) \quad (1)$$

The generator G_{cbct} takes a content code and an attribute code as the input and maps them to domain \mathcal{D}_{cbct} . The decoding from c_x and a reconstructs \hat{x}^{cbct}

and the decoding from c_y and a obtains \hat{y}^{cbct} with attributes of domain \mathcal{D}_{cbct} :

$$\hat{x}^{cbct} = G_{cbct}(c_x, a), \hat{y}^{cbct} = G_{cbct}(c_y, a) \quad (2)$$

The generator G_{ct} takes a content code as the input and maps it to domain \mathcal{D}_{ct} . The decoding from c_y reconstructs \hat{y}^{ct} and the decoding from c_x removes the attribute component of x^{cbct} :

$$\hat{x}^{ct} = G_{ct}(c_x), \hat{y}^{ct} = G_{ct}(c_y) \quad (3)$$

If our disentangle representation is well addressed, \hat{y}^{cbct} could be regarded as a CBCT image in domain \mathcal{D}_{cbct} whose content comes from y^{ct} and attribute comes from x^{cbct} . Therefore, we could recover y^{ct} by reapplying E_{cbct}^c and G_{ct} :

$$\tilde{y}^{ct} = G_{ct}(E_{cbct}^c(\hat{y}^{cbct})) \quad (4)$$

2.3 Loss Function of the Proposed Architecture

It is impossible to minimize the difference between networks outputs and the ground truths using regression losses as the proposed architecture is trained with unpaired data. Therefore, we utilize two discriminators D_{cbct} and D_{ct} to distinguish whether the image is generated by the proposed architecture or sampled from domains \mathcal{D}_{cbct} and \mathcal{D}_{ct} :

$$\begin{aligned} \mathcal{L}_{adv} &= \mathcal{L}_{adv}^{ct} + \mathcal{L}_{adv}^{cbct} \\ &= (\mathbb{E}_{ct} [\log D_{ct}(y^{ct})] + \mathbb{E}_{cbct} [1 - \log D_{ct}(\hat{x}^{ct})]) \\ &\quad + (\mathbb{E}_{cbct} [\log D_{cbct}(x^{cbct})] + \mathbb{E}_{ct, cbct} [1 - \log D_{cbct}(\hat{y}^{cbct})]) \end{aligned} \quad (5)$$

In order to make $\{(E_{cbct}^c, E_{cbct}^a), G_{cbct}\}$ and $\{E_{ct}, G_{ct}\}$ act as autoencoders when performing encoding and decoding with the same image, we introduced reconstruction loss to the proposed architecture:

$$\mathcal{L}_{rec} = \mathbb{E}_{ct, cbct} [\|\hat{x}^{cbct} - x^{cbct}\|_1 + \|\hat{y}^{ct} - y^{ct}\|_1] \quad (6)$$

where \hat{x}^{cbct} (Eq. 2) and \hat{y}^{ct} (Eq. 3) are reconstructed from inputs x^{cbct} and y^{ct} , respectively, and we choose L1 loss to obtain sharper outputs like [11].

The adversarial loss mentioned earlier (Eq. 5) encourages the sCT image \hat{x}^{ct} to resemble a image sampled from domain \mathcal{D}_{ct} , but it is difficult to ensure the anatomy accuracy of \hat{x}^{ct} only using this loss. Therefore, we introduce an attribute consistency loss to improve the anatomy accuracy of

$$\mathcal{L}_{att} = \mathbb{E}_{ct, cbct} [\|(x^{cbct} - \hat{x}^{ct}) - (\hat{y}^{cbct} - y^{ct})\|_1] \quad (7)$$

Then, we can pair \tilde{y}^{ct} with y^{ct} to further verify the effectiveness of the disentangle representation and regularize the disentanglement in Eq. 4 with regression:

$$\mathcal{L}_{self} = \mathbb{E}_{ct, cbct} [\|\tilde{y}^{ct} - y^{ct}\|_1] \quad (8)$$

The adversarial loss and attribute consistency loss ensure the anatomical consistency of \hat{x}^{ct} and x^{cbct} , but we expect to further improve the anatomical accuracy of sCT images \hat{x}^{ct} , so we design a rough shape constraint on the anatomy of \hat{x}^{ct} . Morphological operations are performed to simply segment the different anatomical regions and strengthen the weights of the contours of different regions:

$$\mathcal{L}_{shape} = \mathbb{E}_{ct, cbct} [\|shape_{\hat{x}^{ct}} - shape_{x^{cbct}}\|_1] \quad (9)$$

The overall loss function of the proposed architecture can be described as follows:

$$\mathcal{L} = \lambda_{adv} (\mathcal{L}_{adv}^{ct} + \mathcal{L}_{adv}^{cbct}) + \lambda_{rec} \mathcal{L}_{rec} + \lambda_{att} \mathcal{L}_{att} + \lambda_{self} \mathcal{L}_{self} + \lambda_{shape} \mathcal{L}_{shape} \quad (10)$$

where the λ are hyper-parameters that control the importance of each loss function.

3 Experiments

3.1 Data Acquisition and Preprocessing

The entire dataset is randomly selected from the clinical database, including wholebody planning CT (pCT) and CBCT scans from 52 patients scanned in different days. The parameters of pCT scans are 120 kVp, an in-plane resolution of range 1.01 mm 1.33 mm and a slice thickness of 3.00 mm. The corresponding parameters of CBCT scans are 110 kVp, 0.51 mm 0.91 mm 1.98 mm to 2.00 mm, respectively. Among those patients, 32 is used as training set, 8 as validation set and 12 as testing set. We resample all CBCT and pCT scans to a 1.0 mm 1.0 mm 1.0 mm grid to achieve consistent spatial resolution and align all pCT scans to CBCT scans using AIRLab [15]. In the testing phase, we perform non-rigid pCT-to-CBCT registration using AIRLab to obtain deformed planning CT (dpCT) slices as ground truth. Moreover, we clip the HU values of all the slices to $[-1000, 2000]$ and normalize these slices to the $[-1, 1]$.

3.2 Baselines

In our experiment, two other unpaired image-to-image translation methods, which are described in the following as CycleGAN [10] and CycleGAN-Unet512 [7] were used for comparison under the same datasets. We utilized dpCT images as the ground truth to evaluate the HU value accuracy of the sCT images generated by CycleGAN, CycleGAN-Unet512 and the proposed architecture, respectively. Our evaluation was conducted with similarity measures between sCT images generated by diffent image translation methods and dpCT images, including mean absolute error (MAE), root-mean-square error (RMSE), peak-signal-to-noise ratio (PSNR) and mean structural similarity index (SSIM).

3.3 Training and Testing

We implement the proposed architecture under the PyTorch deep learning framework and use the Adam optimizer with 1×10^{-4} . The hyperparameter values from Eq. 10 were set as follows: $\lambda_{adv} = 1.0$, $\lambda_{rec} = \lambda_{att} = \lambda_{self} = 5.0$ and $\lambda_{shape} = 20.0$. The network model is trained, validated and tested on an NVIDIA GTX 1080 GPU with 8 GB of memory. The batch size was set to 2 during network training and was set to 8 during network validation and testing.

4 Results and Discussion

Figure 4 shows sCT images generated by CycleGAN [10], CycleGAN-Unet512 [7] and the proposed architecture. As shown in Fig. 4, the sCT images generated by CycleGAN and CycleGAN-Unet512 are relatively smooth and close to the smooth HU values distribution of real CT images, but some of the sCT images have unrealistic anatomy. In contrast, the proposed architecture can produce sharper sCT images with higher anatomical accuracy.

Table 1. Mean and standard deviation values of quantitative results. The dpCT images were taken as the ground truth images for similarity measures.

	PSNR (dB)	SSIM	MAE (HU)	RMSE (HU)
CBCT	26.15 ± 5.10	0.77 ± 0.08	70.49 ± 31.52	167.16 ± 63.18
sCT (CycleGAN [10])	28.37 ± 6.55	0.83 ± 0.09	61.82 ± 46.37	143.26 ± 87.40
sCT (CycleGAN-Unet512 [7])	28.88 ± 6.01	0.84 ± 0.08	57.77 ± 42.40	133.96 ± 81.67
sCT (Our method)	29.66 ± 6.73	0.85 ± 0.08	58.71 ± 41.68	132.75 ± 81.67

Table 1 summarizes the quantitative results for all the testing patients. As shown in Table 1, our proposed method significantly outperformed CycleGAN and CycleGAN-Unet512 by PSNR, SSIM, RMSE. Although our method only achieves the second-best performance on the MAE metric, the standard deviation (41.68 HU) of our MAE results is better than the standard deviation (42.40 HU) of the former. Besides, the results of RMSE are more sensitive to outliers than the results of MAE, which means that the proposed architecture shows better stability than two other methods, which is very meaningful for the clinical treatment of IGRT.

So far, the proposed architecture shows the best performance among the unpaired image-to-image translation methods in our experiments, and the image quality of the sCT images generated by this method is close to real CT images. The results of our experiments validate our assumptions about the relation between the latent spaces that images in the CBCT image domain and CT image domain will be map to in the CBCT-to-CT image translation.

However, the dpCT images were used as ground truth to evaluate HU accuracy of sCT images in our experiments, so that the test results may be inaccurate

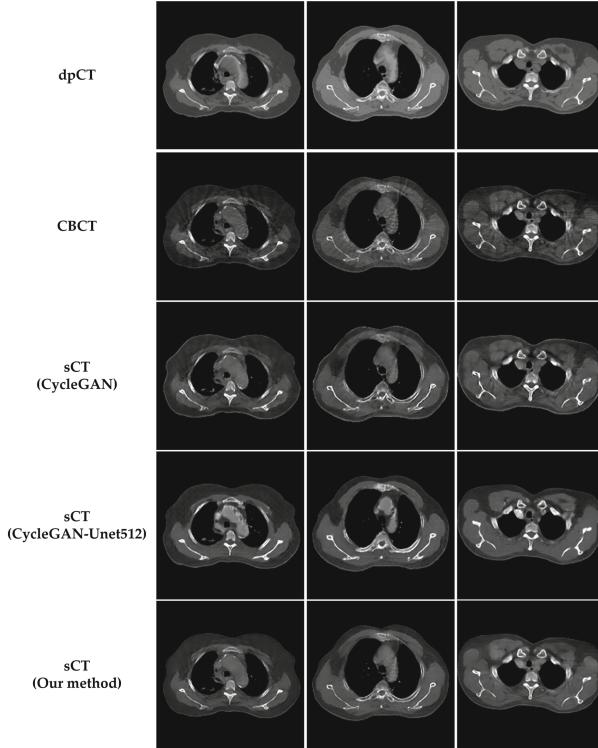


Fig. 4. CBCT, sCT (CycleGAN), sCT (CycleGAN-Unet512), sCT (ADN) and sCT (the proposed architecture) and dpCT images in axial plane from the same part of one patient. Display window is $[-250, 550]$ HU.

because non-rigid registration may change the original HU value distribution of aligned CT. Besides, we assume that the encoded content component of a CT image contains most of the key information of it, and we will only lose negligible information while reconstructing this CT image through its encoded content component. But the information we ignore may contain important attribute information of CT image domain, so the proposed architecture is only an approximate reconstruction. In our future work, we will re-examine the relation between the latent spaces that CBCT and CT images will map to during CBCT-to-CT image translation.

5 Conclusions

In this study, we proposed a novel CBCT-based sCT generation architecture for CBCT correction. The proposed architecture relies on learning a CBCT-to-CT image translation and generated sCT images without paired CT and CBCT for training. The sCT images generated by the proposed architecture are similar

to real CT images with high anatomical accuracy. The proposed architecture can notably improve the mean PSNR from 26.15 dB to 29.66 dB. Although the proposed architecture cannot completely mitigate artifacts and inaccurate HU values in CBCT, it can greatly improve them and the image quality of the sCT generated by the proposed method is much closer to the real CT than CBCT.

The high anatomical accuracy of the proposed architecture improves the practicality of CBCT in monitoring patient changes throughout treatment and is of great significance for the clinical treatment of IGRT. Therefore, the proposed architecture which generates sCT images from CBCT images has a great clinical potential.

References

1. Hvid, C.A., Elstrøm, U.V., Jensen, K., Grau, C.: Cone-beam computed tomography (CBCT) for adaptive image guided head and neck radiation therapy. *Acta Oncologica* **57**, 552–556 (2018). <https://doi.org/10.1080/0284186X.2017.1398414>
2. Raman, N.Y., Kim, S., Deek, M.P., Li, D., Gupta, A., Bond, L., Dwivedi, A., Braver, J.K., Reyhan, M., Mittal, A., Gui, B., Malhotra, J., Aisner, J., Jabbour, S.K.: Daily image-guidance with cone beam computed tomography may reduce radiation pneumonitis in unresectable non-small cell lung cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **101**, 1104–1112 (2018). <https://doi.org/10.1016/j.ijrobp.2018.03.025>
3. Vijayan, R., Silva, T.D., Han, R., Zhang, X., Uneri, A., Doerr, S., Ketcha, M., Pantoja, A.P., Theodore, N., Siewerdsen, J.H.: Automatic pedicle screw planning using atlas-based registration of anatomy and reference trajectories. *Phys. Med. Biol.* **64**, 165020 (2019). <https://doi.org/10.1088/1361-6560/ab2d66>
4. Zachiu, C., De Senneville, B.D., Tijssen, R.H.N., Kotte, A.N.T.J., Houweling, A.C., Kerkmeijer, L.G.W., Lagendijk, J.J.W., Moonen, C.T.W., Ries, M.: Non-rigid CT/CBCT to CBCT registration for online external beam radiotherapy guidance. *Phys. Med. Biol.* **63**, 015027 (2017). <https://doi.org/10.1088/1361-6560/aa990e>
5. Cole, A.J., Veiga, C., Johnson, U., D’Souza, D., Lalli, N.K., McClelland, J.R.: Toward adaptive radiotherapy for lung patients: feasibility study on deforming planning CT to CBCT to assess the impact of anatomical changes on dosimetry. *Phys. Med. Biol.* **63**, 15501 (2018). <https://doi.org/10.1088/1361-6560/aada96>
6. Harms, J., Lei, Y., Wang, T., Zhang, R., Zhou, J., Tang, X., Curran, W.J., Liu, T., Yang, X.: Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. *Med. Phys.* **46**, 3998–4009 (2019). <https://doi.org/10.1002/mp.13656>
7. Liang, X., Chen, L., Nguyen, D., Zhou, Z., Gu, X., Yang, M., Wang, J., Jiang, S.: Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Phys. Med. Biol.* **64**, 125002 (2019). <https://doi.org/10.1088/1361-6560/ab22f9>
8. Kurz, C., Maspero, M., Savenije, M.H.F., Landry, G., Kamp, F., Pinto, M., Li, M., Parodi, K., Belka, C., Van Den Berg, C.A.T.: CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation. *Phys. Med. Biol.* **64**, 225004 (2019). <https://doi.org/10.1088/1361-6560/ab4d8c>

9. Lei, Y., Wang, T., Harms, J., Erfani, G.S., Dong, X., Zhou, J., Patel, P., Tang, X., Liu, T., Curran, W.J., Higgins, K., Yang, X.: Image quality improvement in cone-beam CT using deep learning. In: Medical Imaging 2019: Physics of Medical Imaging, pp. 556–561, SPIE, California (2019). <https://doi.org/10.1117/12.2512545>
10. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: The IEEE International Conference on Computer Vision (ICCV), pp. 2223–2232 (2017). <https://doi.org/10.1109/iccv.2017.244>
11. Liao, H., Lin, W.A., Zhou, S.K., Luo, J.: Artifact disentanglement network for unsupervised metal artifact reduction. *IEEE Trans. Med. Imaging* **39**, 634–643 (2019). <https://doi.org/10.1109/TMI.2019.2933425>
12. Lee, H.Y., Tseng, H.Y., Huang, J.B., Singh, M., Yang, M.H.: Diverse image-to-image translation via disentangled representations. In: The European Conference on Computer Vision (ECCV), pp. 35–51 (2018)
13. Lin, T.Y., Dollar, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2117–2125 (2017). <https://doi.org/10.1109/cvpr.2017.106>
14. Park, S., Plisker, W., Quon, H., Wong, J., Shekhar, R., Lee, J.: Deformable registration of CT and cone-beam CT with local intensity matching. *Phys. Med. Biol.* **62**, 927–947 (2017). <https://doi.org/10.1088/1361-6560/aa4f6d>
15. Sandkuhler, R., Jud, C., Andermatt, S., Cattin, P.C.: AirLab: autograd image registration laboratory (2018). [arXiv:1806.09907](https://arxiv.org/abs/1806.09907)



Design and Application of Wireless Temperature Monitoring System for Diesel Locomotive in Reconditioning Field Based on ZigBee Network

Wenlong Yao^{1,2}, Zhen Pang¹, Keyu Zhuang¹, and Wei Shao^{1(✉)}

¹ School of Automation and Electrical Engineering,
Qingdao University of Science and Technology, Qingdao 266100, China
yaowenlong@qust.edu.cn
² School of Aerospace Engineering,
Beijing Institute of Technology, Beijing 100000, China

Abstract. This paper focuses on the temperature monitoring and management of diesel locomotive in railway reconditioning field. The wireless sensor network monitoring system is utilized to replace the original manual operation. The algorithm of falling error is designed to realize high precision real-time monitoring and management. Firstly, ZigBee wireless communication is adopted to monitor temperature. Realize the remote monitoring of data unattended. Secondly, a portable acquisition terminal with low power consumption was design. Then, the wireless data receiving terminal is connected to the TCP/IP converter. The received data is connected to local area network (LAN), the monitoring and warning upper computer is designed. And the falling error algorithm is added to realize temperature compensation, real-time display, monitoring and warning and data storage. At the same time, it can realize the mobile terminal data viewing and monitoring and early warning. The design of the system has been preliminarily tested in Qingdao locomotive depot and further verified the rationality and feasibility of the scheme.

Keywords: Railway transportation · Monitoring system · Wireless sensor network · Data acquisition · ZigBee wireless communication · Low-power dissipation

1 Introduction

With the increasing social demand and requirements, the development of modern railways has put forward higher requirements for railway transportation and management. It is an urgent task to replace the original production mode with intelligent equipment, reduce energy consumption and gradually realize industrial transformation [1, 2]. At present, various basic research and applied technology research on energy conservation and emission reduction are being actively carried out. This pushed rail-related technology to a new level [3, 4].

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 470–478, 2021.

https://doi.org/10.1007/978-981-15-8458-9_50

At present, domestic railway locomotives are mainly pulled by diesel locomotives. It is necessary to conduct anti-freezing treatment in the locomotive depot. At present, the main method used is manual “full volume, timed temperature beating”. Under the influence of ambient temperature, the distribution of heating oil is unreasonable. These result in low heating efficiency and wasted fuel. In addition, there are huge potential safety hazards and personnel waste when checking the temperature of the car manually in winter [5].

Wireless sensor network technology has crossed many fields of modern life [6,7]. ZigBee wireless network data acquisition system has advantages in node deployment and sensor routing. At present, it has been well applied in environmental information collection, water quality testing, family security monitoring and other scenarios [8,9]. Therefore, the problems such as wide distribution range, dispersive distribution and difficult wiring are considered. This paper designs a wireless temperature detection system for diesel locomotive in reconditioning field based ZigBee wireless transmission mode. It can accurately monitor the locomotive's reconditioning condition on the upper computer, optimize the management and reduce the reconditioning energy consumption.

2 Fine Whole Systems Design

ZigBee wireless transmission technology is selected for stable data transmission and low power. The wireless transmission mode has the freedom to network and is not easily disturbed by the high-voltage transmission lines in the reconditioning field. The system utilizes DS18B20 digital temperature sensor to measure temperature. Based on the unchanged requirements, the design adopts magnetic absorption sensor and portable design. And the design utilizes low-power hardware and independent power supply. Therefore, it is not necessary to use the power supply of the original car, and the temperature measurement system of the original car does not interfere with each other.

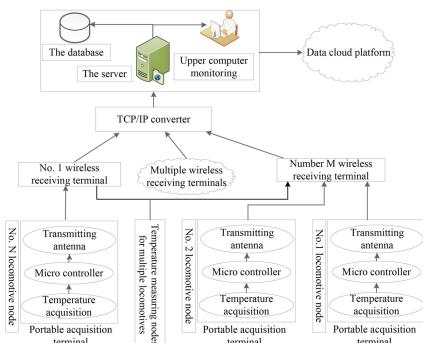


Fig. 1. Overall system architecture

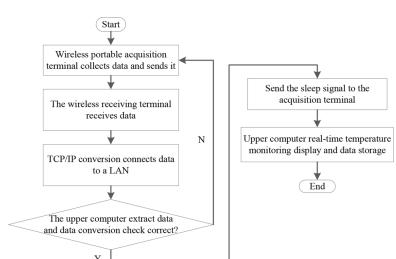


Fig. 2. Data transmission process

The wireless temperature monitoring system is mainly composed of wireless temperature portable acquisition terminal, wireless data receiving terminal, TCP/IP converter, data cloud platform and monitoring and warning host unit. Supported by ZigBee wireless communication technology, the Internet of things system is formed in railway locomotive reconditioning. The overall structure is shown in Fig. 1.

In the above wireless monitoring system, the wireless portable temperature acquisition terminal attaches to the tube wall of the diesel locomotive for temperature acquisition. Then through ZigBee wireless network communication, the measured temperature signal will be sent to the wireless receiving terminal; Multiple wireless receiving terminals in the system self-organize to receive temperature signals according to the signal strength. Due to the difficulty of wiring in the field and the limited transmission distance of ZigBee, the wireless receiving terminal is combined with TCP/IP converter. It utilizes the existing network cable in the reconditioning field to realize long-distance signal transmission; The data cloud platform is connected with the monitoring and warning computer. The field server receives data, realizing temperature drop error, real-time display, monitoring, warning and data storage. At the same time, it will upload the data to the data cloud platform to realize the mobile terminal data view, monitoring and warning. The data transmission process of the system is shown in Fig. 2.

3 Hardware Design

Wireless portable temperature acquisition terminal is an important part of temperature monitoring system. It mainly includes sensor module, control module, wireless transmission module and power module. The device integrates data collection, processing and transmission into a single housing. The whole equipment adopts portable design, which can be used at any time. Since there is a magnet on the outside of the shell, the terminal equipment can be adsorbed on the locomotive during data collection and removed after the completion of reconditioning, without interfering with the original vehicle.

3.1 Sensor Module Design

The sensor module adopts the magnetic absorption digital temperature sensor DS18B20 with insulation layer. It adsorbs directly to the outside of the measuring tube for temperature measurement. At the same time, digital-to-analog conversion is avoided in the hardware circuit. As shown in Fig. 3, the sensor is designed with low power consumption and independent power supply.

3.2 Control Module Design

The core control module of the designed wireless portable acquisition terminal is STM8 chip, which controls the hardware circuit. It mainly includes reset circuit, clock circuit, signal indicator control, frequency adjustment control, temperature sensor control, charging circuit and ZigBee wireless transmission control.

The system utilizes independent power supply. STM8 is equipped with PC oscillation circuit. The system adds crystal oscillator circuit to the peripheral circuit to improve the control accuracy. The design is shown in Fig. 4. In addition, a voltage signal acquisition circuit is designed, as shown in Fig. 5. The control module can collect voltage signal and send it to realize low voltage alarm.

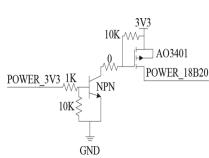


Fig. 3. Control circuit of DS18B20 sensor

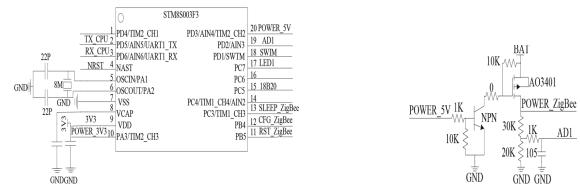


Fig. 4. Design of STM8 peripheral circuit

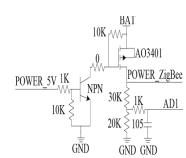


Fig. 5. Voltage signal acquisition circuit

3.3 Design of Wireless Transmission Module

The wireless transmission module adopts ZigBee wireless transmission mode. And sz05-adv (embedded wireless module) is used for data transmission. The circuit design is shown in Fig. 6. The module is designed with low power consumption. It sends the temperature signal collected by the portable terminal, the address number of the device and the voltage signal of the device to the wireless receiving terminal according to the preset protocol. At the same time, it receives the dormancy signal sent by the upper computer and controls the temperature acquisition terminal to enter dormancy.

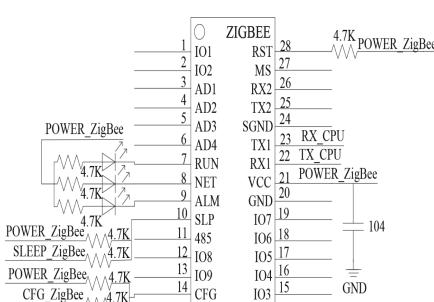


Fig. 6. Circuit design of ZigBee wireless transmission module

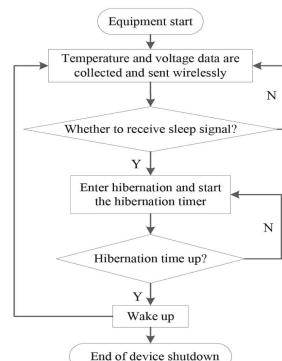


Fig. 7. The workflow of wireless portable acquisition terminal

4 Software Development

4.1 Lower Computer Software Development

In this system, each acquisition terminal has a fixed address to facilitate data management. The STM8 microcontroller works in a timed sleep mode and wakes up every 3 min to measure the temperature signal and send it. The device is equipped with a frequency adjustment button, which can be pressed long enough to increase the signal transmission frequency from once in 3 min to once in 1 s for equipment testing; When the dormancy signal is not received, the temperature signal will be collected again and the transmitting frequency will be increased. The specific workflow is shown in Fig. 7.

4.2 Upper Computer Software Development

Workflow Design. The complete upper computer workflow is shown in Fig. 8. Among them, permission management is added in user login. And different permissions are set for ordinary operators and managers. Management personnel can adjust the locomotive number, alarm temperature and terminal serial port in the upper computer. In addition, the data is stored in the database and uploaded to the data cloud platform. It is convenient to further design mobile terminal APP for remote monitoring. And it improves the flexibility of the system.

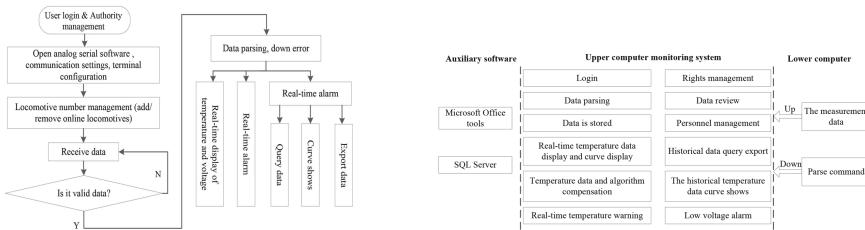
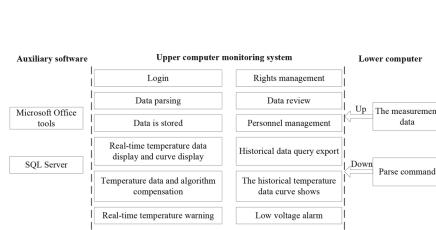


Fig. 8. Work flow chart of upper computer

Fig. 9. Software functional structure design



Software Function Design. The system utilizes a C# language which running on the .net Framework and .net Core released by Microsoft for software development. Based on ZigBee wireless communication mode and the requirements of temperature monitoring, this system designs the monitoring and display interface. The functional structure is shown in Fig. 9.

Data Reduction Error Design. During the working process of the locomotive temperature monitoring system, there are errors in the temperature measurement due to the influence of the external environment temperature, the position of the tube wall measured by the sensor and the difference of the sensor hardware itself. Adaptive weighting algorithm and BP neural network data fusion

are added to the upper computer to reduce measurement error and improve data accuracy. Consider using multiple sensors to monitor and collect multiple monitoring values of the same locomotive tube wall at the same time.

Suppose: the measured value of each sensor is T_1, T_2, \dots, T_n , the mean is \bar{T} , the variance is $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, the total mean squared error is σ^2 , the weighting factor is M_1, M_2, \dots, M_n , the truth value to estimate is T . Where, \bar{T}, M and σ^2 after fusion meet:

$$\begin{aligned}\bar{T} &= \sum_{p=1}^n M_p T_p \\ \sum_{p=1}^n M_p &= 1\end{aligned}$$

$$\sigma^2 = E[(T - \hat{T})] = E \left[\sum_{p=1}^n M_p (T - T_p)^2 + 2 \sum_{p=1}^n \sum_{q=1}^n M_p M_q (T - T_p)(T - T_q) \right]$$

Since T_1, T_2, \dots, T_n is the detection temperature of different sensors, which is independent of each other. And it is the unbiased estimation of T . Therefore, in the above formula, $E[(T - T_p)(T - T_q)] = 0$, namely $E[2 \sum_{p=1}^n \sum_{q=1}^n M_p M_q \times (T - T_p)(T - T_q)] = 0$, formula (3) can be expressed as:

$$\sigma^2 = E \left[\sum_{p=1}^n M_p (T - T_p)^2 \right] = \sum_{p=1}^n M_p^2 \sigma_p^2$$

As can be seen from the above formula, the smaller σ^2 is, the higher the output precision is. Therefore, when demand σ^2 is minimized, the weighting factor M_p is:

$$M_p^* = (\sigma_p^2 \sum_{p=1}^n \frac{1}{\sigma_p^2})^{-1}$$

At this time, σ^2 gets the minimum value of $\sigma_{\min}^2 = (\sum_{p=1}^n \frac{1}{\sigma_p^2})^{-1}$. At the same time, the historical data received by the upper computer is used for estimation. Let the p sensor take the i measurement, and the average value obtained is:

$$\bar{T}_p(i) = \frac{1}{i} \sum_{m=1}^i T_p(m)$$

Then the final estimate of temperature is obtained:

$$\hat{T}_p = \sum_{p=1}^n M_p \bar{T}_p(i)$$

The algorithm utilizes historical data to eliminate the inaccuracy and limitations of the measurement results of a single sensor. Finally, the consistency of the sensor data is obtained to improve the monitoring accuracy.

5 Software Development

The system was tested in the reconditioning yard of Qingdao locomotive depot. The temperature sensor is adsorbed on the temperature measuring tube wall of the temperature hitting locomotive. And the wireless receiving device is installed on the office balcony near the reconditioning field. Using serial port to TCP/IP module, the receiving device is connected to LAN. The upper computer monitoring software is installed on any PC terminal within the LAN.

After the successful installation and debugging of the system, the person on duty only needs to log in the account indoors. After entering the monitoring interface and establishing the communication of the monitoring locomotive, data can be monitored in real time. As shown in the Fig. 10.

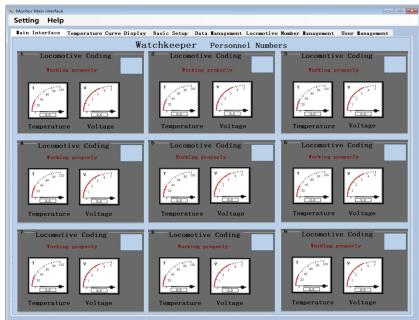


Fig. 10. Monitor interface of upper computer

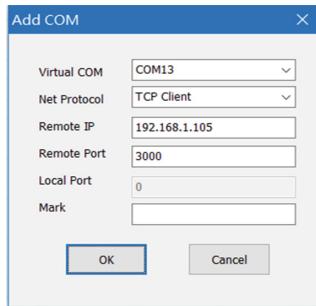


Fig. 11. Create a virtual serial port

Through the test, the application effect of the system, the accuracy and stability of the data received, the alarm situation and whether it can realize the purpose of tempering on demand and saving heating oil are tested. Test methods and effect analysis are as follows:

(1) Testing of wireless data transmission penetration, communication distance, scheme feasibility and reliability

In the railway locomotive reconditioning yard, all the Windows are closed when the locomotive is normally reconditioning. This leads to a decrease in signal transmission intensity. The test was conducted on several locomotives far from the receiving equipment in the reconditioning field. During the test, the wireless temperature measuring terminal is installed in the locomotive. And the locomotive starts. The received data is connected to the LAN through the serial

port to Ethernet module. Each acquisition terminal has a fixed address. After configuration of the serial port server at the PC monitoring end, a virtual serial port is created, as shown in Fig. 11. The upper computer receives data packets and detects the transmission of wireless signals. ZigBee wireless signals can be successfully transmitted and multiple gateways work together. The TCP/IP module accesses the data to the LAN. This enables centralized PC monitoring and communication over unlimited distances.

(2) Test for accuracy and stability of transmitted data

During the test, the locomotive was started into the normal reconditioning state. Five temperature measuring devices with different Numbers are placed on the same monitoring node. The device monitors and sends data every 3 seconds. The actual returned data is recorded on the PC side. At the same time, the temperature data after the system drop error is calculated. Compare and analyze the data, as shown in Fig. 12. Moreover, the relative error analysis is performed, as shown in Fig. 13.

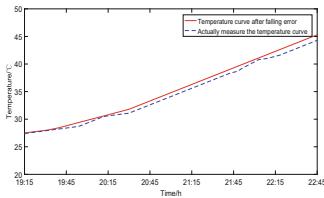


Fig. 12. Comparison and analysis of received data curves

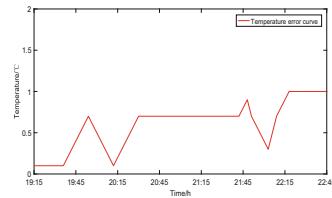


Fig. 13. Error analysis of received data

As can be seen from Fig. 12, the temperature data received is stable within 3 hours. The temperature data after falling error is a little larger than the actual data received. The temperature rise is stable and the measurement accuracy is submitted; According to the error results of falling error data and actual received data in Fig. 13, data transmission is stable. The temperature error of system monitoring is maintained within 1nd. And the temperature measurement is accurate.

6 Actual Measurement and Effect Analysis

The ZigBee network-based reconditioning field wireless temperature monitoring system of diesel locomotive designed in this paper provides a new way of reconditioning monitoring. The system replaces the original manual operation mode. It can accurately monitor the locomotive's reconditioning condition on the upper computer and reasonably control the internal combustion engine. It solves the problems of manpower waste, fuel waste and difficult management in

the reconditioning of the original railway locomotive. The wiring of the system is easy to install and maintain, which improves the flexibility of the system and reduces the cost of operation and maintenance. It is also of great significance to the transformation of industry and the transformation from old to new drivers of growth.

Acknowledgement. This work was supported by Natural Science Foundation of Shandong Province (ZR2017MEE071), National Natural Science Foundation of China (61773227), Shortage Talents Program in Key Supported Regions of Shandong Province (Major office of Shandong Provincial Development and Reform Commission [2019] No. 391).

References

1. Li, X.: Quality management system study on celerity railway engineering. In: 2009 Second International Symposium on Knowledge Acquisition and Modeling, Wuhan, pp. 391–394 (2009). <https://doi.org/10.1109/kam.2009.196>
2. Xukuo, G., Qiong, W.: Research on the mode of present transportation in China and the analysis of railway transportation. In: 2013 6th International Conference on Information Management Innovation Management and Industrial Engineering, Xi'an, pp. 414–416 (2013). <https://doi.org/10.1109/iciii.2013.6703606>
3. Bhuvaneswari, N., Ramesh, L., Ramakrishnan, R.: Indian railways carbon emission reduction and energy performance. In: 2016 International Conference on Control, Instrumentation, Communication and Computational Technologies, Kumarakom, pp. 746–749 (2016). <https://doi.org/10.1109/icccct.2016.7988051>
4. Maghfiroh, H., Hermanu, C., Nizam, M.: Sensorless control of railway traction motor in the energy point of view. In: 2019 6th International Conference on Electric Vehicular Technology, Bali, pp. 267–271 (2019)
5. Komyakov, A.A., Ponomarev, A.V., Erbes, V.V.: Application of an intelligent system for the monitoring of energy efficiency indicators at service locomotive depots. In: 2018 International Multi-Conference on Industrial Engineering and Modern Technologies, Vladivostok, pp. 1–7 (2018). <https://doi.org/10.1109/fareastcon.2018.8602926>
6. Nishikawa, Y., Sasamura, T., Ishizuka, Y., Sugimoto, S., Iwasaki, S., Wang, H., Kurihara, K.: Design of stable wireless sensor network for slope monitoring. In: 2018 IEEE Topical Conference on Wireless Sensors and Sensor Networks, Anaheim, pp. 8–11 (2018). <https://doi.org/10.1109/wisnet.2018.8311550>
7. De la Concepcion, A.R., Stefanelli, R., Trinchero, D.: Adaptive wireless sensor networks for high-definition monitoring in sustainable agriculture. In: 2014 IEEE Topical Conference on Wireless Sensors and Sensor Networks, Newport Beach, pp. 67–69 (2014). <https://doi.org/10.1109/wisnet.2014.6825511>
8. Suryawanshi, V., Khandekar, P.M.: Design and development of Wireless Sensor Network (WSN) for water quality monitoring using Zigbee. In: 2018 Second International Conference on Intelligent Computing and Control Systems, Madurai, pp. 862–865 (2018). <https://doi.org/10.1109/iccons.2018.8663131>
9. Wei-Dong, H., Bo-Xuan, Z.: Smart home wireless system using ZigBee and IEEE802.15.4. In: 2016 Sixth International Conference on Instrumentation & Measurement, Computer, Communication and Control, Harbin, pp. 858–863 (2016). <https://doi.org/10.1109/imccc.2016.168>



IMC-PID Controller Based on Extended State Observer for Second-Order Delayed Unstable Processes with Two Unstable Poles

Yachao Liu, Jian Gao^(✉), and Lanyu Zhang

State Key Laboratory of Precision Electronic Manufacturing Technology and Equipment, Guangdong University of Technology, Guangzhou 510006, China
lyc2018@mail2.gdut.edu.cn

Abstract. The internal model control (IMC) based PID controller for delayed processes often needs to compromise the set-point tracking response and disturbance response. For the complex second-order delayed unstable process with two unstable poles, there are no simple and effective PID controller tuning rules. In this paper, an IMC-PID control method based on extended state observer (ESO) is proposed. Firstly, in order to reduce the estimation burden of traditional ESO and improve the accuracy of disturbance estimation, a synchronous model ESO (SMESO) is designed. The SMESO considers the information of the nominal model, synchronizes the control input and the system response output in time, thus transforming the delayed process into a perfect model form. Then, through the first-order Taylor expansion and pole-zero cancellation, the parameter tuning rules of IMC-PID controller for the second-order delayed unstable process are derived. Finally, the effectiveness of the proposed method is illustrated by simulation examples.

Keywords: Extended state observer · Internal model control · Delayed unstable process

1 Introduction

The industrial processes with time delay are very difficult to control. Especially for the second-order delayed unstable process with two unstable poles, the closed-loop response of the classical PID control usually has large overshoot and long setting time, and the load disturbance rejection performance is not satisfactory. The parameter tuning of PID controller determines the control performance of the closed-loop system, it is urgent to find simple and effective parameter tuning rules for time-delay processes.

The PID parameters tuning based on internal model control (IMC) has become the most practical tuning method after its appearance. The originally IMC principle based PID parameter tuning method was only suitable for delayed stable processes [1]. The IMC-PID parameter tuning method proposed by [2,3]

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 479–486, 2021.

https://doi.org/10.1007/978-981-15-8458-9_51

can be used for integral and unstable processes with time delay. Skogestad [4] proposed a SIMC method for model reduction, which was successfully applied in industry by Haugen [5]. Ghousiha et al. [6] instigated an enhanced IMC-PID parameter tuning rule, which was used for non-minimum phase integral delayed processes. For classical delayed processes, a set of relatively complete PID controller tuning rules based on IMC is derived [7]. However, the classical second-order delayed unstable process with two unstable poles has not been discussed. The disadvantage is that the above tuning rules for PID controller often require a trade-off between set-point tracking and load disturbance rejection.

In recent years, a new active disturbance rejection control (ADRC) technique has been applied to the delayed systems [8–10]. The obvious difference between ADRC and other control methods is that this technology proposes the concept of total disturbance, which unifies the external disturbance and internal dynamics. The core part extended state observer (ESO) of ADRC is used to estimate the total disturbance accurately and compensate it in the control law, that can effectively improve the disturbance rejection ability of the closed-loop system. Zheng et al. [11] proposed a predictive ADRC method for stable processes with time delay. Zhao et al. [12] proposed a modified ADRC method, which can be used for stable and unstable processes with time delay. However, this method did not use the known nominal model information in modified ESO. In addition, Zhang et al. [13] also proposed a novel feed-forward compensation ADRC approach for a class of input time delay systems. However, the control parameters of the feed-forward compensation ADRC are difficult to be tuning.

Therefore, to improve the closed-loop response performance of both set-point tracking and load disturbance rejection, while also providing simple and feasible parameter tuning rules for the controller, this paper proposes an IMC-PID controller based on synchronous model ESO (SMESO) for the second-order delayed unstable process with two unstable poles. The SMESO can reduce the observation burden and improve the observation accuracy by adding the nominal model information and synchronizing the control input and the system response output in time, transform the controlled process into a perfect model form finally. Based on the IMC principle, the first-order Taylor expansion is applied to the time delay part, and the zero-pole cancellation is carried out to obtain the parameter tuning rules of PID controller. The SMESO is used to improve load disturbance rejection performance, and the IMC-PID controller is used to obtain satisfactory set-point tracking performance. The two parts are designed separately, thus there is no need to compromise the set-point tracking response and disturbance response. Compared with the existing methods, the proposed method not only has satisfactory dynamic performance, but also has strong robustness.

2 Problem Statement

Consider a second-order delayed unstable process with two unstable poles and its nominal model are respectively as:

$$G_p(s) = \frac{b_p}{(\tau_{p1}s - 1)(\tau_{p2}s - 1)} e^{-\theta_p s} = \frac{k_p}{s^2 + a_{p1}s + a_{p2}} e^{-\theta_p s} \quad (1)$$

$$G_m(s) = \frac{b_m}{(\tau_{m1}s - 1)(\tau_{m2}s - 1)} e^{-\theta_m s} = \frac{k_m}{s^2 + a_{m1}s + a_{m2}} e^{-\theta_m s} \quad (2)$$

where $k_i = b_i / (\tau_{i1}\tau_{i2})$, $a_{i1} = -(\tau_{i1}\tau_{i2}) / (\tau_{i1}\tau_{i2})$, $a_{i2} = 1 // (\tau_{i1}\tau_{i2})$, $i = p, m$.

For delayed process, the classic IMC-PID feedback control structure is shown in Fig. 1. The IMC controller is designed as $C_{imc}(s) = f(s) / G_{m-}(s)$. $G_{m-}(s)$ is the minimum phase part of the nominal model, $f(s)$ is the IMC filter. If the load disturbance d is considered, the response transfer function of the closed-loop system is:

$$\begin{aligned} y(s) &= \frac{C_{imc}(s)G_p(s)}{1+C_{imc}(s)(G_p(s)-G_m(s))} r(s) \\ &+ \frac{G_p(s)(1-C_{imc}(s)G_m(s))}{1+C_{imc}(s)(G_p(s)-G_m(s))} d(s) \end{aligned} \quad (3)$$

If the nominal model is accurate (i.e., $G_m(s) = G_p(s)$) and the load disturbance is accurately compensated (i.e., $d = 0$), then

$$\frac{y(s)}{r(s)} = C_{imc}(s)G_p(s) = f(s) e^{-\theta_p s} \quad (4)$$

It is shown that when the model is perfect and the disturbance is compensated, the performance of the closed-loop system will be consistent with the IMC filter cascaded with the non-minimum part of the nominal model. That is to say, the satisfactory set-point tracking response can be obtained by selecting appropriate filter parameters. However, the actual system not only has load disturbance, but also has nonlinear dynamics and model errors. Eq. (1) can be rewritten as $\ddot{y} = f(y, \dot{y}, t) + d - a_{p1}\dot{y} - a_{p2}y + k_p u(t - \theta_p)$, $f(y, \dot{y}, t)$ is nonlinear dynamics and model error. Let the unknown total disturbance is $d_u = f(y, \dot{y}, t) + d - (a_{p1} - a_{m1})\dot{y} - (a_{p2} - a_{m2})y + [k_p u(t - \theta_p) - k_m u(t - \theta_m)]$. Therefore, (1) can be further rewritten as $\ddot{y} = d_u - a_{m1}\dot{y} - a_{m2}y + k_m u(t - \theta_m)$. Then the final control input is expressed as $u = u_0 - z_3/k_m$, u_0 is the initial control input by IMC-PID, z_3 is the estimated unknown total disturbance. If z_3 can accurately estimate the unknown total disturbance, we have $k_m u_0 = k_m u + z_3 \approx k_m u + d_u = k_m u + \ddot{y} + a_{m1}\dot{y} + a_{m2}y - k_m u(t - \theta_m)$. That is: $\ddot{y} + a_{m1}\dot{y} + a_{m2}y \approx k_m u_0(t - \theta_m)$.

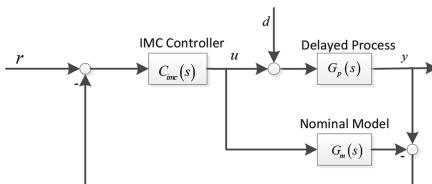


Fig. 1. Classic IMC-PID feedback

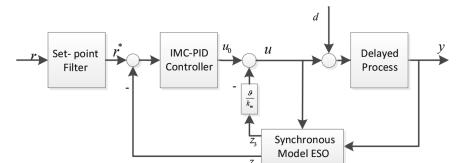


Fig. 2. IMC-PID controller based on SMESO

Thus, the delayed process with uncertainty system can be transformed into the nominal model (2). According to the nominal model, we can deduce the parameter tuning rules of PID controller based on IMC principle. Therefore, the key issue here is how to accurately estimate and compensate the unknown total disturbances to obtain a good disturbance response, and how to design the tuning rules of IMC-PID controller to obtain a satisfactory set-point tracking response.

3 Controller Design

3.1 Synchronous Model ESO

Traditional ESO is used to estimate and compensate the total disturbance. The total disturbance includes external disturbance (such as load disturbance) and internal disturbance (such as nonlinearity, model errors and un-modeled dynamics). However, traditional ESO does not use the known model information, thus it is repeated to estimate the known disturbance. Moreover, the control input and the system response output are not synchronized in time. All these factors increase the estimation burden of ESO and reduce the estimation accuracy.

Suppose d_u is bounded and differentiable. We add the known model information to ESO and delay the control input to get SMESO expressed in the following frequency domain:

$$\begin{cases} z_1(s) = \frac{\beta_1 s^2 + (\beta_2 + a_{m1}\beta_1)s + \beta_3}{\Delta} y + \frac{k_m s}{\Delta} u e^{-\theta_m s} \\ z_2(s) = \frac{(\beta_2 s + \beta_3 - a_{m2}\beta_1)s}{\Delta} y + \frac{k_m(s + \beta_1)s}{\Delta} u e^{-\theta_m s} \\ z_3(s) = \frac{(s^2 + a_{m1}s + a_{m2})\beta_3}{\Delta} y - \frac{k_m\beta_3}{\Delta} u e^{-\theta_m s} \end{cases} \quad (5)$$

where $\Delta = s^3 + (\beta_1 + a_{m1})s^2 + (\beta_2 + a_{m2} + \beta_1 a_{m1})s + \beta_3$, $[z_1 \ z_2 \ z_3]$ is the estimate of $[y \dot{y} \ d_u]$, $[\beta_1 \ \beta_2 \ \beta_3] = [3w_o \ 3w_o^2 \ w_o^3]$, w_o is the observer bandwidth. It is easy to prove that choosing large enough w_o can ensure that SMESO accurately estimate the state variable of each order, especially the total unknown disturbance, thereby improving the load disturbance rejection performance of the closed-loop system.

3.2 IMC-PID Controller

For the second-order delayed unstable process with two unstable poles, the minimum phase is $G_{m-}(s) = b_m / ((\tau_{m1}s - 1)(\tau_{m2}s - 1))$. Select the appropriate IMC filter as $f(s) = (\alpha_2 s^2 + \alpha_1 s + 1) / (\lambda s + 1)^3$. Therefore, the IMC-PID controller is designed as:

$$\begin{aligned} C_{pid}(s) &= \frac{C_{imc}(s)}{1 - C_{imc}(s)G_m(s)} \\ &= \frac{(\tau_{m1}s - 1)(\tau_{m2}s - 1)(\alpha_2 s^2 + \alpha_1 s + 1)}{b_m[(\lambda s + 1)^3 - (\alpha_2 s^2 + \alpha_1 s + 1)e^{-\theta s}]} \end{aligned} \quad (6)$$

The non-minimum phase part is approximated by the first-order Taylor expansion, i.e., $e^{-\theta s} \approx 1 - \theta s$. We have:

$$C_{pid}(s) = \frac{[\tau_{m1}\tau_{m2}s^2 - (\tau_{m1} + \tau_{m2})s + 1](\alpha_2s^2 + \alpha_1s + 1)}{b_m(3\lambda - \alpha_1 + \theta_m)s \left(\frac{\lambda^3 + \alpha_2\theta_m}{3\lambda - \alpha_1 + \theta_m}s^2 + \frac{3\lambda^2 - \alpha_2 + \alpha_1\theta_m}{3\lambda - \alpha_1 + \theta_m}s + 1 \right)} \quad (7)$$

In order to realize the controller and ensure the internal stability, the following zero-pole cancellation processing is adopted:

$$\begin{cases} \frac{\lambda^3 + \alpha_2\theta_m}{3\lambda - \alpha_1 + \theta_m} = \tau_{m1}\tau_{m2} \\ \frac{3\lambda^2 - \alpha_2 + \alpha_1\theta_m}{3\lambda - \alpha_1 + \theta_m} = -(\tau_{m1} + \tau_{m2}) \end{cases} \quad (8)$$

To improve the robustness and dynamic performance of the closed-loop response, a lead-lag compensator $R(s) = (0.5\theta_ms + 1) / (0.01s + 1)$ is connected in series behind the PID controller. The finally IMC-PID controller is obtained as follows:

$$\begin{aligned} C_{pid}(s) &= \frac{\alpha_2s^2 + \alpha_1s + 1}{b_m(3\lambda - \alpha_1 + \theta_m)s} R(s) \\ &= \frac{\alpha_1}{b_m(3\lambda - \alpha_1 + \theta_m)} \left(1 + \frac{1}{\alpha_1 s} + \frac{\alpha_2}{\alpha_1} s \right) \frac{0.5\theta_ms + 1}{0.01s + 1} \end{aligned} \quad (9)$$

3.3 IMC-PID Controller Based on SMESO

The following control input is recommended as $u = u_0 - \vartheta z_3/k_m$, ϑ is the disturbance compensation coefficient, which is bounded by $0 < \vartheta \leq 1$. In addition, to reduce the overshoot of set-point tracking response, we add a set-point filter: $F_r(s) = \alpha_2s^2 + \alpha_1s + 1$. The block diagram of the proposed structure is shown in Fig. 2.

4 Simulation Examples

Two examples are used to show the advantages of the proposed method in closed-loop response performance and robustness compared with the existing methods. To evaluate the performance and robustness of the control system, the integral of absolute error (IAE), the integral of square error (ISE) and total variation (TV) are considered as the evaluation indices.

4.1 Example 1

Considering the second-order delayed unstable process with a critical unstable pole [6]: $G_p(s) = e^{-0.2s} / (s(s - 1))$. To apply the proposed method, the process can be approximated as: $G_p(s) = 100e^{-0.2s} / ((100s - 1)(s - 1))$.

Suppose the nominal model is perfect, i.e., $b_m = 100$, $\tau_{m1} = 100$, $\tau_{m2} = 1$, $\theta_m = 0.2$. Take IMC filter parameter as $\lambda = 0.5$. From (8), we calculated that $\alpha_2 = 1.523$, $\alpha_1 = 1.696$. Other parameters are selected as $w_o = 10$, $\vartheta = 0.2$.

Ghousiya's method [6] and Zhao's method [12] are applied to control the process. To achieve a more convincing comparison, the filter parameters of Ghousiya's method also take $\lambda = 0.5$. The other parameters are $k_c = 3.682$, $\tau_i = 2.079$, $\tau_d = 1.0795$, $F_r(s) = (0.5s + 1) / (2.0681s^2 + 1.9944s + 1)$. The observer bandwidth of Zhao's method also take $w_o = 10$, the controller bandwidth is $w_c = 1$.

A unit step input to the set-point at $t = 0$ s and a negative step load disturbance with a magnitude 0.5 is given at $t = 20$ s. The closed-loop response is shown in Fig. 3. It can be seen that when the nominal model is perfect, the set-point tracking performance of the proposed method is similar to that of Ghousiya's method, but the load disturbance rejection performance is significantly improved. And the set-point tracking response of the proposed method is faster than Zhao's method. Consider the model parameters are mismatch within a certain. Case 1: Suppose b_p and θ_p both have +20% mismatch, i.e., $G_p(s) = 1.2e^{-0.24s} / (s(s - 1))$. Case 2: Suppose τ_{p1} and τ_{p2} both have -10% mismatch, i.e., $G_p(s) = e^{-0.2s} / (0.9s(0.9s - 1))$. Keep the control parameters unchanged, the resulting closed-loop responses are shown in Fig. 4 and Fig. 5. It is obvious that the proposed method can still retain smooth control behavior under large model errors. The corresponding IAE, ISE and TV indices for perfect case and mismatch case are given in Table 1.

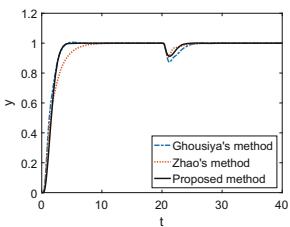


Fig. 3. Example 1

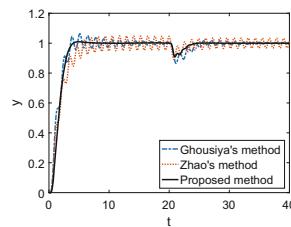


Fig. 4. Case 1

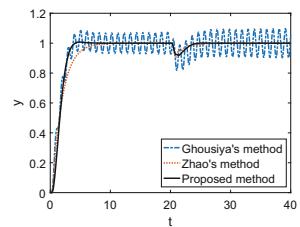


Fig. 5. Case 2

Table 1. Performance indices of Example 1

Method	Perfect model			Case 1			Case 2		
	IAE	ISE	TV	IAE	ISE	TV	IAE	ISE	TV
Ghousiya's method	0.720	0.051	7.642	0.897	0.057	54.48	2.197	0.165	302.2
Zhao's method	2.160	1.206	49.09	2.936	1.248	182.0	2.160	1.203	51.66
Proposed method	0.524	0.032	48.97	0.471	0.024	53.68	0.473	0.025	49.66

4.2 Example 2

Considering the second-order delayed unstable process with two unstable poles [3]: $G(s) = 2e^{-0.3s} / ((3s - 1)(s - 1))$.

If the nominal model is perfect, i.e., $b_m = 2$, $\tau_{m1} = 3$, $\tau_{m2} = 1$, $\theta_m = 0.3$. We calculated that $\alpha_2 = 4.148$, $\alpha_1 = 1.871$ by selecting $\lambda = 0.7$ in (8). The observer bandwidth and the disturbance compensation coefficient are $w_o = 8$ and $\vartheta = 0.2$, respectively. We compare the proposed method with Lee's method [3] and Zhao's method [12]. The corresponding controller tuning of Lee's method are $\lambda = 0.7$, $k_c = 2.3153$, $\tau_i = 1.7843$, $\tau_d = 1.8859$, $F_r(s) = 1/(3.252s^2 + 1.7147s + 1)$. The parameters of Zhao's method are tuning as $w_o = 8$ and $w_c = 0.8$. The closed-loop responses are shown in Fig. 6. It is obvious that the proposed method has better set-point tracking response and disturbance response. Considering the model mismatches. Case 3: Suppose b_p and θ_p both have $+10\%$ mismatch, i.e., $G_p(s) = 2.2e^{-0.33s}/((3s-1)(s-1))$. Case 4: Suppose τ_{p1} and τ_{p2} both have -10% mismatch, i.e., $G_p(s) = 2e^{-0.3s}/((2.7s-1)(0.9s-1))$. Keep the control parameters unchanged, and the corresponding closed-loop responses are shown in Fig. 7 and Fig. 8. It can be seen that the robustness of Lee's method is poor and the response of Zhao's method is slow. The corresponding IAE, ISE and TV indices for perfect case and mismatch case are given in Table 2.

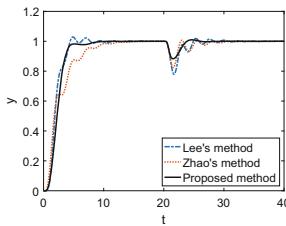


Fig. 6. Example 2

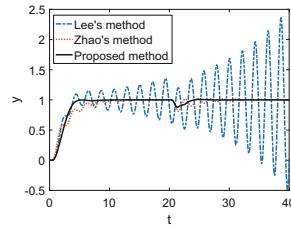


Fig. 7. Case 3

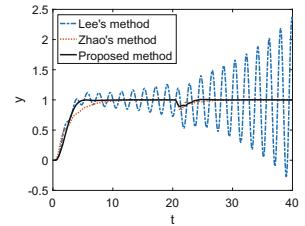


Fig. 8. Case 4

Table 2. Performance indices of Example 2

Method	Perfect model			Case 3			Case 4		
	IAE	ISE	TV	IAE	ISE	TV	IAE	ISE	TV
Lee's method	1.595	0.194	15.40	12.49	7.692	449.7	11.14	6.305	505.6
Zhao's method	3.204	1.689	17.29	3.277	1.697	30.11	3.190	1.683	16.68
Proposed method	1.674	0.214	10.84	1.614	0.197	12.59	1.576	0.189	14.28

5 Conclusions

In this paper, a parameter tuning method of IMC-PID controller based on SMESO is proposed for the second-order delayed unstable process. Compared with the traditional ESO, SMESO reduces the estimation burden, improves the estimation accuracy, and transforms the process into the nominal model. For the

delayed unstable process with two unstable poles, the simple and effective IMC-PID controller parameter tuning rules are derived. The IMC-PID controller is designed separately from the SMESO, which can improve the set-point tracking response and disturbance response together. The simulation results show that the proposed method has good closed-loop response performance and strong robustness, and can be applied to the actual unstable industrial process with time delay.

Acknowledgements. This project is supported by the Project of National Natural Science Foundation of China (Grant No. 51675106, No. U1601202, No. 51905108), and by the Guangdong Provincial R&D Key Projects under Grant 2018B090906002.

References

1. Rivera, D.E., Morari, M., Skogestad, S.: Internal model control: PID controller design. *Ind. Eng. Chem.* **25**, 2163 (1986)
2. Quinn, S.B., Sanathanan, C.K.: Controller design for integrating and runaway processes involving time delay. *AIChE J.* **35**, 923–930 (1989)
3. Lee, Y., Lee, J., Park, S.: PID controller tuning for integrating and unstable processes with time delay. *Chem. Eng. Sci.* **55**, 3481–3493 (2000)
4. Skogestad, S.: Simple analytic rules for model reduction and PID controller tuning. *Model. Ident. Control* **13**, 291–309 (2003)
5. Haugen, F.: Comparing pi tuning methods in a real benchmark temperature control system. *Model. Ident. Control* **31**, 79–91 (2010)
6. Ghousiya Begum, K., Seshagiri Rao, A., Radhakrishnan, T.K.: Enhanced IMC based PID controller design for non-minimum phase (NMP) integrating processes with time delays. *ISA Trans.* **68**, 223–234 (2017)
7. Wang, Q., Lu, C., Pan, W.: IMC PID controller tuning for stable and unstable processes with time delay. *Chem. Eng. Res. Des.* **5**, 120–129 (2016)
8. Chen, S., Xue, W., Huang, Y.: Analytical design of active disturbance rejection control for nonlinear uncertain systems with delay. *Control Eng. Pract.* **84**, 323–336 (2019)
9. Geng, X., Hao, S., Liu, T.: Generalized predictor based active disturbance rejection control for non-minimum phase systems. *ISA Trans.* **87**, 34–45 (2019)
10. Wang, X., Zhou, Y., Zhao, Z., Wei, W., Li, W.: Time-delay system control based on an integration of active disturbance rejection and modified twice optimal control. *IEEE Access* **7**, 130734–130744 (2019)
11. Zheng, Q., Gao, Z.: Predictive active disturbance rejection control for processes with time delay. *ISA Trans.* **53**, 873–881 (2014)
12. Zhao, S., Gao, Z.: Modified active disturbance rejection control for time-delay systems. *ISA Trans.* **53**, 882–888 (2014)
13. Zhang, D., Yao, X., Wu, Q.: ADRC based control for a class of input time delay systems. *J. Syst. Eng. Electron.* **28**, 1210–1221 (2017)



Fast Load Control of Supercritical Unit Through ADRC Optimized by Cross Quantum Pigeon

Guolian Hou¹, Ting Huang^{1(✉)}, Huilin Su¹, Linjuan Gong¹, Congzhi Huang¹, Bo Hu², and Bing Wang²

¹ School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China
90509753@ncepu.edu.cn

² State Grid Liaoning Electric Power Supply Co., Ltd., Shenyang 110004, China

Abstract. In order to absorb the random fluctuations caused by the large scale integration of new energy into the power grid, it is essential to improve the flexible operation ability of cool-fired especially supercritical units. In this paper, a fast load control strategy based on the active disturbance rejection control (ADRC) theory is proposed for the coordinated control system in supercritical units. Then, a novel cross quantum pigeon optimization algorithm is used to optimize the ADRC controller parameters automatically for coordinated control system. Finally, extensive simulation results demonstrate the feasibility and effectiveness of the optimal control strategy in fast load control.

Keywords: Supercritical units · Fast load control · ADRC · Cross quantum pigeon algorithm

1 Introduction

The instability of the power grid caused by large-scale integration of new energy is an urgent problem needed to be solved. As the most mature and stable power generation technic, the flexibility operation of supercritical units is crucial to this problem [1].

The flexibility focuses on the rapidity and depth of load regulation of supercritical units. As an important control system of the supercritical units, coordinated control system (CCS) has a direct impact on the safe and stable operation of units and the active power regulation level of the power grid. Therefore, it is of great significance to explore a coordinated control method for improving the rapidity and depth of load variation in CCS [2]. Aiming at the control of CCS, an energy-saving predictive control algorithm was adopted in [3] and showed excellent energy-saving performance. The flexibility of different heat-power combined technologies was summarized in [4], and the application of those technologies

was discussed from the perspective of power plant power generation-side and demand-sides. Based on the mechanical model subcritical circulating fluidized bed unit system, the control strategy based on dynamic matrix control was designed in [5] to realize fast and large-scale load variation of the unit.

The active disturbance rejection control (ADRC) was introduced as a new practical control technology [6] successfully and widely used in industrial fields. A genetic algorithm-based ADRC three-axis inertial stable platform parameter optimization method and inertially stabilized platform strategy were presented in [7]. The results show that, the proposed parameters tuning method has significant disturbance rejection ability which can improve the stabilization accuracy obviously. Moreover, an ADRC method for the gain scheduling of all-condition supercritical units was proposed in [8] to improve the tracking and disturbance rejection performance in both power output and throttle pressure loops of thermal power plants. ADRC control strategy shows excellent control performance in aviation and other industrial fields, but is rarely used in power industry. Therefore, this paper will combine ADRC control strategy with CQPIO to improve the load rapid tracking capability of Supercritical unit.

The automatic optimization of parameters still is an essential problem to be addressed in ADRC design. With the rapid development of bionic swarm optimization algorithms, numerous algorithms such as particle swarm optimization, whale algorithm, and ant colony algorithm had been widely used in ADRC parameter tuning (see [9–11]). The pigeon group optimization (PIO) algorithm inspired by homing behavior showed great advantages in exploration and exploitation. Individuals in the pigeon group locate their nests through map and compass operator navigation, and landmark operator navigation (see [12] and [13]). A control strategy based on Cauchy mutant pigeon heuristic optimization algorithm was adopted in [14] to improve the performance of multi-unmanned aerial vehicle path planning. Besides, a simulated annealing pigeon heuristic optimization algorithm was introduced in [15] to accomplish the target detection task for Unmanned Aerial Vehicles. The PIO algorithm had been applied in optimization process of some industrial fields successfully, but it is rarely used in power systems. The convergence speed and accuracy of standard PIO still had considerable expansion capacity. Therefore, the PIO algorithm will be improved this paper for the parameter setting of the ADRC in CCS to improve the flexible operation capability of the supercritical units.

The contributions of this paper are summarized as:

- A novel cross quantum PIO (CQPIO) is proposed by combining the cross theory and quantum behavior rules;
- The CQPIO is applied to the parameter optimization of ADRC approach;
- A fast load control strategy based on the designed ADRC is adopted for the CCS of supercritical units.

The remaining of this paper is organized as follows: Sect. 2 is the introduction of the CCS in supercritical unit. Then, the fundamentals of the ADRC design

process based on CQPIO are described in Sect. 3. Furthermore, extensive simulation results are displayed in Sect. 4. Finally, some conclusions are given in Sect. 5.

2 Coordinated Control System in Supercritical Unit

The main task of CCS is to coordinate the balance the load demands between of the power generation side and the power energy supply side. The boiler and steam turbine are controlled as a whole in CCS to ensure the fast load respond and maintain the overall stability of the plant.

Assuming that the power plant can meet the requirement of stable combustion in the furnace so that the coal-water ratio and the coal-air ratio are maintained at ideal levels, then the CCS of the supercritical unit is simplified to a multi-variable system with three inputs and three outputs.

The strong nonlinear coupling caused great difficulties in the control of CCS. A multiple-input multiple-output system (MIMO) is considered as combination of several multiple-input single-output (MISO) systems. In that case, the CCS can be regarded as a system with three subsystems, namely unit load subsystem, main steam pressure subsystem, and intermediate point temperature subsystem.

Based on the modeling method proposed in [16], the model of a supercritical unit is obtained:

$$Y(s) = G(s)U(s) \quad (1)$$

where, $Y(s) = [y_1(s), y_2(s), y_3(s)]^T$, $U(s) = [u_1(s), u_2(s), u_3(s)]^T$.

$$G(s) = \begin{bmatrix} \frac{-0.832s^2 - 0.713s - 0.067}{s^2 + 1.254s + 0.414} & \frac{0.079s^2 + 0.013s + 0.034}{s^2 + 1.254s + 0.414} & \frac{-0.003s^2 - 0.046s + 0.004}{s^2 + 1.254s + 0.414} \\ \frac{0.002s^2 - 0.03s - 0.05}{10s^2 + 11.03s + 1.22} & \frac{0.002s^2 - 0.002s - 0.004}{10s^2 + 11.03s + 1.22} & \frac{-0.0002s^2 + 0.0003s + 0.0008}{10s^2 + 11.03s + 1.22} \\ \frac{0.070s^2 - 0.222s - 0.725}{s^2 + 3.269s + 2.239} & \frac{0.022s^2 + 0.035s - 0.010}{s^2 + 3.269s + 2.239} & \frac{0.0002s^2 + 0.0002s - 0.0016}{s^2 + 3.269s + 2.239} \end{bmatrix} \quad (2)$$

where, u_1 is opening degree of main steam valve, u_2 is fuel flow, u_3 is feed water flow. y_1 is unit load, y_2 is main steam pressure and y_3 is intermediate point temperature.

Considering the importance of the load regulation in the flexibility operation of the supercritical unit, the control of the load subsystem mainly discussed in this paper. An optimal control strategy which is based on the ADRC theory and CQPIO algorithm is proposed to improve the rapid load variation capability and flexibility of supercritical unit.

3 ADRC Design Based on Cross Quantum Pigeon Optimization Algorithm

The ADRC controller is composed of tracking differentiator, feedback control law and extended state observer (ESO). The key of ADRC design focuses on ESO, which provides the state information inside the system. The parameter setting of

ESO is related to the control performance of the entire ADRC controller directly. Thus, the CQPIO algorithm is used to automatically turn the ADRC parameters in this paper, and a second-order ADRC is taken as an example.

3.1 Controller Design

Considering the following second-order system:

$$\ddot{y} = -a\dot{y} - by + w + (b - b_0)u + b_0u = f + b_0u \quad (3)$$

where, y is the output variable vector, u is the input variable vector, w is the external interference, a and b are unknown parameters. $b_0 \approx b$, f is the total disturbance of the system.

\hat{f} represents the estimated value of f , which is used to acquire the control law $u = (-\hat{f} + u_0)/b_0$. Then, the system can be reduced to a unit-gain double integrator control problem: $\ddot{y} = (f - \hat{f}) + u_0$, $(f - \hat{f})$ is disturbance.

Then, the state space form of the controlled object can be obtained:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 + b_0u \\ \dot{x}_3 = h \\ y = x_1 \end{cases} \quad (4)$$

where, $x_3 = f$ is the expansion state and $h = \dot{f}$ is the unknown disturbance. Therefore, f can be observed by a state observer based on the state space model.

The state space observer, denoted as the linear ESO, is constructed as Eq. (5).

$$\begin{aligned} \dot{z} &= Az + Bu + L(y - \hat{y}) \\ \hat{y} &= Cz \end{aligned} \quad (5)$$

where, $L = [\beta_1 \ \beta_2 \ \beta_3]^T$ is the gain vector of the ESO, which can be obtained by the pole placement method. Define ω_0 as the observer bandwidth, and configure the poles of the gain matrix at $-\omega_0$, and then,

$$L = [3\omega_0, 3\omega_0^2, \omega_0^3]^T \quad (6)$$

The PD controller can be designed as following control law:

$$u = \frac{-z_3 + u_0}{b_0} \quad (7)$$

Then, the system is reduced to a unit gain double integrator if ignoring the estimation error of z_3 :

$$\ddot{y} = (f - z_3) + u_0 \approx u_0 \quad (8)$$

Therefore, the control law can be expressed as Eq. (9):

$$u_0 = k_p(r - z_1) - k_dz_2 \quad (9)$$

where, r is the setpoint value.

In conclusion, the choice of b_0 , ω_o , k_p , and k_d is critical in the design of the ADRC controller. The bionic swarm intelligence optimization algorithm has been applied successfully in controller parameter tuning since its appearance. Therefore, the CQPIO proposed in this paper will be used to automatically optimize the parameters of ADRC for better control performance.

3.2 Parameters Optimization of ADRC Based on CQPIO Algorithm

In the CQPIO algorithm, the introduction of the wave function in the quantum rule can auxiliary the update of the individual position in the pigeon group, and the binomial crossing of the genetic algorithm combining with the global optimal value can promote the diversity of the population in the optimization process. The CQPIO algorithm reveals remarkable convergence speed and optimization accuracy. In the parameters optimization process of ADRC, each optimal value of optimized parameters is regarded as the location of the pigeon nest. In the initial stage of homing, the position of the pigeon nest is unknown, thus the current optimal candidate solution (the individual with the smallest fitness value) is assumed to be the pigeon nest. Take the error integral of the load model output displayed in *Eq. (10)* as the fitness function:

$$F(x_i^t) = \alpha_1 \left(\int_0^{t_s} (\eta_1 t |e(t)| + \eta_2 |u(t)| dt) \right) + \alpha_2 M_p \quad (10)$$

where, x_i is determined by CQPIO. $e(t)$ is the error between the system output and the setpoint value when the controller parameter is x_i^t . α_1 and α_2 represent the weight coefficients of the integral objective function and the index type objective function, respectively, and $\alpha_1 + \alpha_2 = 1$. η_1 and η_2 represent the weight coefficients of the error integral and the manipulated variable integral, respectively, and $\eta_1 + \eta_2 = 1$. t_s is the settling time of the system response, M_p represents the overshoot of the system response.

After selecting the current optimal candidate solution X_{gbest} , the quantum wave function $\ln(\frac{1}{\vartheta})$ is introduced in the navigation stage of the map and compass operator. Thus, the remaining pigeon individuals update their positions according to *Eq. (11)*:

$$X_i^t = P \pm \alpha |X_{mbest}^{t-1} - X_i^{t-1}| \ln \left(\frac{1}{\vartheta} \right) \quad (11)$$

$$\alpha = \frac{(1 - 0.5)(t - T_1)}{T_1} + 0.5 \quad (12)$$

$$P = \frac{\gamma X_{pbest_i}^{t-1} + \eta X_{gbest}^{t-1}}{(\gamma + \eta)} \quad (13)$$

where, $\gamma, \eta, \vartheta \in (0, 1)$; t is the current iteration number of the compass operator; T_1 is the total iteration number of the compass operator navigation; X_i is the

position vector of the i th pigeon, X_{mbest} is the average position of all individuals in the pigeons flock; X_{pbest_i} is the optimal position of the i th flock; X_{gbest} is the current global optimal solution. Combine X_{gbest} with Binomial Crossover:

$$sol_i^t = \begin{cases} X_{gbest}^t + \beta(\phi - 0.5) \left(X_{gbest}^t - X_i^t \right) & v < \sigma \\ X_i^t & others \end{cases} \quad (14)$$

where, $\phi, \sigma, v \in (0, 1)$, $\beta \in (-1, 1)$, sol_i^t represents the position of the pigeons after crossing.

Based on the fitness value of the individual, the individuals are sorted from small to large according to the bubbling method in the navigation stage of landmark operators. Then, the first half of the individuals is retained and the second half of the individuals with less navigation ability will be discarded. In addition, the center position of the remaining group will be calculated according to the retained individual position information, and be regarded as the maximum possible pigeon nest position (landmark). Besides, all the pigeon individuals update their position information according to Eq. (15):

$$x_i = x_i^{t-1} + \delta(x_{center}^{t-1} - x_i^{t-1}) \quad (15)$$

$$x_{center}^{t-1} = \frac{\sum_{i=1}^{N^{t-1}} x_i^{t-1} F(x_i^{t-1})}{\sum_{i=1}^{N^{t-1}} F(x_i^{t-1})} \quad (16)$$

where, $\delta \in (0, 1)$, $N^t = \frac{N^{t-1}}{2}$.

New position is generated from Eq. (17). Then, a greedy choice will be made between original ones and new position through Eq. (18):

$$\begin{aligned} X_{new_i} &= X_i^{t-1} + \kappa(X_{gbest}^{t-1} - \beta X_{mbest}^{t-1}) \\ \beta &= round(1 + \tau) \end{aligned} \quad (17)$$

$$X_i^t = \begin{cases} X_{new_i}, f(X_{new_i}) < f(X_i^{t-1}) \\ X_i^{t-1}, f(X_{new_i}) > f(X_i^{t-1}) \end{cases} \quad (18)$$

where, $\kappa, \tau \in (0, 1)$, X_{new_i} is the newly generated location of the i th pigeon, greedy choice of new individual and the original individual.

Similarly, the current global optimal individual X_{gbest}^t of each iteration is cross-combined with the binomial, and make a greedy choice for a better solution. After the maximum number of iterations is reached, the optimization process is finished and the global optimal individual X_{gbest} is obtained.

4 Simulation Results

In this section, the proposed quantum pigeon group algorithm is used in the design of linear ADRC based on the model of Eq. (1) and Eq. (2), and a simulation experiment of load fast tracking was carried out.

According to the previous theoretical analysis, the four parameters that need to be optimized are k_p , k_d , b_0 , w_o . The upper and lower limits of the search range are: PopMax = [0.05 0.2 0.2 0.05]; PopMin = [0 0 0 0]; The parameters of CQPIO algorithm are set as follows: dimension $D = 4$; Population size $N_o = 30$; Map and compass operator navigation maximum iteration number MaxDt1 = 20; landmark operator navigation maximum iteration number MaxDt2 = 10.

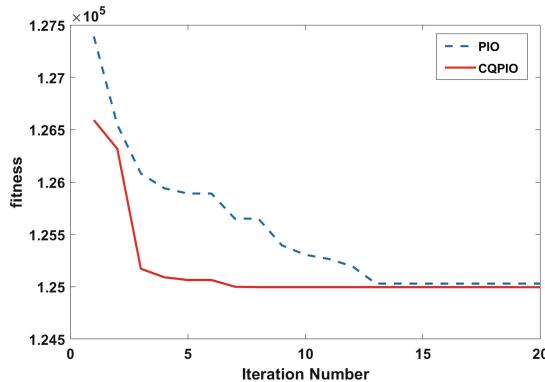


Fig. 1. Comparison of convergence curves between CQPIO and PIO algorithms

The optimization result of CQPIO algorithm is: $k_p = 0.00007591$; $k_d = 0.1164$; $b_0 = 0.1471$; $w_o = 0.0008224$. The convergence curves of the two comparison algorithms are shown in Fig. 1. The solid red line is convergence curves of the CQPIO algorithm, and the blue dashed line is convergence curves of the standard PIO algorithm. It can be seen that the optimization effect of the CQPIO algorithm is much better than the standard PIO, and the convergence speed is faster.

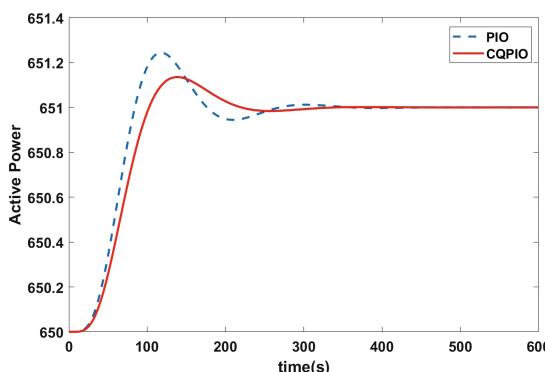


Fig. 2. Response curve of system load when the setting value r changes

In the simulation experiment, only the coal flow was changed while the other two input variables, namely, the steam turbine valve opening and feedwater flow keeping at steady-state values. When the setpoint of active load varies from 650 to 651, the corresponding load tracking result is showed as Fig. 2.

It can be seen from Fig. 2 that, system load can quickly track the rated value in a short time. This shows that ADRC has good tracking performance. Besides, the tracking performance of ADRC based on CQPIO algorithm is better than standard PIO algorithm.

In order to verify the effectiveness of the CQPIO algorithm in optimizing the parameters of the ADRC controller, four optimization algorithms namely CQPIO, standard PIO, PSO, and Emperor penguin optimizer(EPO)were compared. The results are shown in Table 1. Obviously, the Ave and MSE of the CQPIO algorithm are smaller than other algorithms, showing the highest optimization accuracy.

Table 1. Comparison of optimization errors of four algorithms

	Max	Min	AVE	MSE
CQPIO	1.1354	-5.42E-07	0.1045	0.1123
PIO	1.2444	-5.43E-07	0.1472	0.1282
EPO	1.1632	-5.61E-07	0.1153	0.1247
PSO	1.2095	-5.31E-07	0.1384	0.1266

5 Conclusion

Aiming at the safe and stable operation of the power grid under large-scale integration of new energy power generation, an ADRC strategy based on CQPIO optimization is proposed in this paper for fast load control of supercritical unit. In the design of ADRC, the classic PIO algorithm is improved by combining cross-cutting idea and quantum behavior rules, and applied to the parameters optimization. The simulation results show that the CQPIO algorithm has faster convergence rate and higher accuracy than the other compare algorithms. Thus, the adopted ADRC strategy shows remarkable control performance with excellent rapidity and accuracy. Therefore, the ADRC approach based on the CQPIO optimization can enhance the rapid load control capability and flexibility of the thermal power unit.

Acknowledgement. This work is supported by the National Key Research and Development Project (Grant No. 2019YFB1505403). The reviewers insightful comments and valuable suggestions are also greatly appreciated.

References

1. Wang, W., Liu, J.Z., Zeng, D.L., et al.: Modeling and flexible load control of combined heat and power units. *Appl. Therm. Eng.* **166**, 114624 (2019)
2. Wang, J.W., Shi, Y., Yi, Z., et al.: Flexibility of combined heat and power plants: a review of technologies and operation strategies. *Appl. Energy* **252**, 113445 (2019)
3. Hou, G.L., Gong, L.J., Huang, C.Z., et al.: Novel fuzzy modeling and energy-saving predictive control of coordinated control system in 1000 MW ultra-supercritical unit. *ISA Trans.* **86**, 48–61 (2019)
4. Zhang, H.F., Gao, M.M., Hong, F., et al.: Control-oriented modelling and investigation on quick load change control of subcritical circulating fluidized bed unit. *Appl. Therm. Eng.* **163**, 114420 (2019)
5. Zeng, D.L., Gao, Y.K., Hu, Y., et al.: Optimization control for the coordinated system of an ultra-supercritical unit based on stair-like predictive control algorithm. *Control Eng. Pract.* **82**, 185–200 (2019)
6. Han, J.Q.: Auto-disturbances-rejection controller and its applications. *Control Decis.* **13**, 1–5 (1998). (in Chinese)
7. Zhou, X.Y., Gao, H., Zhao, B.L., et al.: A GA-based parameters tuning method for an ADRC controller of ISP for aerial remote sensing applications. *ISA Trans.* **81**, 318–328 (2018)
8. Wu, Z.L., Li, D.H., Xue, Y.L., et al.: Gain scheduling design based on active disturbance rejection control for thermal power plant under full operating conditions. *Energy* **185**, 744–762 (2019)
9. Hu, D.D., Zhang, Y.C.: ADRC controller optimization design based on improved PSO algorithm for quad-rotor. *Appl. Res. Com.* **36**, 1762–1766 (2019). (in Chinese)
10. Yi, X., Wu, H.: Parameter tuning of ADRC based on double population chaotic whale algorithm. *Comput. Appl. Soft.* **36**, 281–285 (2019). (in Chinese)
11. Yin, Z.G., Du, C., Liu, J., et al.: Research on autodisturbance-rejection control of induction motors based on an ant colony optimization algorithm. *IEEE Trans. Ind. Electron.* **65**(4), 3077–3094 (2018)
12. Duan, H.B., Qiao, P.X.: Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning. *Int. J. Intel. Comput. Cybern.* **7**(1), 24–37 (2014)
13. Wang, H., Zhang, Z., Dai, Z., et al.: Heterogeneous pigeon-inspired optimization. *Sci. Chin. Inf. Sci.* **62**(7), 1–9 (2019)
14. Wang, B.H., Wang, D.B., Ali, Z.A.: A Cauchy mutant pigeon-inspired optimization—based multi-unmanned aerial vehicle path planning method. *Manag. Control* **53**(1–2), 83–92 (2020)
15. Li, C., Duan, H.B.: Target detection approach for UAVs via improved pigeon-inspired optimization and edge potential function. *Aero. Sci. Tech.* **39**, 352–360 (2014)
16. Hou, G.L., Yang, Y., Zhang, J.H., et al.: A new approach of modeling an ultra-supercritical power plant for performance improvement. *Energies* **9**(5), 310 (2016)



Finite-Time Containment Consensus Control of Second-Order Multi-agent Systems

Yan Cui and Qi Xue^(✉)

College of Physics and Information Engineering,
Shanxi Normal University, Linfen 041000, Shanxi, China
xueqi9406@163.com

Abstract. Finite-time containment consensus control problem is considered for multi-agent systems. A innovative finite-time containment control algorithm is designed to make followers converge to the leader's convex hull faster. In addition, based on piecewise function control method, a finite-time containment consensus control algorithm is proposed. Then by utilizing Lyapunov stability and homogeneous finite-time stability theorem, the control algorithm is proved to be able to make followers achieve finite-time consensus in leaders convex hull. Finally, lots of numerical simulation examples are given to demonstrate the rationality of the control algorithm.

Keywords: Finite-time · Consensus · Multi-agent · Containment control

1 Introduction

Containment control means that a group of followers converge and keep moving in the smallest geometric space (convex hull) enclosed by the leaders. Containment control has a number of potential applications in practice, for example, Polytopic Type Uncertainty [1], 4WS Vehicles [2].

Because of the advantages of finite-time consensus controller and the potential application of containment control, the study of multi-agent finite-time containment control has important significance. In literature [3–6], containment control of first-order, second-order, higher-order and fractional-order multi-agent systems are investigate respectively. The containment control problem based on state observer is studied in literature [7]. Literature [8–10] study containment control problems with parameter uncertainty, noise and multi-dynamic leaders, respectively. The above studies about containment control of multi-agent systems are all asymptotic containment control. But, in practical engineering applications, the system is generally required to achieve objective control in finite-time, it is also necessary to study the finite-time consensus problem of multi-agent

This work is supported by the NSFC (61503231, 61473015).

system. Literature [11] proposed a controller that can make multi-agent systems realize finite-time containment control faster than the traditional controller. Literature [12,13] studied the finite-time containment control of multi-agent systems by constructing interference observer and adaptive observer respectively. Although literature [11–13] have studied the finite-time containment control of multi-agents system, followers do not achieve consensus in the leader's convex hull. The innovation of this paper is that a piecewise function controller is proposed to study the finite-time containment consensus control of multi-agent systems.

2 Basic Knowledge and Model Design

2.1 Basic Knowledge

Assuming that \bar{G} is used to represent the leader-following multi-agent system and G is used to represent the followers system. The weighted adjacency matrix is expressed as $A = [a_{ij}] \in R^{N \times N}$ ($a_{ij} > 0$), the Laplacian matrix L of multi-agent system can be expressed as $L = \begin{bmatrix} 0_{m \times m} & 0_{m \times (n-m)} \\ L_{FL} & L_F \end{bmatrix}$.

Definition 1 ([14]). Consider the following continuous nonlinear system: $\dot{p} = f(p)$, $p(0) = p_0 \in R^N$ and the degree $\kappa \in R$ with the dilation coefficient $r = (r_1, r_2, \dots, r_n)$, $r_i > 0$ is homogeneous, if you have any $\varepsilon > 0$, $p \in R^n$, $f_i(\varepsilon^{r_1} p_1, \varepsilon^{r_2} p_2, \dots, \varepsilon^{r_n} p_n) = \varepsilon^{\kappa+r_i} f_i(p)$, $i = 1, 2, \dots, n$.

Lemma 1 ([15]). Suppose that system $\dot{p} = f(p)$, $p(0) = p_0 \in R^n$ is homogeneous of degree $\kappa < 0$ with the dilation $r = (r_1, r_2, \dots, r_n)$, $r_i > 0$, the function $f(p)$ is continuous and $p = 0$ is asymptotically stable point, then the system is finite-time stable.

Lemma 2 ([16]). Suppose that the topology \bar{G} corresponding to the multi-agent system has a directed spanning tree, then all the eigenvalues of L_F have positive real parts, every element of $-L_F^{-1}L_{FL}$ is non-negative, and the sum of the elements of each row of the matrix is 1.

2.2 Model Design

The model of multi-agent system can be depicted as:

$$\begin{cases} \dot{p}_i(t) = q_i(t) \\ \dot{q}_i(t) = u_i(t) \end{cases} \quad i = 1, 2, \dots, m, \dots, n \quad (1)$$

where, $p_i(t)$, $q_i(t)$, $u_i(t)$ represents the position, velocity status and control input of the i th agent. Let $\Re = \{1, 2, \dots, m\}$, $F = \{m + 1, \dots, n\}$.

This paper considers the situation of the static leaders, that is $q_i(t) = 0, i \in \mathfrak{R}$, the dynamic model (1) can be changed into:

$$\begin{cases} \dot{p}_i(t) = q_i(t) \\ \dot{q}_i(t) = \begin{cases} 0 & i \in \mathfrak{R} \\ u_i(t) & i \in F \end{cases} \end{cases} \quad (2)$$

Definition 2. For any initial state, the multi-agent system is able to achieve asymptotic consensus, , if exist $\lim_{t \rightarrow \infty} |p_i(t) - p_j(t)| = 0, \lim_{t \rightarrow \infty} |q_i(t) - q_j(t)| = 0, i, j \in F$.

Literature [15] proposes the following algorithm (3) to make the followers converge to the leader's convex hull faster than the traditional algorithm (4).

$$u_i(t) = \text{sig}^{\alpha_1} \left(\sum_{j \in N_i} a_{ij} (p_j(t) - p_i(t)) \right) + \text{sig}^{\alpha_2} \left(\sum_{j \in N_i} a_{ij} (q_j(t) - q_i(t)) \right) \quad (3)$$

$$u_i(t) = \text{sig} \left(\sum_{j \in N_i} a_{ij} (p_j(t) - p_i(t)) \right) + \text{sig} \left(\sum_{j \in N_i} a_{ij} (q_j(t) - q_i(t)) \right) \quad (4)$$

To make multi-agent realize finite-time containment control faster, the controller is depicted as follows:

$$\begin{aligned} u_i(t) = & \text{sig}^{\alpha_1} \left(3 \sum_{j \in N_i} a_{ij} (p_j(t) - p_i(t)) \right) + \text{sig}^{\alpha_2} \left(3 \sum_{j \in N_i} a_{ij} (q_j(t) - q_i(t)) \right) \\ & + \beta_1 \sum_{j \in N_i} a_{ij} (p_j(t) - p_i(t)) + \beta_2 \sum_{j \in N_i} a_{ij} (q_j(t) - q_i(t)) \end{aligned} \quad (5)$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are real numbers, $0 < \alpha_1 < 1, \alpha_2 = \frac{2\alpha_1}{\alpha_1+1}, \text{sig}^\alpha(p) = |p|^\alpha \text{sign}(p), \text{sign}(\cdot)$ is symbolic function.

Since the controller (5) can only cause every follower to converge into the leader's convex hull, it cannot make the followers converge to consensus within the leader's convex hull, in order to make the followers achieve consensus in the leaders convex hull, in this paper, a piecewise function control algorithm is designed as:

$$u(t) = \begin{cases} \text{sig}^{\alpha_1}(-3Lp(t)) + \text{sig}^{\alpha_2}(-3Lq(t)) \\ \quad - \beta_1 Lp(t) + \beta_2 Lq(t) & t < T \\ \text{sig}^{\alpha_1}(-3L1p(t)) + \text{sig}^{\alpha_2}(-3L1q(t)) \\ \quad - \beta_1 L1p(t) + \beta_2 L1q(t) & t \geq T \end{cases} \quad (6)$$

where $L = \begin{bmatrix} 0_{m \times m} & 0_{m \times (n-m)} \\ L_{FL} & L_F \end{bmatrix}, L1 = \begin{bmatrix} 0_{m \times m} & 0_{m \times (n-m)} \\ 0_{(n-m) \times m} & L_F \end{bmatrix}, L_F \in R^{(n-m) \times (n-m)}, L_{FL} \in R^{(n-m) \times m}$.

3 Finite-Time Containment Consensus Analysis

To prove simplification, let $p_L(t) = \{p_1(t) \cdots p_m(t)\}^T$, $q_L(t) = \{q_1(t) \cdots q_m(t)\}^T$, $p_F(t) = \{p_{m+1}(t) \cdots p_n(t)\}^T$, $q_F(t) = \{q_{m+1}(t) \cdots q_n(t)\}^T$.

First, demonstrate the leader's finite-time containment control for followers.

Theorem 1. *The controller (6) can enable the multi-agent system to realize finite-time containment control, if there is $T_0 \in [0, T]$ and the final state of each agent can be satisfied:*

- (1) $\lim_{t \rightarrow T_0} \|p_F + L_F^{-1}L_{FL}p_L\| = 0$, for $\forall t \geq T_0$, $p_F = -L_F^{-1}L_{FL}p_L$.
- (2) $\lim_{t \rightarrow T_0} \|q_F + L_F^{-1}L_{FL}q_L\| = 0$, for $\forall t \geq T_0$, $q_F = -L_F^{-1}L_{FL}q_L$.

Proof. When $t < T$ and $T_0 < T$

$$\dot{q}(t) = \text{sig}^{\alpha_1}(-3L_F p_F(t) - 3L_{FL}p_L(t)) + \text{sig}^{\alpha_2}(-3L_F q_F(t) - 3L_{FL}q_L(t)) \quad (7)$$

$$- \beta_1(L_F p_F(t) + L_{FL}p_L(t)) - \beta_1(L_F q_F(t) + L_{FL}q_L(t))$$

The containment error function can be denoted as:

$$\begin{cases} \bar{p}_F(t) = p_F(t) + L_F^{-1}L_{FL}p_L(t) \\ \bar{q}_F(t) = q_F(t) + L_F^{-1}L_{FL}q_L(t) \end{cases} \quad (8)$$

According to the system dynamics model:

$$\dot{q}(t) = \text{sig}^{\alpha_1}(-3L_F \bar{p}_F(t)) + \text{sig}^{\alpha_2}(-3L_F \bar{q}_F(t)) - \beta_1 L_F \bar{p}_F(t) - \beta_2 L_F \bar{q}_F(t) \quad (9)$$

Let

$$\begin{cases} m = -L_F \bar{p}_F(t) \\ n = -L_F \bar{q}_F(t) \end{cases} \quad (10)$$

Equation (9) can be expressed as:

$$\dot{q}(t) = \text{sig}^{\alpha_1}(3m) + \text{sig}^{\alpha_2}(3n) - \beta_1 m - \beta_2 n \quad (11)$$

Select Lyapunov function candidate:

$$V(t) = V_1(t) + V_2(t)$$

where $V_1(t) = \frac{1}{2} \bar{q}_F^T(t) L_F \bar{q}_F$, $V_2(t) = \sum_{i=1}^n V_{2i}(t)$

$$V_{2i}(t) = \int_0^{\left(\sum_{j \in N_i} a_{ij}(p_j(t) - p_i(t))^T\right)} (\text{sig}^{\alpha_1}(3s) + \beta_1 s) ds$$

$$\begin{aligned}
V_1 &= \bar{q}_F^T L_F \dot{\bar{q}}_F = -n^T \text{sig}^{\alpha_1}(3m) - n^T \text{sig}^{\alpha_2}(3n) - n^T \beta_1 m - n^T \beta_2 n \\
V_{2i}(t) &= \sum_{j \in N_i} a_{ij} (q_j(t) - q_i(t))^T \times (\text{sig}^{\alpha_1}(3 \sum_{j \in N_i} a_{ij} (p_j(t) - p_i(t))^T) \\
&\quad + \beta_1 \sum_{j \in N_i} a_{ij} (p_j(t) - p_i(t))^T)) \\
\dot{V}_2 &= \sum_{i=1}^n \dot{V}_{2i} = \dot{V}_{21} + \dot{V}_{22} \dots + \dot{V}_{2n} \\
&= \sum_{j \in N} a_{ij} (q_j(t) - q_i(t))^T \times (\text{sig}^{\alpha_1}(3 \sum_{j \in N_i} a_{ij} (p_j(t) - p_i(t))^T) \\
&\quad + \beta_1 \sum_{j \in N_i} a_{ij} (p_j(t) - p_i(t))^T)) = n^T \text{sig}^{\alpha_1}(3m) + n^T \beta_1 m
\end{aligned}$$

Taking the derivative of V gives:

$$\begin{aligned}
\dot{V}(t) &= \dot{V}_1(t) + \dot{V}_2(t) = -n^T \text{sig}^{\alpha_2}(3n) - n^T \beta_2 n \\
&= -(|n_1| |3n_1|^{\alpha_2} + \dots + |n_n| |3n_n|^{\alpha_2}) - \beta_2 (n_1^2 + \dots + n_n^2) \leq 0
\end{aligned}$$

When $\dot{V}(t) = 0$, $n = 0, m = 0$ can be obtained from Eq. (7), and $\bar{p}_F(t) = 0, \bar{q}_F(t) = 0$ can be obtained from Lemma 2 and Eq. (9), so $\dot{V}(t) = 0$ is only available at the equilibrium point. According to Lyapunov's second theorem, the multi-agent system is asymptotically stable at equilibrium, further obtain $p_F(t) = -L_F^{-1} L_{FLPL}(t)$, $q_F(t) = -L_F^{-1} L_{FLQL}(t)$, according to Eq. (7), it can be inferred that the multi-agent system can achieve containment control.

By defining 1, the dilation coefficient can be set to $(2, 2, \dots, 2, \alpha_1 + 1, \alpha_1 + 1, \dots, \alpha_1 + 1)$, Let $r_2 = \alpha_1 + 1, \kappa = \alpha_1 - 1 < 0$

The system equation can be written as:

$$\begin{cases} f_1(p_i, q_i) = q_i \\ f_2(p_i, q_i) = u_i \end{cases}$$

There are

$$\begin{aligned}
f_1(\varepsilon^{r_1} p_i, \varepsilon^{r_2} q_i) &= \varepsilon^{r_2} q_i = \varepsilon^{\kappa+r_1} f_1(p_i, q_i) \\
f_2(\varepsilon^{r_1} p_i, \varepsilon^{r_2} q_i) &= \text{sig}^{\alpha_1} \left(\sum_{j \in N_i} a_{ij} (\varepsilon^{r_1} p_j - \varepsilon^{r_1} p_i) \right) + \text{sig}^{\alpha_2} \left(\sum_{j \in N_i} a_{ij} (\varepsilon^{r_2} q_j - \varepsilon^{r_2} q_i) \right) \\
&\quad + \beta_1 \sum_{j \in N_i} a_{ij} (\varepsilon^{r_1} p_j - \varepsilon^{r_1} p_i) + \beta_2 \sum_{j \in N_i} a_{ij} (\varepsilon^{r_2} q_j - \varepsilon^{r_2} q_i)
\end{aligned}$$

That gives us that $\varepsilon^{r_1 \alpha_1} = \varepsilon^{r_2 \alpha_2} = \varepsilon^{\kappa+r_2}$, so the above equation can be expressed as $f_2(\varepsilon^{r_1} p_i, \varepsilon^{r_2} q_i) = \varepsilon^{\kappa+r_2} f_2(p_i, q_i)$, the multi-agent system is identical to degree $\kappa = \alpha_1 - 1 < 0$ with dilation coefficient $(\underbrace{2, 2, \dots, 2}_n, \underbrace{\alpha_1 + 1, \alpha_1 + 1, \dots, \alpha_1 + 1}_n)$, therefore, according to Lemma 1, the system can achieve finite-time convergence. To sum up, the multi-agent system can realize finite-time containment control in $t < T_0$.

Nextly, it is proved that the followers can achieve finite-time consensus in the leaders's convex hull in finite-time.

Theorem 2. The controller (6) can enable the followers to realize finite-time consensus, if there is $T_1 \in [T, +\infty)$ and the final state of each agent can be satisfied:

- (1) $m+1 \leq i, j \leq n, \lim_{t \rightarrow T_1} p_i(t) - p_j(t) = 0$, for $\forall t \geq T_1, p_i(t) = p_j(t)$.
- (2) $m+1 \leq i, j \leq n, \lim_{t \rightarrow T_1} q_i(t) - q_j(t) = 0$, for $\forall t \geq T_1, q_i(t) = q_j(t)$

Proof. when $t \geq T$ and $T_1 \geq T$, according to the system dynamics model:

$$\begin{aligned} \dot{q}(t) &= \text{sig}^{\alpha_1}(-3L_F p_F(t)) + \text{sig}^{\alpha_1}(-3L_F q_F(t)) \\ &\quad - \beta_1(-L_F q_F(t)) - \beta_2(-L_F q_F(t)) \end{aligned} \quad (12)$$

Let

$$g = -L_F p_F(t), h = -L_F q_F(t) \quad (13)$$

Equation (12) can be expressed as:

$$\dot{q}(t) = \text{sig}^{\alpha_1}(3g) + \text{sig}^{\alpha_1}(3h) - \beta_1 g - \beta_2 h \quad (14)$$

Select Lyapunov function condidate:

$$V(t) = V_1(t) + V_2(t)$$

$$\text{where } V_1(t) = \frac{1}{2} \bar{q}_F^T(t) L_F \bar{q}_F, V_2(t) = \sum_{i=1}^n V_{2i}(t)$$

$$V_{2i}(t) = \int_0^{\left(\sum_{j \in N_i} a_{ij}(p_j(t) - p_i(t))^T\right)} (\text{sig}^{\alpha_1}(3s) + \beta_1 s) ds$$

$$\dot{V}_1 = \bar{q}_F^T L_F \dot{\bar{q}}_F = -h^T \text{sig}^{\alpha_1}(3g) - h^T \text{sig}^{\alpha_2}(3g) - h^T \beta_1 g - h^T \beta_2 h$$

$$\begin{aligned} \dot{V}_{2i}(t) &= \sum_{j \in N_i} a_{ij}(q_j(t) - q_i(t))^T (\text{sig}^{\alpha_1}(3 \sum_{j \in N_i} a_{ij}(p_j(t) - p_i(t))^T) \\ &\quad + \beta_1 \sum_{j \in N_i} a_{ij}(p_j(t) - p_i(t))^T) \end{aligned}$$

$$\dot{V}_2 = \sum_{i=1}^n \dot{V}_{2i} = \dot{V}_{21} + \dot{V}_{22} + \dots + \dot{V}_{2n} = h^T \text{sig}^{\alpha_1}(3g) + h^T \beta_1 g$$

Taking the derivative of V gives

$$\begin{aligned} \dot{V}(t) &= \dot{V}_1(t) + \dot{V}_2(t) = -h^T \text{sig}^{\alpha_2}(3h) - h^T \beta_2 h \\ &= -(|h_1| |3h_1|^{\alpha_2} + \dots + |h_n| |3h_n|^{\alpha_2}) - \beta_2(h_1^2 + \dots + h_n^2) \leq 0 \end{aligned}$$

When $\dot{V}(t) = 0$ and $h = 0, g = 0$ can be obtained from Eq. (14), and $p_F(t) = 0, q_F(t) = 0$ can be obtained from Lemma 2 and Eq. (13), so $\dot{V}(t) = 0$ is

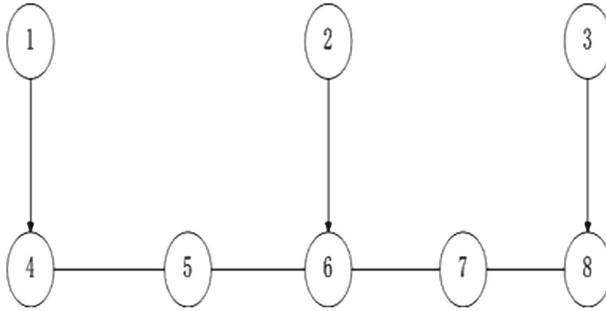


Fig. 1. The communication topology of multi-agent systems

only available at the equilibrium point. According to Lyapunov's second theorem, further obtain $p_i(t) = p_j(t), q_i(t) = q_j(t)$, by Definition 2, the multi-agent system can achieve consensus and global asymptotic stability. The finite-time convergence proof method of the followers is similar to Theorem 1. Finally, according to Lemma 1, the followers can realize finite-time convergence, to sum up, the followers can achieve finite-time consensus in $t \in (T_1, +\infty)$.

According to Theorems 1 and 2, it can be concluded that the second-order multi-agent system can realize containment consensus control in finite-time.

4 Numerical Simulation

Assuming that there are 3 leaders and 5 followers, the communication topology of the multi-agent system is shown in Fig. 1.

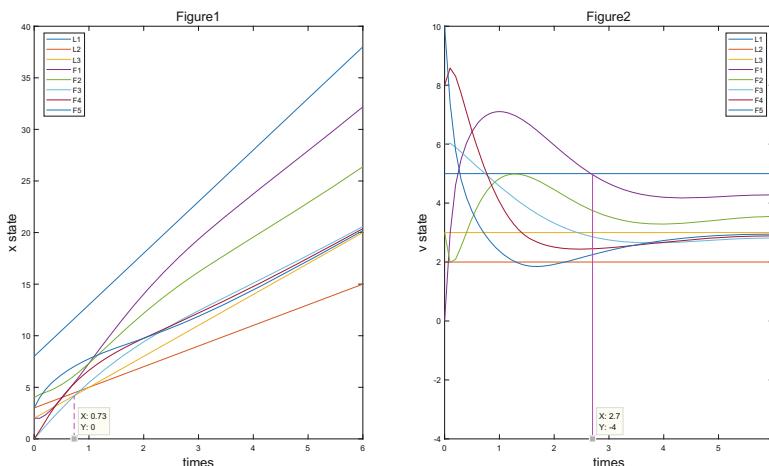


Fig. 2. Position and velocity curves of every agent with controller 1

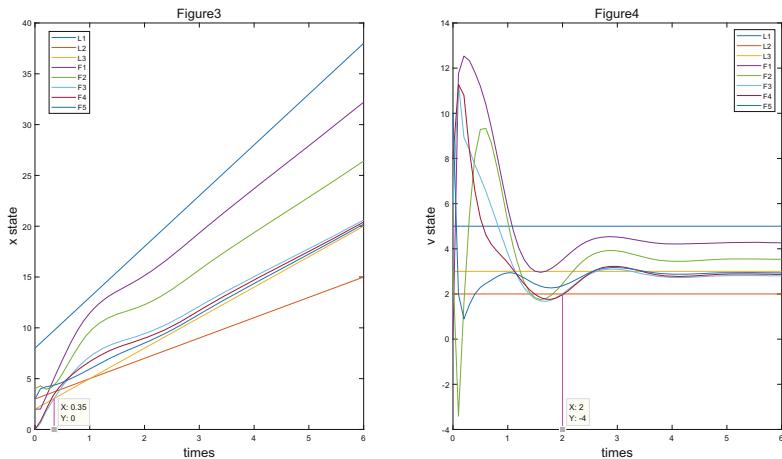


Fig. 3. Position and velocity curves of every agent with controller 2

In Laplacian matrix L , L_{FL} and L_F are

$$L_{FL} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, L_F = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 3 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix}.$$

Take the parameters $\alpha_1=0.8, \alpha_2=0.8889, \beta_1=9, \beta_2=2$ involved in the controller. Suppose the initial position of the three leaders are: $p_1(0) = 8, p_2(0) =$

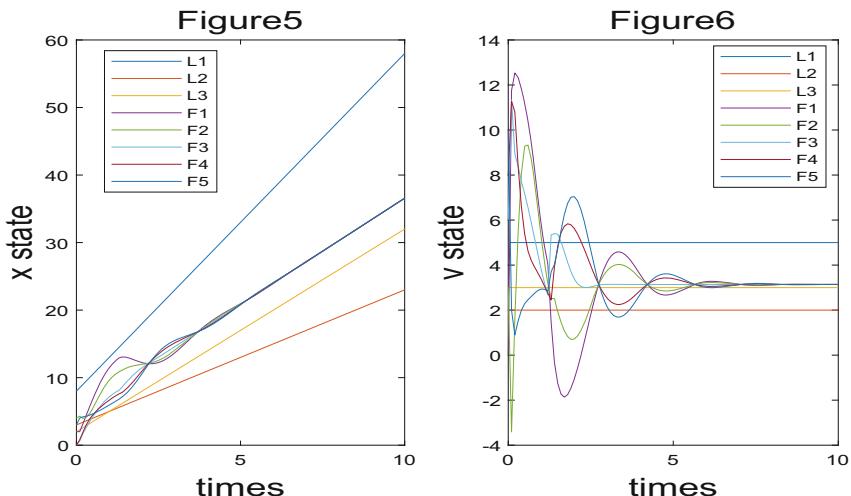


Fig. 4. Position and velocity curves of every agent with controller (6)

$3, p_3(0) = 2$, and the initial velocity are: $q_1(0) = 5, q_2(0) = 2, q_3(0) = 3$. The initial position of the five followers are: $p_4(0) = 2, p_5(0) = 4, p_6(0) = 0, p_7(0) = 0, p_8(0) = 3$, and the initial velocity are: $q_4(0) = 0, q_5(0) = 3, q_6(0) = 6, q_7(0) = 8, q_8(0) = 10$.

The algorithm (3) in Literature [15] is set as controller 1, and the algorithm (5) in this paper is set as controller 2. The simulation figure of controller 1 and 2 are shown in Fig. 2 and Fig. 3.

It can be seen from Figs. 2 and 3 that the controller designed in this paper can enable followers to converge into the leader's convex hull faster in finite-time, the rationality of the controller 2 designed in this paper are verified.

Take the parameters $T = 1.2$ in controller (6), the simulation figure of controller (6) are shown in Fig. 4.

From Fig. 4, we can see that the followers achieve finite-time consensus within the leader's convex hull.

5 Conclusion

A innovative finite-time containment consensus algorithm is proposed by utilizing piecewise control function method in this paper. The algorithm is proved to be able to make every follower converge to the leaders convex hull and achieve consensus within the convex hull in finite-time. Through simulation examples, the validity and rationality of the algorithm are further proved. Future work will focus on the problem of finite-time containment consensus control for multi-agent systems with time-delay and parameter uncertainty.

References

1. Jia, Y.M.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000). <https://doi.org/10.1109/87.845885>
2. Jia, Y.M.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003). <https://doi.org/10.1109/tac.2003.815033>
3. Wang, F.Y., Yang, H.Y., Zhang, S.N., et al.: Containment control for first-order multi-agent systems with time-varying delays and uncertain topologies. *Commun. Theor. Phys.* **66**(08), 249–255 (2016). <https://doi.org/10.1088/0253-6102/66/2/249>
4. Ma, L., Min, H.B., Wang, S.C., et al.: Distributed containment control of networked nonlinear second-order systems with unknown parameters. *IEEE/CAA J. Automatica Sinica* **5**(01), 232–239 (2018). <https://doi.org/10.1109/jas.2016.7510235>
5. Zhao, Y., Liu, Y.F.: Specified finite-time containment tracking for multiple high-order linear multi-agent systems. In: the 36th Chinese Control Conference, pp. 290–295 (2017). <https://doi.org/10.23919/chicc.2017.8028711>
6. Yang, H.Y., Wang, F.Y., Han, F.J.: Containment control of fractional order multi-agent systems with time delays. *IEEE/CAA J. Automatica Sinica* **5**(03), 727–732 (2018). <https://doi.org/10.1109/jas.2016.7510211>

7. Li, Z.G., Xie, G.M., Gao, P.: Observer-based containment control of multi-agent systems with linear dynamic. In: the 24th Chinese Control and Decision Conference, pp. 404–410. IEEE Press, Tai Yuan (2012). <https://doi.org/10.1109/ccdc.2012.6244088>
8. Xu, C.J., Su, H.S., Zheng, Y., et al.: Distributed adaptive containment for linear multi-agent systems using output information. In: the 34th Chinese Control Conference, Hang Zhou, pp. 890–895 (2015). <https://doi.org/10.1109/chicc.2015.7260847>
9. Wang, Y.P., Cheng, L., Hou, Z.G., et al.: Containment control of multi-agent systems in a noisy communication environment. *Automatica* **50**(7), 1922–1928 (2014). <https://doi.org/10.1016/j.automatica.2014.05.018>
10. Liu, H.Y., Cheng, L., Tan, M., et al.: Containment control of general linear multi-agent systems with multiple dynamic leaders: a fast sliding mode based approach. *IEEE/CAC J. Automatica Sinica* **1**(02), 134–140 (2014). <https://doi.org/10.1109/jas.2014.7004542>
11. Zhuang, H., Yang, H.Y.: Finite-time containment control of second-order multi-agent systems with jointly connected topologies. *CAAI Trans. Intell. Syst.* **12**(02), 188–195 (2017). <https://doi.org/10.11992/tis.201605013>
12. Li, Y.L., Yang, H.Y., Liu, F., et al.: Finite-time containment control of second-order multi-agent systems with mismatched disturbances. *Acta Automatica Sinica* **45**(09), 1783–1790 (2019). <https://doi.org/10.16383/j.aas.2018.c170571>
13. Yuan, C.Z., Zeng, W.: Output containment control of heterogeneous multi-agent systems with leaders of bounded inputs: an adaptive finite-time observer approach. *J. Franklin Inst.* **356**(6), 1–21 (2019). <https://doi.org/10.1016/j.jfranklin.2018.12.022>
14. Xiao, Q.Y.: Finite-Time Consensus Problems of Multi-agent Systems. University of The South (2015)
15. Yang, H.Y., Wang, F.Y., Weng, C.: Maximum consistence of complex multi-agent systems. *Comput. Simul.* **32**(6), 403–406 (2015). <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=JSJZ201506089&DbName=CJFQ2015>
16. Meng, Z.Y., Ren, W., You, Z.: Distributed finite-time attitude containment control for multiple rigid bodies. *Automatica* **46**(12), 2092–2099 (2010). <https://doi.org/10.1016/j.automatica.2010.09.005>



Track Tracking Control of Unmanned Intelligent Sweeping Vehicles Based on Improved MFAC

Wenlong Yao^{1,2}, Zhen Pang¹, Ronghu Chi¹, and Wei Shao^{1(✉)}

¹ School of Automation and Electrical Engineering,
Qingdao University of Science and Technology, Qingdao 266100, China
yaowenlong@qust.edu.cn

² School of Aerospace Engineering, Beijing Institute of Technology,
Beijing 100000, China

Abstract. Aiming at the trajectory tracking steering angle control problem of unmanned intelligent sweeper, the unknown parameters of sweeper system were considered, and an improved data-driven model-free adaptive control scheme was proposed. Specifically, under the given driving path, the trajectory tracking control problem of unmanned vehicle was transformed into the tracking control problem of pre-sighting deviation angle. Based on input and output data, a dynamic linearized data model for the transverse control of sweeping car was established. An time-varying proportional control item was added to the steering angle control rate. Then an improved MFAC trajectory tracking control scheme was designed and its convergence was proved. Simulation results show that: the trajectory tracking control scheme based on improved MFAC has strong adaptability and portability for trajectory tracking of different intelligent sweeping vehicle in different driving environments, with high tracking accuracy and tracking performance.

Keywords: Control engineering · Intelligent sweeping vehicles · Model-free adaptive control · Trajectory tracking · Unmanned · Pre-sighting deviation angle

1 Introduction

At present, with the improvement of road mechanical cleaning rate, the automatic operation of sweeper has become an important application scene in the field of unmanned driving [1]. Trajectory tracking is one of the most basic problems in motion control of driverless vehicles [2,3]. Since the vehicle is a complex nonlinear system with high coupling degree, it is difficult to establish an accurate vehicle dynamics system model. Moreover, the control system requires high speed and flexibility. So the trajectory tracking control of driverless vehicles is always a key and difficult point [4].

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 506–515, 2021.

https://doi.org/10.1007/978-981-15-8458-9_54

In recent years, many research achievements have been made on the lateral control of trajectory tracking of unmanned vehicle [5–7]. Kodagoda et al. proposed an uncoupled direct fuzzy PD/PI control scheme [5]. Feng D et al. designed an H optimum controller for the 4WS vehicle based on model following technology [6]. Yin GD et al. proposed a model predictive control with Legendre function based on three degrees of freedom vehicle model with magic formula’ tire model [7]. In practical application, it is very difficult to precisely establish the dynamic model of vehicle and environment due to the variable model parameters of different vehicles. When the vehicle type is different, the parameters of the algorithm also need to be adjusted again, with low flexibility and poor portability.

At present, in order to get rid of the dependence on mathematical models of controlled systems, a large number of data-driven control methods have been applied to mechanical, electric and transportation systems [8,9]. Model-free adaptive control (MFAC) is a typical data-driven control method proposed. This paper proposes an improved model-free adaptive control scheme for the trajectory tracking corner control of the intelligent sweeper, considering the requirements of response speed, accuracy, adaptability and portability of the control system.

2 Unmanned Trajectory Tracking Problem Description

2.1 Kinematic Model of Intelligent Sweeper

This paper focuses on four-wheel sweeper. Considering the low speed of the sweeper and the stable running condition of the vehicle, the kinematics model of the four-wheel drive is adopted in this paper. As is shown in Fig. 1.

In the figure, (x, y) represents the center point coordinate of the rear axle, v represents the current vehicle speed, φ is the heading Angle of the current vehicle, and represents the wheelbase of the front and rear axles. χ is the front wheel Angle. Thus, the kinematic model of car can be expressed as:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{bmatrix} = \begin{bmatrix} \cos \varphi \\ \sin \varphi \\ \tan \chi/n \end{bmatrix} v \quad (1)$$

2.2 Preview Deviation Angle Concept

Preview tracking means that depending on the speed of the car. The driver will focus on different points in the desired trajectory to control the direction. Preview tracking is more in line with the real driver operation habits. It is simple in structure and easy to realize. Figure 2 is the trajectory tracking diagram given according to the preview point.

Where, (x, y) is the position of the vehicle, (P_x, P_y) is the preview point on the target path, β is the heading Angle of the unmanned vehicle ($\beta + \varphi = \pi/2$), ϕ is the heading Angle of the preview point, α is the included Angle between

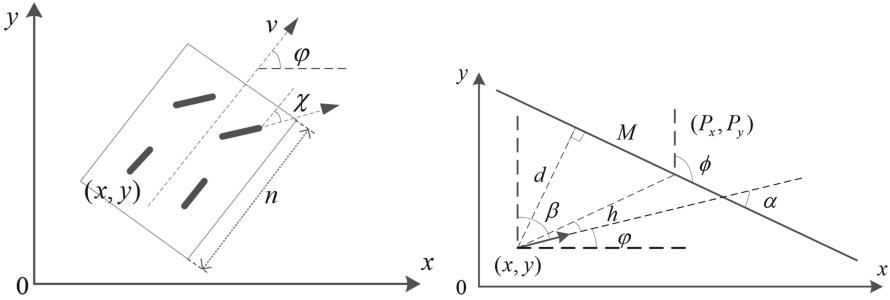


Fig. 1. Kinematics model of intelligent **Fig. 2.** Preview deviation angle model sweeper

the target path and the vehicle heading Angle extension line, d is the distance between the vehicle and the target path, M is the preview distance. h is the included Angle between the line connecting the car and the preview point and the direction of the car, that is, the preview deviation Angle.

2.3 Control Objectives

According to the theory of preview and combined with the working conditions of the unmanned intelligent sweeper, it is generally driven at a low speed. The selected preview distance is basically unchanged. From the trajectory tracking of the preview deviation Angle, it can be seen that the real-time position coordinate (x, y) of the vehicle can be given by the preview Angle through the kinematic model. When the driverless intelligent sweeper is running, the speed is low and the preview distance remains unchanged. The design controls the preview deviation Angle h by controlling the front wheel Angle χ , so that it can track to 0 quickly and without overshoot. So it can keep the sweeper on track.

3 Design of Improved MFAC Scheme for Track Tracking of Intelligent Sweeper

3.1 Improved MFAC Algorithm Design

Let $h(t)$ represent the preview deviation Angle at time t , and $u(t)$ represent the steering wheel Angle at time t . Where, $h(t)$ is the output at time t , and $u(t)$ is the input at time t . The discrete time nonlinear system is presented as follows:

$$h(t+1) = f(h(t), \dots, h(t-n_h), u(t), \dots, u(t-n_u)) \quad (2)$$

where, n_h and n_u are two unknown parameters, and $f(\dots)$ is an unknown nonlinear function. First, the following assumptions and theorems are proposed:

Hypothesis 1. $f(\dots)$ has a continuous partial derivative with respect to system input variable $u(t)$.

Hypothesis 2. The system (2) satisfies the generalized Lipschitz condition, That is, given any $u(t_1) \neq u(t_2)$ ($t_1 \neq t_2$ and $t_1, t_2 \geq 0$), there is:

$$|h(t_1 + 1) - h(t_2 + 1)| \leq m |u(t_1) - u(t_2)| \quad (3)$$

Where, $m > 0$ is a constant.

Theorem 1. For the nonlinear system (2) satisfying Hypothesis 1 and Hypothesis 2, when $\Delta u(t) \neq 0$, there must be time-varying parameter vector $\xi(t)$ to transform the system into the following dynamic linearized data model:

$$\Delta h(t + 1) = \xi(t) \Delta u(t) \quad (4)$$

Where, $\Delta u(t) = u(t) - u(t - 1)$, $\xi(t)$ is the pseudo-partial derivative of the system. Then the dynamic linearization model of the tracking system of the preview deviation Angle of the unmanned intelligent sweeper can be expressed as:

$$h(t + 1) = h(t) + \xi(t) \Delta u(t) \quad (5)$$

Consider the following input criterion function:

$$G(u(t)) = |h_d(t + 1) - h(t + 1)|^2 + \eta |u(t) - u(t - 1)|^2 \quad (6)$$

where, $\eta > 0$ is the weight factor.

Substitute Eq. (5) into Eq. (6). Take the derivative of $u(t)$ and set it equal to zero to obtain the control rate as follows:

$$u(t) = u(t - 1) + \frac{\mu \xi(t)}{\eta + |\xi(t)|^2} (h_d(t + 1) - h(t)) \quad (7)$$

where, $\mu \in (0, 1]$ is the step factor, which makes the control algorithm (7) more flexible. Since parameter $\xi(t)$ is unknown, it needs to be estimated.

Consider the following parameter estimation criteria function:

$$G(\xi(t)) = |h(t) - h(t - 1) - \xi(t) \Delta u(t - 1)|^2 + \rho |\xi(t) - \hat{\xi}(t - 1)|^2 \quad (8)$$

where, $\rho > 0$ is the weight factor.

By finding the extreme value of $\xi(t)$ in formula 8, the estimation algorithm of PID can be obtained as follows:

$$\hat{\xi}(t) = \hat{\xi}(t - 1) + \frac{\lambda \Delta u(t - 1)}{\rho + |\Delta u(t - 1)|^2} \times (\Delta h(t) - \hat{\xi}(t - 1) \Delta u(t - 1)) \quad (9)$$

where, $\hat{\xi}(t)$ is the estimated value of $\xi(t)$, and $\lambda \in (0, 1]$ is the step factor which also makes the control algorithm design more flexible.

Combined with the actual needs of intelligent sweeper track tracking, the algorithm has a better ability to track time-varying parameters. If $|\hat{\xi}(t)| \leq \tau$ or $|\Delta u(t-1)| \leq \tau$ or $\text{sign}(\hat{\xi}(t)) \neq \text{sign}(\xi(1))$, the reset algorithm is given as follows:

$$\hat{\xi}(t) = \hat{\xi}(1) \quad (10)$$

Where, τ is a small positive number and $\hat{\xi}(1)$ is the initial value of $\hat{\xi}(t)$.

The above is the design of the traditional compact format model-free adaptive control scheme. The current control rate only considers the influence of time-varying error integral on the system. And the response speed is slow. So it is difficult to achieve a good dynamic control effect. Therefore, based on the traditional model-free adaptive control algorithm, the design control rate is as follows:

$$u(t) = u(t-1) + \frac{\mu \hat{\xi}(t)}{\eta + |\hat{\xi}(t)|^2} \times (k_1(e(t) - e(t-1)) + k_2 e(t-1)) \quad (11)$$

where, $e(t) = h_d(t+1) - h(t)$, $k_1 > 0$, $k_2 \geq 0$, $k_1 \geq k_2$.

In the above formula, the time-varying proportion of the error of the preview Angle is introduced to improve the tracking speed of the steering Angle; The error integral control is adopted at the last moment to reduce the computation and save the computation time.

3.2 Convergence Analysis

In order to verify the convergence of tracking error and boundedness of system input and output in the improved model-free adaptive trajectory tracking algorithm, the following assumptions and theorem guarantees are given.

Hypothesis 3. The first element symbol of the system time-varying vector parameter $\xi(t)$ remains unchanged. And $0 < \xi(k) \leq \bar{\xi}$ is satisfied.

Theorem 2. For the discrete nonlinear system of intelligent sweeper (2), when Hypothesis 1–3 is satisfied and $2k_1 - k_2 < \min\left\{\frac{4\sqrt{\eta}}{\mu\xi}, \frac{2\sqrt{\eta}}{\mu}\right\}$ & $h_d(t+1) = h_d$ are guaranteed to be constants. It can be pointed out that there is a positive number $\eta_{\min} > 0$, when $\eta > \eta_{\min}$: the system output tracking error is asymptotically convergent, and $\lim_{t \rightarrow \infty} |h_d - h(t+1)| = 0$; The closed loop system is BIBO stable.

Proof. The proof of Theorem 2 is divided into two parts. In part 1, it is proved that the system output tracking error can converge to 0.

According to Eqs. (5) and (11), the following can be obtained:

$$e(t+1) = h_d - h(t+1) = (1 - \xi(t)\theta(t)k_1)e(t) + (k_1 - k_2)\xi(t)\theta(t) \times e(t-1) \quad (12)$$

Where, $\theta(t) = \mu\hat{\xi}(t)/(\eta + |\hat{\xi}(t)|^2)$. According to $0 < \xi(k) \leq \bar{\xi}$, $\eta > 0$ and $\mu \in (0, 1]$, $\theta(t) > 0$.

Let $a_1(t) = \xi(t)\theta(t)k_1$ and $a_2(t) = \xi(t)\theta(t)(k_1 - k_2)$, $a_1(t) > a_2(t) \geq 0$ can be obtained. And Eq. (12) is converted to:

$$|e(t+1)| \leq (|1 - a_1(t)| + a_2(t)) \max \{|e(t)|, |e(t-1)|\} \quad (13)$$

Take the constant $b = \max \{|1 - a_1(t)| + a_2(t)\}$. Equation (13) can be converted into:

$$|e(t+1)| \leq b \max \{|e(t)|, |e(t-1)|\} \quad (14)$$

According to $0 \leq b < 1$:

$$\max \{|e(t)|, |e(t-1)|\} \leq b \max \{|e(t-2)|, |e(t-3)|\} \quad (15)$$

According to Eqs. (14) and (15), the following can be obtained:

$$\begin{aligned} |e(t+1)| &\leq b \max \{|e(t)|, |e(t-1)|\} \leq b^2 \max \{|e(t-2)|, |e(t-3)|\} \dots \\ &\leq \begin{cases} b^{\frac{t}{2}+1} \max \{|e(0)|, |e(-1)|\}, & t \text{ is even number} \\ b^{\frac{t+1}{2}} \max \{|e(1)|, |e(0)|\}, & t \text{ is odd number} \end{cases} \end{aligned} \quad (16)$$

According to $0 \leq b < 1$ and the above formula, we can get:

$$\lim_{t \rightarrow \infty} |h_d - h(t+1)| = \lim_{t \rightarrow \infty} |e(t+1)| = 0 \quad (17)$$

In part 2, we prove that the input and output of the closed-loop system are bounded. The first part proves that the tracking error is bounded, and h_d is a constant. So the system output $h(t)$ is bounded. In addition, we have to prove that the input u is bounded.

The absolute value of the deformation in Eq. (11) can be obtained:

$$|\Delta u(t)| = |\theta(t)(k_1 e(t) + (k_1 - k_2)e(t-1))| \quad (18)$$

Let $d_1(t) = k_1\theta(t)$ and $d_2(t) = (k_1 - k_2)\theta(t)$, from $k_1 > 0$, $k_2 \geq 0$ and $k_1 \geq k_2$, Eq. (18) can be converted into:

$$|\Delta u(t)| \leq (d_1(t) + d_2(t)) \max \{|e(t)|, |e(t-1)|\} \quad (19)$$

Let $c = \max \{d_1(t) + d_2(t)\} < 1$, Eq. (19) can be converted into:

$$|\Delta u(t)| \leq c \max \{|e(t)|, |e(t-1)|\} \quad (20)$$

We can get:

$$|u(t)| \leq |u(0)| + \frac{2c}{1-b} \max \{|e(2)|, |e(1)|, |e(0)|\} \quad (21)$$

Thus, the input $u(t)$ is bounded, and the input and output of the closed-loop system are bounded.

4 Analysis of Simulation Result

Based on the improved MFAC scheme and the PID control scheme, we respectively present the simulation comparison experiment for two models of intelligent sweeper (1.5 m and 2.5 m) under the two operating conditions of the park's cycling condition (2.5 m/s) and the suburban driving condition(1.5 m/s). The tracking effects of trajectory tracking and preview deviation Angle tracking are also given. The incremental PID control algorithm is used for comparison. The controller is as follows:

$$du(t) = K_p(e(t) - e(t-1)) + K_i e(t) + K_d(e(t) - 2e(t-1) + e(t-2))$$

Where, $e(t) = h_d(t+1) - h(t)$ is the deviation Angle error of preview.

In this paper, the preview distance is fixed at 5 m. The initial position of the sweeper is (0, 0) and the initial heading Angle is 0. And the expected tracking route is a straight line one meter from the vehicle. Controller parameters $\hat{\xi}(1)$ is $[1, 1, 1]^T$, k_1 is 10, k_2 is 9, μ is 1, η is 1, λ is 1, ρ is 50, and τ is 10^{-5} . Considering the two models, the optimal PID parameters were obtained by cut-and-trial, K_p was 1.2, K_i was 0.1, and K_d was 0.11. Simulation of the two sweeper models is as follows.

4.1 Intelligent Sweeping Vehicles NO. 1 Model

For no. 1 sweeper, the comparison results of trajectory tracking simulation under two working conditions are given, as shown in Fig. 3.

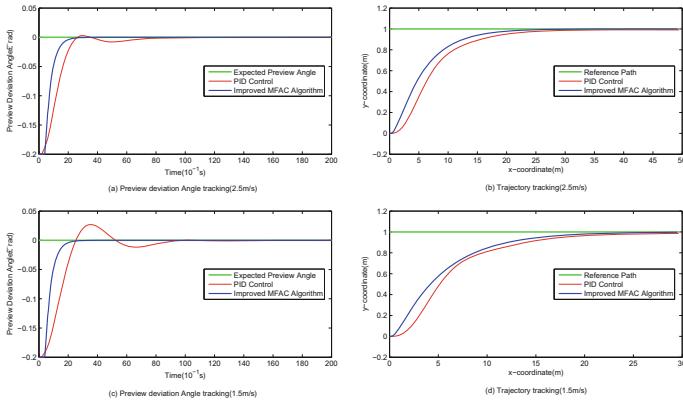


Fig. 3. Intelligent sweeping vehicles NO. 1 model

Figure 3(a) and (b) show the operating conditions in the park. It can be seen from the comparison that the tracking of the deviation Angle of preview under the improved MFAC scheme tends to 0 after 2 s. The tracking process is smooth

without overshoot. The deviation Angle of preview under PID control scheme has slight overshoot at about 3 s. The error of track tracking under MFAC is reduced to less than 0.01 m in 25 m x-coordinate, and the tracking accuracy is better than that of PID control scheme.

Figure 3(c) and (d) show the suburban driving condition. As can be seen that the deviation Angle of preview does not change significantly compared with the circulating conditions in the park, and it is basically tracked to 0 around 2 s, which has strong adaptability and tracking accuracy. The deviation Angle of preview under the PID control has a relatively obvious overshoot, reaching 0.025 rad. The driving and tracking error under the improved MFAC is less than 0.02 m at 20 m transverse, which is less than the PID control error and has better control accuracy.

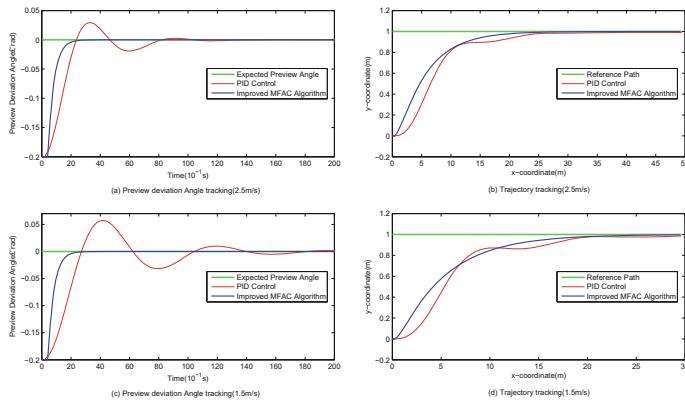


Fig. 4. Intelligent sweeping vehicles NO. 2 model

4.2 Intelligent Sweeping Vehicles NO. 2 Model

The wheelbase of no. 2 Sweeping Vehicle is increased, the turning radius is also increased, and the control difficulty of track tracking is increased, which requires higher control algorithm. If the algorithm parameters remain unchanged, the simulation results of trajectory tracking under different working conditions are shown in Fig. 4.

Figure 4(a) and (b) show the operating conditions in the park. It can be seen from the comparison that the deviation Angle of preview under the improved MFAC approaches 0 at about 2.5 s. The error of the driving trajectory is reduced to less than 0.01 m when it is about 25 m in abscissa. Under PID control scheme, the pre-sighting deviation Angle at 3.5 s and 6 s showed an overshoot of about 0.03 rad and 0.02 rad, respectively. The tracking effect became worse. Figure 4(c) and (d) show the suburban driving condition. The pre-sighting deviation Angle of no. 2 sweeper under the improved MFAC was tracked to 0 at about 3 s. The

tracking effect is stable. The deviation Angle of preview under PID control was overshoot in 4.2 s, 8 s, 12 s, etc., with a maximum of 0.055 rad. The tracking effect is relatively poor.

5 Conclusion

In this paper, an improved model-free adaptive control scheme based on the pre-sighting deviation Angle is proposed for the trajectory tracking steering Angle control of unmanned intelligent sweeper. The design of the controller only utilize the I/O data of the intelligent vehicle system, which overcomes the limitation that the traditional vector control relies on some knowledge of the system; A time-varying proportional control of preview Angle error is added to the steering input, which improves the flexibility of the system. The simulation results verify the effectiveness of the improved model-free adaptive control scheme applied to the track tracking system of the intelligent sweeper. It has good tracking speed and tracking accuracy. And can adapt to different models under different conditions of track tracking, has a good adaptability and portability.

Acknowledgement. This work was supported by Natural Science Foundation of Shandong Province (ZR2017MEE071), National Natural Science Foundation of China (61773227), Shortage Talents Program in Key Supported Regions of Shandong Province (Major office of Shandong Provincial Development and Reform Commission [2019] No. 391).

References

1. Cui, L., Zhang, T., Yin, X., Qi, Y.: Road Garbage cleaning device based on ZigBee gateway and image recognition. In: 2019 Chinese Control and Decision Conference, Nanchang, pp. 5603–5607 (2019). <https://doi.org/10.1109/CCDC.2019.8833453>
2. Song, X., Shao, Y., Qu, Z.: A vehicle trajectory tracking method with a time-varying model based on the model predictive control. IEEE Access **8**, 16573–16583 (2020). <https://doi.org/10.1109/access.2019.2963291>
3. Bhaskar, P.K., Yong, S.-P.: Image processing based vehicle detection and tracking method. In: 2014 International Conference on Computer and Information Sciences, Kuala Lumpur, pp. 1–5 (2014). <https://doi.org/10.1109/iccoins.2014.6868357>
4. Song, D., Tharmarasa, R., Kirubarajan, T., Fernando, X.N.: Multi-vehicle tracking with road maps and car-following models. IEEE Trans. Intell. Transp. Syst. **19**(5), 1375–1386 (2018). <https://doi.org/10.1109/tits.2017.2723575>
5. Kodagoda, K.R.S., Wijesoma, W.S., Teoh, E.K.: Fuzzy speed and steering control of an AGV. IEEE Trans. Control Syst. Technol. **10**(1), 112–120 (2002). <https://doi.org/10.1109/87.974344>
6. Du, F., Li, J., Li, L., Xue, Y., Liu, Y., Jia, X.: Optimum control for active steering of vehicle based on H model following technology. In: 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics, Wuhan, pp. 341–344 (2010). <https://doi.org/10.1109/car.2010.5456531>
7. Yin, G., Wang, Z., Jin, X.: Active steering of autonomous vehicle using model predictive control with legendre function. In: 2016 Chinese Control and Decision Conference, Yinchuan, pp. 3277–3281 (2016). <https://doi.org/10.1109/ccdc.2016.7531548>

8. Mei, J., Ren, Y., Jin, S., Hou, Z.: Freeway and side road balancing control scheme using MFAILC approach. In: 2017 6th Data Driven Control and Learning Systems, Chongqing, pp. 661–666 (2017). <https://doi.org/10.1109/ddcls.2017.8068151>
9. Li, X., Jia, X., Yan, S., Wang, S.: Data-driven model-free adaptive control of high power converters in super-large-scale energy storage station. In: 2019 Chinese Control Conference, Guangzhou, pp. 7245–7249 (2019). <https://doi.org/10.23919/chicc.2019.8865798>



Rapid Trajectory Design for Air-Launched High Speed Near Space Vehicle

Jianhui Liu^(✉), Mingang Zhang, Dongfei Hu, and Ningning Zhang

Science and Technology on Space Physics Laboratory, Beijing 100076, China
13693058549@163.com

Abstract. The trajectory design problem for air-launched rocket boosted high speed near space vehicle subject to multiple process and terminal constraints is discussed. A rapid ascent phase trajectory and descent phase trajectory design method is presented. The ascent trajectory is divided into a powered boost phase and an unpowered gliding phase. For ascent phase trajectory, the angle of attack is designed as an parametric function of time, and determined by two design parameters. The trajectory's state variables and terminal functions are computed by numerical integration, and by solving the nonlinear equations formulated by terminal conditions, the design parameters are determined efficiently. By choosing appropriate design parameters and regulating, descent phase trajectory is designed rapidly. Numerical simulation for demonstration test vehicle has been done and results are given, verifying that these methods are correct and effective.

Keywords: Trajectory · Design · Near space · Vehicle · Air-launched

1 Introduction

In recent ten years, the air-breathing powered vehicle and boost-glide hypersonic vehicles are attracting the interest of more and more academic and industrial institutions all over the world. Most of these vehicles are usually launched from ground or delivered by aircraft, then boosted by a single stage rocket or multi-stage rockets. The integrated system design and the design and test of vehicle's propulsion system, the GNC system, the thermal protection system (TPS) design and test are all dependent on the vehicle's flight trajectory, and are usually severely coupling with the trajectory. Therefore, the trajectory design and optimization of vehicle is critical to the integrated performance optimization design of hypersonic vehicle. In literature, there are some research works discussing about trajectory design, optimization issues for atmospheric flight vehicles [1–7]. However, the research literature on the topics about air-launched supersonic and hypersonic vehicle's trajectory design is relatively little. In Refs. [1,2], the flight dynamics theory and design methods for multi-stage launch vehicle and

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 516–523, 2021.

https://doi.org/10.1007/978-981-15-8458-9_55

atmospheric re-entry vehicles were described extensively. In Ref. [3], on board trajectory generation approach for entry vehicle was addressed. In Ref. [4], a boost phase trajectory optimization method based on genetic algorithm for ground launched hypersonic flight test model was discussed. In Refs. [5–7], the ascent trajectory design and optimization problems for air-breathing powered supersonic and near space vehicle were discussed. In this paper, the design and optimization problem subject to multiple process and terminal constraints for high speed near space vehicle is discussed, and a rapid flight trajectory design method for both the powered ascent phase and unpowered descent gliding phase of atmospheric vehicle is proposed.

2 Mathematical Model

In this section, the vehicle's mathematical model is formulated, including the equations of motion, the atmosphere model, the aerodynamic model, and the rocket engine model.

2.1 Equations of Motion

We use a non-rotating flat earth model, the gravitational acceleration is approximated to a constant $g_0 = 9.80665 \text{ (m/s}^2\text{)}$. The vehicle is assumed to be a controllable point mass, its equations of motion in vertical plane can be described as follows.

$$\dot{s} = V \cos \theta \quad (1)$$

$$\dot{h} = V \sin \theta \quad (2)$$

$$m \dot{V} = P \cos \alpha - D - mg \sin \theta \quad (3)$$

$$m v \dot{\theta} = P \sin \alpha - L - mg \cos \theta \quad (4)$$

$$\dot{m} = -m_d \quad (5)$$

In above equations, s is the flight range, h is the altitude, V is the velocity, P is the thrust of rocket, θ is the flight path angle, α is angle of attack, L is aerodynamic lift force, D is aerodynamic drag force, m is the total mass of vehicle (the initial mass is m_0), m_d is the fuel consumption rate.

2.2 Atmospheric Model

The density of atmosphere is modelled by an exponential function of altitude h .

$$\rho = \rho_0 e^{-h/\beta} \quad (6)$$

where, $\beta = 6700 \text{ (m)}$, is the scaled height, ρ_0 is the density of atmosphere at sea level. The atmospheric pressure p_h and speed of sound in the air V_s are functions of altitude, and they can be calculated by piecewise linear interpolation from

tabulated data listed in U.S. Standard Atmosphere 1976 [8]. Mach number Ma is defined and computed as

$$Ma = V/V_s \quad (7)$$

$$q = 0.5\rho V^2 \quad (8)$$

2.3 Rocket Engine Model

In boost phase, let m_{f0} denote the total mass of rocket's propellant, and T_0 denote the burning time of propellant. The average rate of propellant consumption of the rocket engine is computed as follows.

$$m_d = m_{f0}/T_0 \quad (9)$$

The rocket engine's thrust is computed as follows.

$$P = g_0 I_{sp} m_d - S_n p_h \quad (10)$$

In above equation, I_{sp} is the specific impulse of rocket in vacuum, S_n is the area of rocket engine's exit nozzle.

2.4 Aerodynamic Model

The aerodynamic drag force, lift force are computed by following equations.

$$L = q S_{ref} C_L(Ma, \alpha) \quad (11)$$

$$D = q S_{ref} C_D(Ma, \alpha) \quad (12)$$

In above equations, C_L and C_D denote the aerodynamic lift and drag coefficients, respectively. S_{ref} is the vehicle's reference area.

3 Ascent Phase Trajectory Design

3.1 Initial Conditions

The demo test vehicle is delivered by an aircraft horizontally, at a specified altitude and velocity. The initial mass is m_0 (kg), the initial altitude is $H_0(m)$, and the initial velocity is $V_0(m/s)$.

$$m(0) = m_0, h(0) = H_0, V(0) = V_0, \theta(0) = 0 \quad (13)$$

3.2 Terminal Conditions

Let T_1 denote the flight time of ascent phase trajectory. The expected terminal flight path angle of the trajectory is 0, and terminal altitude is H_1 , and the terminal velocity is V_1 . (H_1 and V_1 are to be determined.)

$$m(T_1) = m_1, h(T_1) = H_1, V(T_1) = V_1, \theta(T_1) = 0 \quad (14)$$

3.3 Constraints Model

For boost phase trajectory design, there are many constraints, the main constraints are the minimum and maximum available values for angle of attack, the minimum dynamic pressure, and the maximum value bound for normal overload ny.

$$\alpha_{min} \leq \alpha \leq \alpha_{max} \quad (15)$$

$$q \geq q_{min} \quad (16)$$

$$ny \leq ny_{max} \quad (17)$$

3.4 Trajectory Design Method

The objective of boost phase trajectory design is to generate a feasible trajectory effectively and rapidly. As usually, for 3DOF boost trajectory design, the angle of attack (AOA) is chosen as the design variable, and it is designed as a function of time t, with three design parameters to be determined, as follows.

$$\alpha(t) = \begin{cases} 0; & t < 1.0 \\ 4\alpha_m e^{-c(t-1)}(1 - e^{-c(t-1)}), & 1 < t < t_m \\ 4\alpha_m e^{-c(t_m-1)}(1 - e^{-c(t_m-1)}), & t_m < t \leq T_0 - 1 \\ 4\alpha_m e^{-c(t_m-1)}(1 - e^{-c(t_m-1)})(1 - t/T_0), & T_0 - 1 < t < T_0 \\ 0, & t > T_0 \end{cases} \quad (18)$$

In Eq. (18), t_m and α_m are design parameters. c is a control constant used to specify the time when the rocket reached the maximum negative angle of attack. In this paper, $c = 0.3$.

By appropriately choosing a pair of values of α_m and t_m parameters, the fourth order explicit Runge-Kutta's integration formula can be used to integrate Eqs. (1)–(5) to compute process and terminal state values and constraints values of powered boost phase and unpowered ascent phase. The terminal altitude h_f , velocity v_f , flight path angle θ_f , as well as the dynamic pressure q_f are obtained. α_f and θ_f are nonlinear functions of independent variables t_m and α_m . From a system of two nonlinear Eqs. (19) and (20), we can solve two parameters t_m and α_m satisfying the terminal conditions.

$$\theta_f(t_m, \alpha_m) = 0 \quad (19)$$

$$q_f(t_m, \alpha_m) - q_{min} = 0 \quad (20)$$

In practical implementation, we can pre-specify the value of α_m by a proportion of maximum available angle of attack α_{max} , then only t_m is a design parameter to be determined, such that the terminal dynamic pressure condition is satisfied. Let

$$\alpha_m = k\alpha_{max}, \quad 0.5 \leq k \leq 1.0$$

where k is a designed coefficient, it is used to control the maximum normal overload in ascent phase. We take $k = 0.9$ in this paper.

After powered boost phase, the flight path angle will monotonically decrease to zero finally, and the dynamic pressure is also monotonically decreasing, so the highest point of ascent phase is just the point with minimum dynamic pressure.

Therefore, in the integration process of Eqs. (1)–(5), we can use flight path angle as the termination criterion. When flight path angle is equal to or less than zero, the integration is terminated, terminal altitude, velocity, and dynamic pressure are computed and saved. So, the design problem of t_m is changed into a problem of solving following nonlinear equation.

$$q_f(t_m) - q_{min} = 0 \quad (21)$$

Since dynamic pressure q is a monotonically decreasing function of t_m , let $t_{max} = T_0 - 1$, and by choosing appropriate minimum value t_{min} of t_m (for example, we can take $t_{min} = T_0/3$), then Eq.(21) can be solved simply by bisection method rapidly.

Now, the ascent phase trajectory can be rapidly designed by integrating Eqs. (1)–(5).

4 Descent Phase Trajectory Design

4.1 Initial Conditions

The initial point of descent phase is just the terminal point of ascent phase, and it is also the highest point of the whole trajectory (the point with maximum altitude). The initial conditions of this phase are as follows.

$$m(0) = m_1, h(0) = H_1, V(0) = V_1, \theta(0) = 0 \quad (22)$$

4.2 Terminal Conditions

Let T_2 denote the flight time of descent phase trajectory. The expected terminal altitude and flight path angle are given as follows.

$$h(T_2) = H_g, \quad \theta(T_2) = 0$$

4.3 Constraints Model

In descent phase, in addition to maximum available angle of attack constraint, the vehicle is also subject to maximum dynamic pressure and maximum heating rate constraints.

$$\alpha_{2min} \leq \alpha \leq \alpha_{2max}$$

$$q \leq q_{max}$$

$$Q \leq Q_{max}$$

Where Q is the heating rate at stagnation point, it is computed by following equation.

$$Q = k_n \sqrt{\rho} V^{3.15}$$

Where k_n is a constant related with shaping of the vehicle's front nose cap.

4.4 Descent Trajectory Design

When the vehicle reaches the highest point, the ascent phase is ended, and the descent phase is started. The design purpose for descent phase is to pull up the trajectory. The angle of attack is chosen as trajectory design parameter in this phase, and it is designed as a piecewise constant function of altitude h .

$$\alpha(t) = \begin{cases} \alpha_1, & h \leq H2 \\ \alpha_2, & h > H2 \end{cases} \quad (23)$$

Where $\alpha_{2min} \leq \alpha_1 \leq \alpha_{2max}$, and $\alpha_{2min} \leq \alpha_2 \leq \alpha_{2max}$.

In above equation, $H2$, α_1 , α_2 are design parameters. By appropriately choosing $H2$ parameter and regulating α_1 and α_2 parameters, the terminal conditions of descent phase can be satisfied. And the trajectory state variables are obtained by numerical integration formula efficiently. Design process is described as follows.

First, let $\alpha_1 = \alpha_2 = \alpha_{2max}$, by using Eq. (23) and integrating Eqs. (1)–(4), we can get the terminal altitude of descent phase, $h(T2) = H_{2a}$.

If H_{2a} is greater than Hg , take $H2 = H_{2a}$, and let $\alpha_2 = \alpha_1$, and α_1 can be determined by bisection method, such that $h(T2) = Hg$.

If H_{2a} is less than Hg , take $H2 = H_{2a}$, and let $\alpha_1 = k_p \alpha_{2max}$ ($k_p \leq 1$), and α_2 can be determined by bisection method, such that $h(T2) = Hg$.

5 Numerical Simulation Results

A demo test vehicle example is used to verify methods discussed in the paper. Trajectory design parameters results and trajectory state variables such as altitude, velocity, flight path angle, and dynamic pressure results are given.

Some design parameters for the demo test vehicle are given as follows.

$H0 = 10\text{ km}$, $V0 = 240\text{ m/s}$, $m0 = 1500\text{ kg}$, $mf0 = 1000\text{ kg}$, $T1 = 25\text{ s}$, $Isp = 200\text{ s}$, $Hg = 28\text{ (km)}$.

For the ascent phase trajectory, designed α_m and tm results are: $\alpha_m = 11.8\text{ (deg)}$, $tm = 20\text{ (s)}$.

For the descent phase trajectory, designed parameters results are: $\alpha_1 = 9.1\text{ (deg)}$, $\alpha_2 = 7.0\text{ (deg)}$, $H2 = 32\text{ (km)}$.

Figure 1 shows the designed angle of attack profile result for ascent and descent phase trajectory.

Figure 2 shows the generated trajectory altitude and velocity profile result.

Figure 3 shows the generated trajectory flight path angle and dynamic pressure results.

The computation time of whole trajectory design process is less than 0.1 s on an ordinary desktop computer.

The numerical computation results prove that the trajectory design methods presented in this paper are correct and effective.

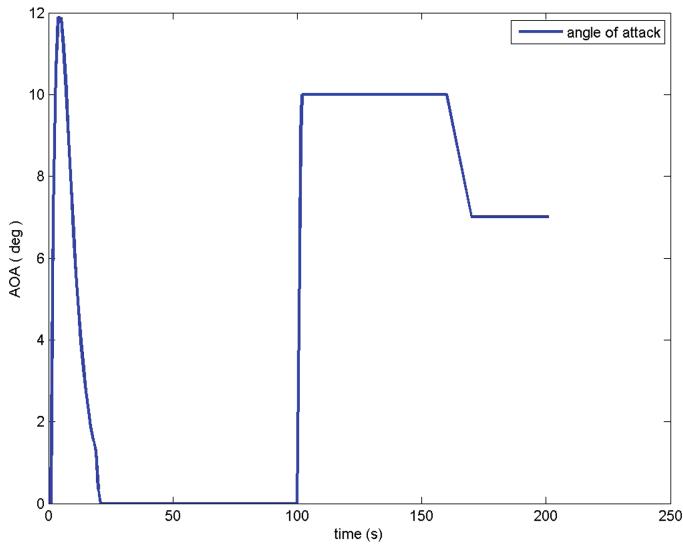


Fig. 1. Angle of attack design result

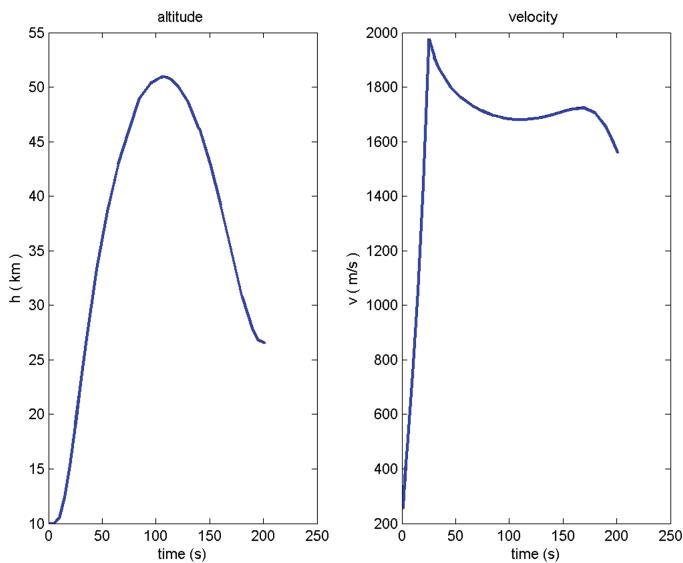


Fig. 2. Altitude and velocity design results

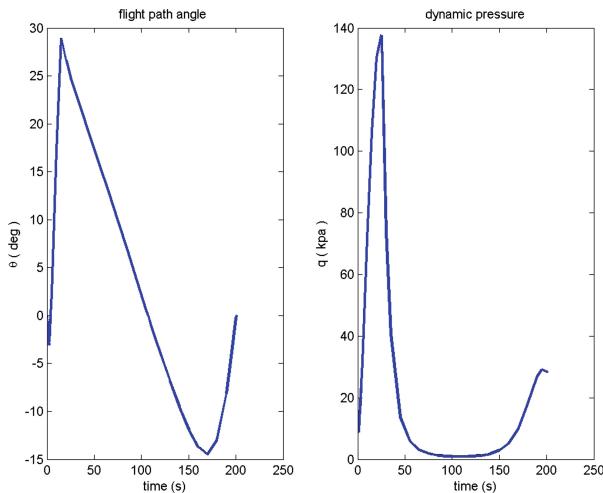


Fig. 3. Flight path angle and dynamic pressure design results

6 Conclusions

For an air-launched high speed near space vehicle, the boost phase and ascent phase trajectory design methods are discussed. The numerical computation and simulation results demonstrate that the methods presented in this paper are effective. The methods can be used in vehicle's conceptual engineering design process, and are also suitable for on board rapid trajectory design and reference flight trajectory profile generation of air-launched rocket boosted glide vehicle.

References

1. Vinh, N.X.: Optimal Trajectories in Atmospheric Flight. Elsevier Scientific, New York (1981)
2. Chen, K., Liu, L., Meng, Y.: Launch Vehicle Flight Dynamics and Guidance. National Defense Industry Press, Beijing (2014)
3. Shen, Z.J., Lu, P.: On-board generation of three-dimensional constrained entry trajectories. *J. Guid. Control Dyn.* **26**(1), 111–121 (2003)
4. Zhang, Z., Jiang, J., Wang, X.: Investigation on boost trajectory optimization method of hypersonic flight test. *J. Ballist.* **28**(3), 12–16 (2016)
5. Lu, P.: An inverse dynamic approach for trajectory design and guidance for aerospace plane. In: AIAA 92-4331-CP (1992)
6. Ming, C., Sun, R., Liang, Z., et al.: Trajectory Optimization design and analysis for air-breathing supersonic missile. *J. Solid Rocket Technol.* **39**(6), 833–838 (2016)
7. Zong, Q., Tian, B.L., Dou, L.Q.: Ascent phase trajectory optimization for near space vehicle based on Gauss pseudospectral method. *J. Astronaut.* **31**(7), 1775–1781 (2010)
8. U.S. 1976 Standard Atmosphere, U.S. Government's Printing Office (1976)



Advanced Simulation of Electric Trimmable Horizontal Stabilizer Actuation System Based on AMESim

Wei Li¹, Xudong Han², Wensen Zhang², Jian Fu^{2(✉)}, and Liming Yu³

¹ Nanjing Engineer Institute of Aircraft System Jincheng, AVIC,
Nanjing 211100, China

² School of Mechanical Engineering and Automation, Beihang University,
Beijing 100191, China
fujian@buaa.edu.cn

³ Shenyuan Honors College, Beihang University, Beijing 100191, China

Abstract. In the aircraft's flight control system, the trimmable horizontal stabilizer actuator (THSA) is used to balance the aircraft's pitch attitude, which is an important part for safety. In this paper, system scheme of an electrical type of THSA is firstly described and then the power transformation and multidisciplinary model are analyzed. Finally, the simulation analysis under different duty cycle are carried out in virtual simulation environment AMESim. The simulation results verify the accuracy of the established model and show the dynamic performance of the actuation system under various working conditions.

Keywords: Electrical actuation · Electromechanical actuator · Modeling and simulation · AMESim

1 Introduction

More-electric aircraft (MEA) has become the development trend of modern civil aircraft. The most important characteristics of MEA are the partial replacement of traditional hydraulic and mechanical systems through the electrical power system, and the wide application of power-by-wire (PbW) actuation systems. The electro-hydrostatic actuators (EHA) and electro-mechanical actuators (EMA) are the two most widely used types of PbW actuator systems [1–3]. Compared with traditional actuators, PbW has the advantages of higher power-to-weight ratio and cleaner [4].

As part of the aircraft's auxiliary control system, the horizontal stabilizer and elevator constitute the horizontal tail of the aircraft, which together achieve the pitch trim of the aircraft. The trimmable horizontal stabilizer actuator (THSA) can adjust the attitude angle of the horizontal stabilizer according to the instructions of the flight control computer. The screw nut is an important part of the horizontal stabilizer actuator. When the screw is rotated by the hydraulic motor

or electrical motor, the nut moves linearly on the lead screw, which drives the horizontal stabilizer connector to rotate a certain angle, and finally makes the horizontal stabilizer rotates. So that the deflection angle of the horizontal stabilizer can be controlled.

In addition to driving the horizontal stabilizer to rotate to accomplish the pitch trim of the aircraft, the horizontal stabilizer actuator can also provide support for the horizontal stabilizer, offsetting the hinge torque generated by the aerodynamic load on the stabilizer rotation axis. As a key component, the design and analysis of THSA is very important to the safety of the aircraft. Therefore, with the development of PbW, it is very necessary to carry out the performance simulation and analysis of electric THSA under different working conditions.

Reference [5] comprehensively analyzes the new THSA concept and presents the criteria to evaluate innovative architectures. Reference [6] designs a multivariable speed controller to adjust the rotational speed of the power source, which is implemented into the nonlinear simulation model. Reference [7] proposes several optimized algorithms of H_∞ controller, which are used to solve the THSA control strategy problem. Then these algorithms are verified on the test bench and compared with respect to the performance. However, there is little research about the advanced simulation on AMESim of ETHSA, which is important to analyze the dynamic performance.

This paper introduces the architecture scheme of an electric THSA and describes the composition of each part in Sect. 2. The power transformation and system working mode are analyzed in Sect. 3. Then AMESim model and parameters are presented, and simulations under different speed commands and load conditions are performed in Sect. 4. Finally, the simulation results prove that the system has good dynamic performance under different types of the load, and the active and backup motors can be switched quickly and steadily.

2 System Description

The horizontal tail is the wing surface that maintains the longitudinal stability and control of the aircraft. On most airplanes, it is installed symmetrically on the tail. Generally, the horizontal stabilizer and elevator form the horizontal tail. The horizontal stabilizer actuator is the object of this study. When the aircraft is affected by various updrafts or lateral winds, the navigation attitude will change. At this time, THSA will drive the horizontal stabilizer to realize the static stability in the pitch direction of the aircraft.

2.1 System Structure

The main components of ETHSA are a power unit, a control electronics, two motors, a gear reducer, a torque limiting clutch, a no-back mechanism and a ball screw. Figure 1 describes its main architecture scheme. The nut screw and gear reducer transmit the torque output by the motor to the horizontal stabilizer,

which can offset the load torque and accomplish the purpose of trimming. The two motors adopt the active and backup working mode, one of which works and the other as a backup. When the air load increases beyond a certain threshold, the torque-limiting clutch can prevent the motor from overloading by changing the friction state. When the horizontal stabilizer is stationary, the no-back mechanism can prevent the lead screw nut from rotating reversely under the drive of air load [8]. In this way, the horizontal stabilizer can be prevented from deflection.

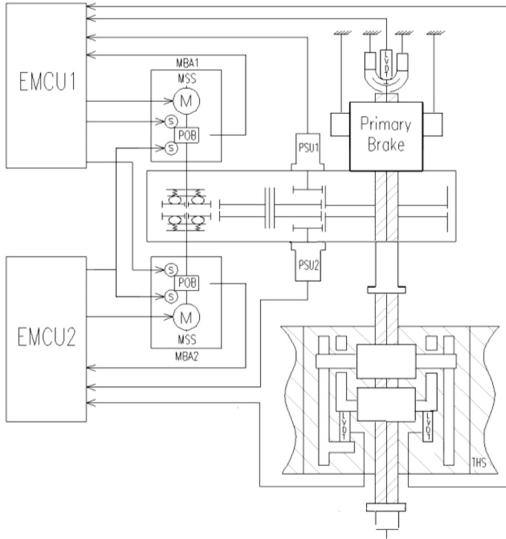


Fig. 1. Architecture scheme

2.2 Control Scheme

As shown in Fig. 2, the driver adopts classic three-loop control strategy to control the servo motor, which works in the position mode. The angle command is sent to the motor to adjust the output speed and torque by controller. The outermost loop is used to collect the speed of horizontal stabilizer to realize speed control. The controller adopts the traditional PID algorithm. In order to improve the control performance in the future, other modern control theories can be considered [9, 10].

3 Analysis of System Working Mode

As a complex system in which multiple energy forms are coupled, the simulation analysis of ETHSA is difficult. The conversion of different energy forms is diverse.

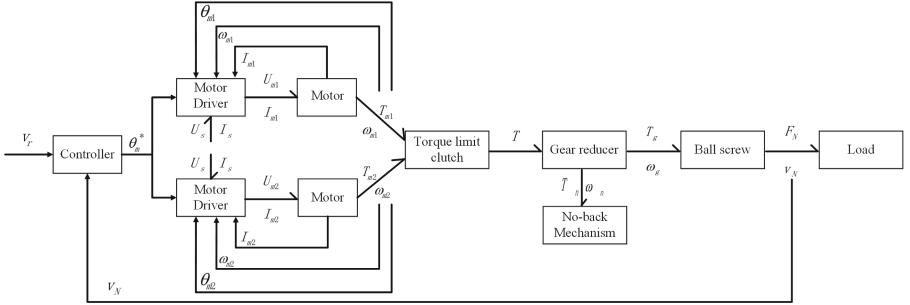


Fig. 2. Control signal and power flow of ETHSA

Based on the previous research work [8], the established mathematical models of the main parts of ETHSA, the working mode of the system can be analyzed. According to Fig. 2, the two motors adopt active and backup working methods, which are selected by control signals. The torque outputs by the motors are summed into the torque limiting clutch. According to the mathematical model of torque limiting clutch, when the torque applied by the air load increases to exceed the limit value, the original static friction between the friction discs will be converted into dynamic friction [8]. At this time, the torque output T by the torque limiting clutch is:

$$T = \begin{cases} T_{m1} + T_{m2} & T_{m1} + T_{m2} \leq T_{lim} \\ T_{lim} & T_{m1} + T_{m2} > T_{lim} \end{cases} \quad (1)$$

where T_{m1} and T_{m2} are the torque outputs by the two motors, and T_{lim} is the torque limited by the torque limiting clutch.

After the output of the torque limiting clutch enters the gear reducer, the torque can be increased to improve the load capacity of the motor. A part of the output torque of the gear reducer makes the screw rotate, and the other part enters the no-back mechanism. When the horizontal stabilizer is stationary, the no-back mechanism can prevent the lead screw nut from rotating reversely under the drive of air load. Therefore, the torque T_n entering the no-back mechanism is related to the axial force F_s received by the screw and the direction of the speed ω_s of the screw shaft:

$$T_n = \begin{cases} T_s & F_s \cdot \omega_s \leq 0 \\ f_c F_s & F_s \cdot \omega_s > 0 \end{cases} \quad (2)$$

where T_s and f_c are static friction force and coulomb friction coefficient of no-back mechanism, respectively.

4 System Model and Simulation

Based on the above analysis of the system, this paper developed a model of the ETHSA system under the AMESim environment as shown in Fig. 3. In this

model, the system mechanical structure, controller and nonlinear links are modeled, and the model can accurately reflect the real characteristics of the current system. For the electric trimmable horizontal stabilized surface actuation system, the system startup time and speed closed-loop stability are important indicators, and the switching of the motor working mode is also the focus of research. Table 1 shows the main simulation parameters in the model. The following is the simulation of ETHSA under several different working conditions.

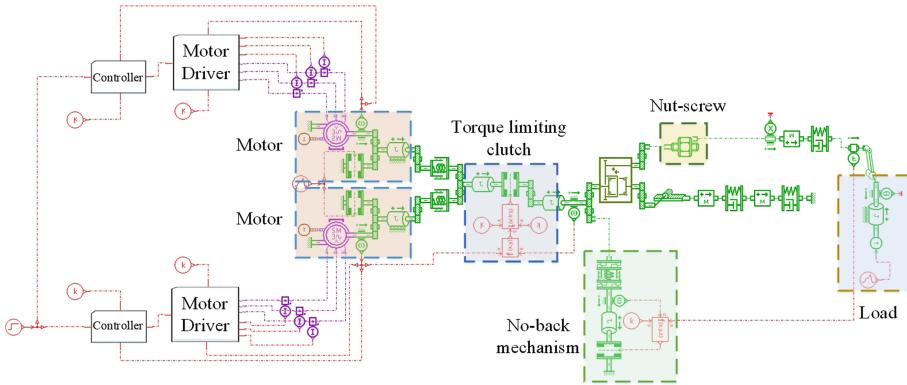


Fig. 3. AMESim model of THSA

4.1 Simulation of Constant Speed and Constant Load

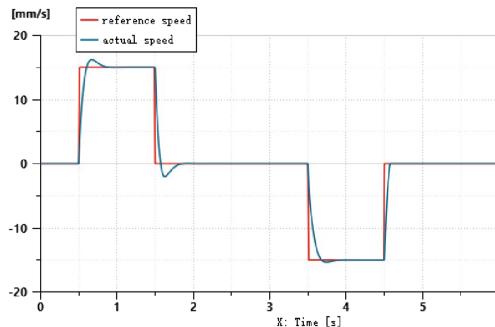
In this case, ETHSA receives a constant speed command and receives a constant air load on the load side. The comparison between the reference speed and the actual speed is clearly shown in Fig. 4. When the reference speed stepped from 0 to 15 mm/s at 0.5 s, the actual speed quickly followed this command and stabilized after about 0.3 s of overshoot. When the reference speed command is reversed, the actual speed can still quickly respond to this change. Figure 5 shows the current curve of the Q axis of the active motor.

4.2 Simulation of Variable Speed and Constant Load

In this situation, the reference speed changes within 15 mm/s with the law of ramp signal, and the rudder surface is still receiving 29 kN air load force at 0.5 s. Figure 6 shows the speed response curve of the system under variable speed and constant load. The actual speed curve only has a slight vibration when just adding the air load and changing the speed direction.

Table 1. Main parameters of the ETHSA models.

Parameters	Symbols	Units	Values
DC supply	U_s	V	270
Motor pole pairs	n_p		3
Motor resistance	R_m	Ω	0.89
Motor inductance	L_m	mH	1.95
Motor rotor inertia	J_m	kgm^2	1.410^{-4}
Permanent magnet flux linkage	Phi	wb	0.067
Diode forward voltage drop	U_d	V	0.85
Transistor forward voltage drop	U_t	V	0.75
Gear ratio	i		70.54
Stick torque of the torque limiter	T_s	Nm	14.9
Stick friction coefficient of primary brake	f_s		0.09
Coulomb friction coefficient of primary brake	f_c		0.078
Lead of screw	pit	mm	12.7
Inertia of screw	J_s	kgm^2	0.09
Structural stiffness	k_s	N/m	210^8
Structural damping	B_s	Ns/m	110^4
Inertia of flight surface	J_f	kgm^2	800

**Fig. 4.** Movement speed of the horizontal stabilizer

4.3 Simulation of Constant Speed and Variable Load

This situation is the opposite of the previous section, the rudder surface is subjected to an air load varying with a ramp signal between 29 kN, and at the same time, the reference speed is changed from 0 to 15 mm/s in 0.5 s. It can be seen from the Fig. 7 that with the fluctuation of the air load, the actual speed curve will have a slight vibration with an amplitude of about 0.1 mm/s. Overall, this result shows that ETHSA can resist changes in air load within a certain range, while keeping the speed response at a stable value with only small fluctuations.

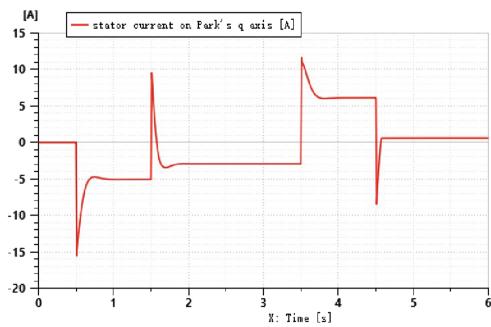


Fig. 5. Q-axis current of active motor

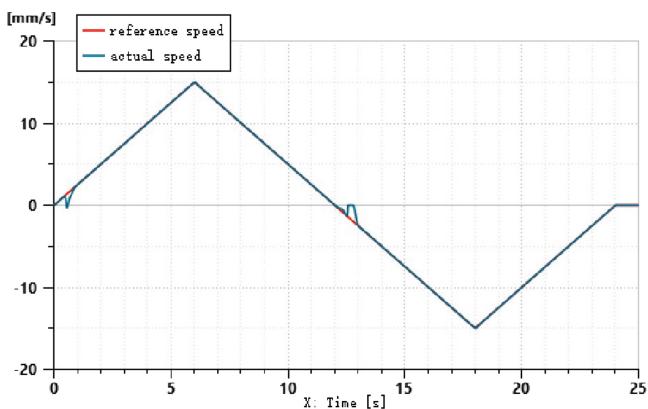


Fig. 6. System trim speed tracking response

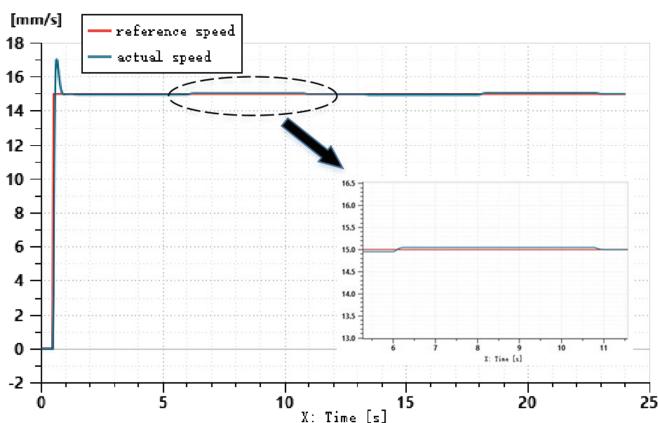


Fig. 7. Movement speed of the horizontal stabilizer

4.4 Switching Simulation of Active and Backup Motors

This simulation condition is under constant speed and constant load. The active motor is switched to the backup state at 3 s, and at the same time, the backup motor is switched to the active state. The change of trim speed of the system is shown in Fig. 8. Figure 9 describes the change process of the output torque of the two motors. It can be seen that the switching process has no obvious effect on the system trimming process, but there is a slight speed mutation at the moment of switching. After about 0.2 s, the system speed response quickly returns to the normal servo state.

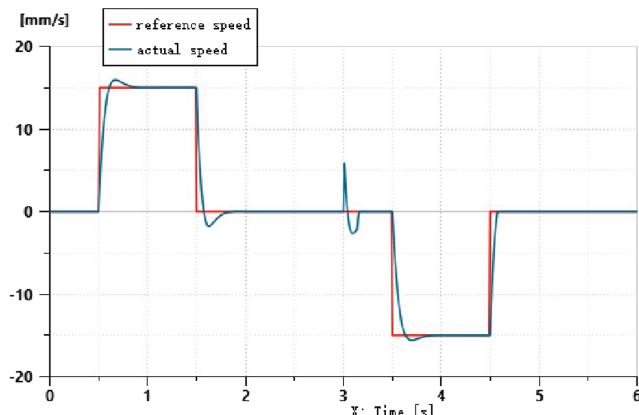


Fig. 8. Movement speed of the horizontal stabilizer

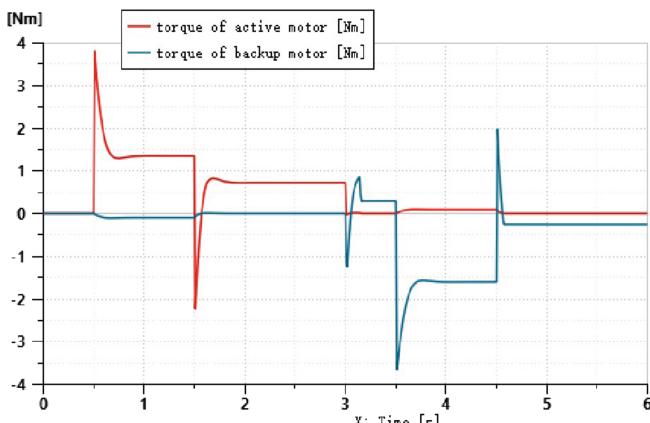


Fig. 9. Torque of active motor and backup motor

5 Conclusion

This paper introduces the architecture and control scheme of ETHSA. The simulations are accomplished in the software of AMESim according to the established mathematical models. The established model appropriately considers various physical effects. The simulation results prove that the system has a good dynamic performance under different duty cycle, and the switching of dual motors shows that the dualchannel active and backup working mode is effective. When more parasitic effects are taken into account at more complex level, the model can be used for energy loss analysis and efficiency calculation.

References

1. Mar, J.-C., Fu, J.: Review on signal-by-wire and power-by-wire actuation for more electric aircraft. *Chin. J. Aeronaut.* **30**(03), 857–870 (2017). <https://doi.org/10.1016/j.cja.2017.03.013>
2. Jinker, P., Claeysen, F.: New actuators for aircraft and space applications. In: 10th International Conference on New Actuators, Bremen, Germany, pp. 14–16 (2006)
3. Botten, S.L., Whitley, C.R., King, A.D.: Flight control actuation technology for next-generation all-electric aircraft. *Technol. Rev. J.* **8**(2), 55–68 (2000). Millennium Issue. Fall/Winter
4. Fu, J., Mar, J.-C., Fu, Y.L., et al.: Incremental modelling and simulation of power drive electronics and motor for flight control electro-mechanical actuators application. In: IEEE International Conference on Mechatronics and Automation, Beijing, China, (2015). <https://doi.org/10.1109/ICMA.2015.7237676>
5. Johnsen, S., Thielecke, F.: Integration analysis of trimmable horizontal stabilizer actuators and technology evaluation. *CEAS Aeronaut. J* **2**(1–4), 11–19 (2011). <https://doi.org/10.1007/s13272-011-0029-1>
6. Wachendorf, N., Thielecke, F., Carl, U., et al.: Multivariable controller design for a trimmable horizontal stabilizer actuator with two primary load paths. In: ICAS Secretariat - 26th Congress of International Council of the Aeronautical Sciences, vol. 4, pp. 1796–1807 (2008)
7. Huang, Y., Pe, T., Popov, A.P., et al.: Control of a two-load-path trimmable horizontal stabilizer actuator of an aircraft comparison of H_∞ design approaches. In: 49th IEEE Conference on Decision and Control (CDC) (2010)
8. Han, X.D., Ma, J.S., Fu, J., et al.: Modelling and simulation of an electric trimmable horizontal stabilizer actuator based on bond graph. In: 2019 15th Chinese Intelligent Systems Conference (CISC), Hai Kou, China, pp. 417–425 (2019). https://doi.org/10.1007/978-981-32-9698-5_47
9. Jia, Y.M.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000). <https://doi.org/10.1109/87.845885>
10. Jia, Y.M.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003). <https://doi.org/10.1109/TAC.2003.815033>



Characteristic Analysis of Mechatronic Equipment of Flight Control System for More Electric Aircraft

Tuanhui Guo, Kunxu Cui, Jian Fu^(✉), and Yongling Fu

School of Mechanical Engineering and Automation, Beihang University,
Beijing 100191, China
I.fujian@buaa.edu.cn

Abstract. Actuation system is one of the key subsystems to realize the flight control of aircraft. With the future development of the more electric aircraft concept, the electric actuation could replace the traditional centralized hydraulic actuation system. This paper presents a new architecture of flight control actuation system more electric aircraft. Firstly, the main components and architecture of the system are introduced. Then, the types and power levels of key electrical equipment are given. Under the environment of multi-disciplinary simulation platform AMESim, the models from a single electric actuator to the whole flight control actuation system are established. Finally, the dynamic performance of the aircraft flight control system is simulated and analyzed under the specific flight conditions. The simulation results verify the practicability of the model, and reflect the electrical characteristics of the electrical equipment of the flight control actuation system during the operation of the aircraft.

Keywords: More electric aircraft · Actuation system · Modelling and simulation · Power · AMESim

1 Introduction

In recent years, higher requirements for economy and safety have promoted the development of more electric aircraft (MEA) and all electric aircraft (AEA) [1,2]. In addition, the technologies of actuators are developing continuously and power-by-wire (PbW) actuation systems are also increasingly used on aircraft [3,4]. Electro-hydrostatic actuator (EHA) and electro-mechanical actuators (EMA) are two of the most widely used power-by-wire actuators [5–8]. It is an inevitable trend that PbW replaces traditional hydraulic servo actuator technology [9].

High performance and reliability of flight control actuation system is an important guarantee for aviation safety, which is of great significance for the development of large airliner and military aircraft [10]. The function of flight control system is to control the flight attitude of aircraft by controlling aileron,

elevator, rudder, flap, slat and spoiler. EHA and EMA are gradually applied to high-power rudder surface actuator, and also serve as backup of main flight control actuator and other key hydraulic driving rudder surface.

With the development of power electronic equipment, high-speed motor and control electronic technology, the practical design and application of more electric aircraft have been preliminarily realized in foreign civil and military aircraft such as F-22, F-35, military transport aircraft A400M, civil aircraft B787, A380, etc. all adopt multi electric or all electric technology. Multi electrochemical technology is widely used in flight control system, electric fuel system, electric environmental control system and electric anti icing system.

The speed regulating motor, servo motor and permanent magnet motor are used in electric actuators, electric fuel pumps and electric compressors. The running state of these devices, such as starting, sudden loading and unloading, will bring severe transient power impact to power supply system, generate energy feedback during braking, and affect the stability of the system. And the starting power of high-power EHA and EMA motors can reach tens of kilowatts.

The extensive application of multi electrification technology makes aircraft power system much more complex. The increasing number of airborne electric drive loads and the increasing demand for electric power put forward higher requirements for the fault tolerance and reliability of the aircraft power distribution system. Multi electrified load is the main influencing factor of power supply quality. It is important to study the characteristics of multi electrified load to improve the reliability and stability of electrical system.

In this paper, the characteristics of the aircraft flight control system of a kind of more electric aircraft are analyzed. Firstly, the architecture of the flight control system is designed and the appropriate electric actuators are configured for each rudder surface. Secondly, the power levels of the actuators are calculated. Then, according to the design results, the model of each actuator is built in the software, and the system model is integrated. Finally, the electrical performance of the designed flight control actuation system is analyzed in the software. The simulation results verify the correctness and practicability of the model. It reflects the electrical characteristics of the electrical equipment of the flight control actuation system during the operation of the aircraft.

2 More Electric Architecture Design

Two independent hydraulic systems that were used in the traditional aircraft respectively extract mechanical energy from the engine accessory gearbox by two sets of independent hydraulic pumps. The two sets of hydraulic systems work at the same time, providing hydraulic energy for the main flight surface including rudder, elevator and aileron and the flap as the lifting device to complete various flight actions. If one hydraulic system fails, the other one can provide enough pressure to drive the flight control mechanism, but the movement speed of the mechanism will be reduced accordingly. The hydraulic energy of other hydraulic users like landing gear retraction system, front wheel steering system, brake

system and hatch actuation system are also independently provided by the two sets of hydraulic systems.

In the face of the future development of aircraft multi electrification, this paper mainly takes the flight control actuation system as the research object and designs a multi electrification. The energy system architecture of the flight control actuation system is shown in Fig. 1. The electric actuator system is applied to the rudder surface of the aircraft. Five electrostatic hydraulic actuators are used in aileron, elevator and rudder. These five electro-hydrostatic actuators undertake the control tasks of all the main control surfaces of the aircraft. The flaps are driven by rotating EMA. Two independent 270 VDC airborne power supply networks are designed. Each rudder actuator is powered by two sets of independent power supply system E1 and E2.

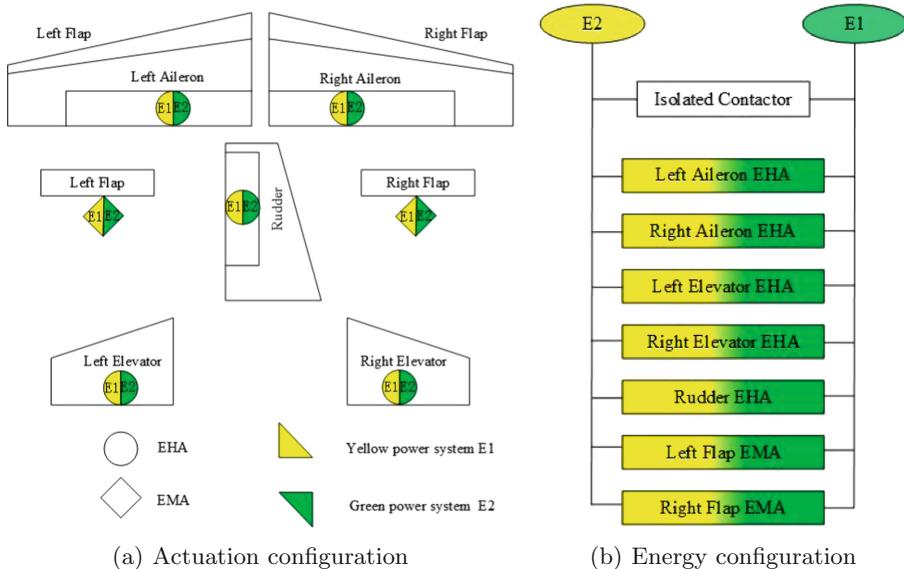


Fig. 1. Diagram of more electric aircraft flight control actuation system

3 Modeling of Electric Actuator

3.1 EHA and EMA Model

The EHA used in this paper is a closed electro-hydraulic servo system without servo valve, which can adjust the pump speed quantitatively. The integrated EHA actuator is formed by integrating the DC brushless motor, hydraulic pump, accumulator, hydraulic valve group, cylinder, sensors and power electronics of EHA.

Ignoring the influence of hydraulic valve, pipeline and accumulator, the model of closed system of pump controlled hydraulic cylinder is established as shown in Fig. 2.

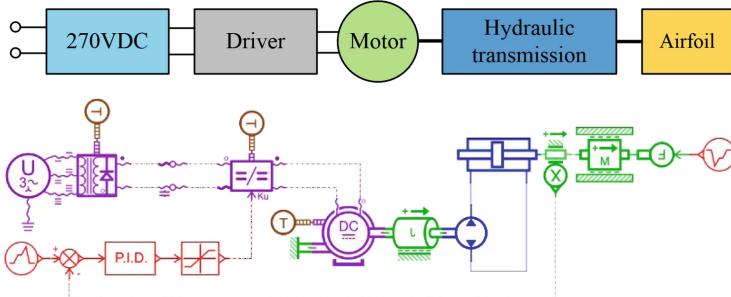


Fig. 2. Structural schematic diagram and model of EHA

EMA servo drive system consists of driver, DC brushless motor, mechanical transmission mechanism and detection device. According to the use requirements, the EMA is designed as a rotary EMA, and its model diagram is shown in Fig. 3.

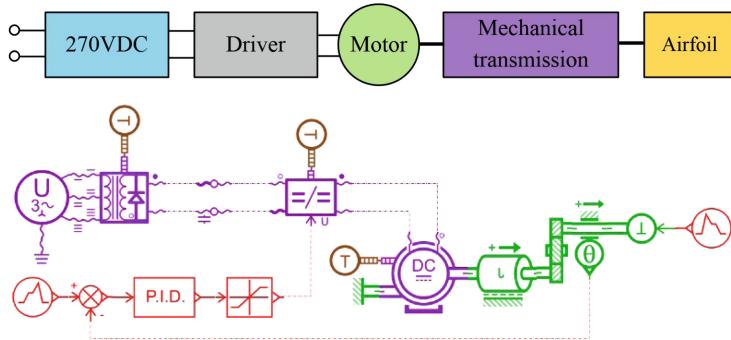


Fig. 3. Structural schematic diagram and model of EMA

3.2 System Model Integration

So far, the multi electric analysis and design of multi electric aircraft flight control actuation system are carried out, and the single electric actuator model is established.

In addition to the models of the main components mentioned above, the power system also needs to be considered. In order to make the model more

complete and convenient for simulation analysis, a 270 V power supply is used instead of the power grid. By integrating the above model into the aircraft power grid, the power consumption equipment of the flight control actuation system can obtain power from the aircraft power grid as required. The complete system model is shown in Fig. 4.

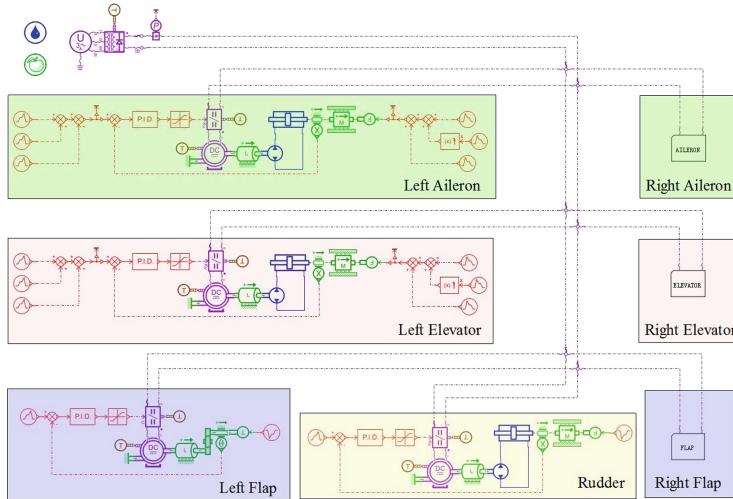


Fig. 4. Flight control actuation system model of more electric aircraft

4 Simulation Test and Analysis

4.1 System Simulation Parameters

The power of each rudder surface of the aircraft can be calculated from the ultimate load moment and the maximum speed of the rudder surface. The power estimation results of steering gear for each surface of the aircraft are shown in Table 1.

In order to analyze the dynamic characteristics of the flight control system in limit state and the impact on the power grid, the flight processes of the aircraft including combat cycle time are simulated.

The operation state of the aircraft includes five processes: take-off, level flight, combat cycle time, level flight and landing.

Combat cycle time (CCT) is an index to test the maneuverability of fighter. Four actions are defined in combat cycle time, which correspond to the limit working states of different electric actuators. The flight control surface actuator action corresponding to each process is obtained by decomposing the aircraft operation process. The load signal and position signal of each electric actuator are set in the software.

Table 1. Power estimation of surface actuator of aircraft

Parameter	Aileron	Elevator	Rudder	Flap
Torque/(N·m)	8000	11000	7000	8000
Speed/(deg/s)	± 80	± 60	± 120	± 25
Displacement/(deg)	± 21.5	± 25	± 30	(0,25)
EHA/EMA peak power/kW	8.06	10.26	7.82	2.44
EHA/EMA design power/kW	11.2	14.25	10.86	2.86

Process 1: take off. The start stage of take-off process includes elevator action, aileron linkage action and flap action.

Process 2: level flight. The take-off process is over, and the start stage of the level flight process includes elevator, aileron and flap actions.

Process 3: combat cycle time, which consists of four stages.

- 1) The aileron and elevator are both in differential motion. The aircraft completes 90° roll, and changes from level flight to side flight. At this time, the rudder acts to maintain the lift and the aircraft enters the turning state.
- 2) The aileron and elevator are actuated to complete a 180° turn in side flight.
- 3) Aileron, elevator and rudder are applied, and reverse roll 90° to level the aircraft.
- 4) In the acceleration phase of level flight, there is no rudder action in this phase.

Process 4: level flight, the end of the air combat cycle, and the aircraft flying at a constant speed.

Process 5: landing, the initial stage of landing process includes elevator action, aileron linkage action and flap action.

The above process involves sudden loading and unloading of aileron, elevator, rudder and flap, and the actions of each rudder surface are simplified. The corresponding time of each process is shown in the table. For the convenience of simulation, the take-off, level flight and landing time are reasonably shortened.

4.2 Simulation Results and Analysis

Running the simulation model, the power consumption of the whole flight control actuation system is obtained. Figure 5 shows the voltage and current changes of the whole flight control actuation system. According to the simulation results, the voltage fluctuation of the power grid is not obvious, and the current changes with the change of aircraft flight status. Figure 6 shows the change of power consumption of the whole flight control actuation system. Because the voltage fluctuation of the power grid is not obvious, the change rule of power is basically the same as that of current. The peak values of power and current are at 17.25 s, 27.16 s and 43.47 s. The corresponding aircraft operation state is that the aircraft changes from level to side, and the aircraft changes from side to level and enters

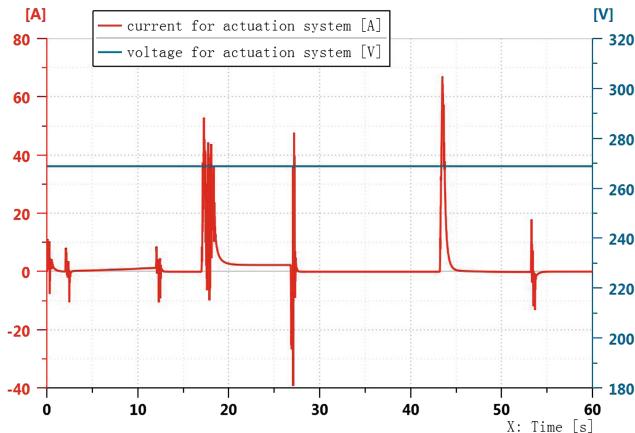


Fig. 5. Voltage and current of flight control actuation system

the landing stage. From the power and current fluctuation at 27 s, it can be seen that the power consumption of the actuator system is negative and a negative current is generated, which means that there is energy feedback in this process. Figure 7 shows the power consumption of left aileron EHA. The peak power of 17.25 s and 27.08 s corresponds to the aileron action when the aircraft changes from level to side and changes from side to level. At 17.58 s and 27.26 s, the peak power is negative, and the force vector and displacement direction of left aileron EHA output are opposite. At this time, the EHA has energy feedback.

Energy feedback and peak current will do great harm to the power grid, so the impact of airborne high-power equipment on the power grid needs to be considered. In order to ensure its normal working state, it is necessary to reduce its impact on the power grid as much as possible.

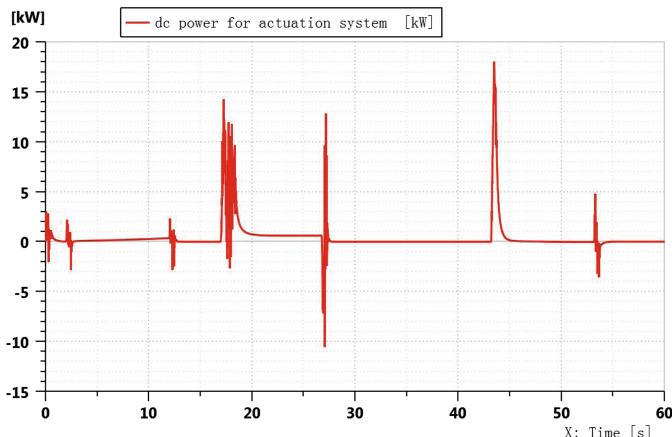


Fig. 6. Power consumption of flight control actuation system

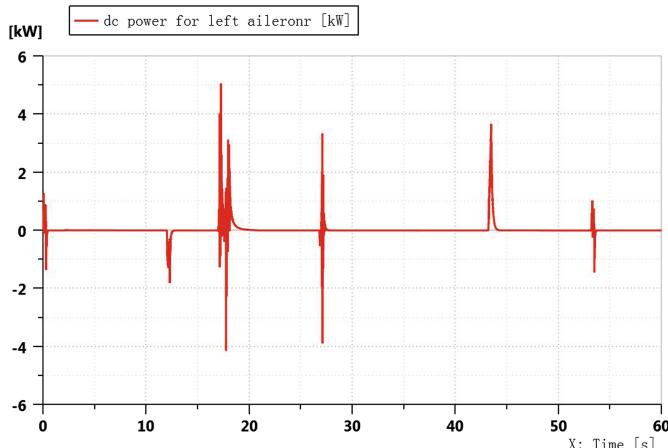


Fig. 7. Power consumption of left aileron EHA

5 Conclusion

In this paper, the architecture of the flight control actuation system of the more electric aircraft is designed firstly. After estimating the power level of electric actuators in flight control system, the power consumption model of flight control actuation system is established, simulated and analyzed by system modeling. The simulation results show that the basic model is correct and practical. Considering different engineering needs, the model can be combined with aircraft power grid model to analyze the impact of aerodynamic load on the power grid. If more specific load conditions of the electrical equipment of the flight control actuation system is considered, such as the more detailed structural model and control model of the actuator, a more practical model can be established, and other electrically applied actuation systems such as landing gear and braking, electric environmental control system and electric fuel system can also be implemented.

References

- Chakraborty, I., Trawick, D.R., Jackson, D.: Electric control surface actuator design optimization and allocation for the more electric aircraft. In: AIAA Aviation 2013 Conference (2013). <https://doi.org/10.2514/6.2013-4283>
- Chakraborty, I., Mavris, D.N., Emeneth, M.: A system and mission level analysis of electrically actuated flight control surfaces using Pacelab SysArc. In: 2014 AIAA Science and Technology Forum and Exposition (SciTech) (2014). <https://doi.org/10.2514/6.2014-0381>
- Shang, Y.X., Liu, X.C., Jiao, Z.X.: An integrated load sensing valve-controlled actuator based on power-by-wire for aircraft structural test. *Aerospace Sci. Technol.* (2018). <https://doi.org/10.1016/j.ast.2018.02.030>

4. Charrier, J.J., Kulshreshtha A.: Electric actuation for flight and engine control; evolution and current trend. In: 45th AIAA Aerospace Sciences Meeting and Exhibit (2013). <https://doi.org/10.2514/6.2007-1391>
5. Maré, J.C., Fu, J.: Review on signal-by-wire and power-by-wire actuation for more electric aircraft. Chin. J. Aeronaut. (2017). <https://doi.org/10.1016/j.cja.2017.03.013>
6. Jinker, P., Claeysen, B.: New actuators for aircraft and space applications. In: 11th International Conference on New Actuators, Bremen, Germany (2008)
7. Botten, S.L., Whitley, C.R., King, A.D.: Flight control actuation technology for next-generation all-electric aircraft. Technol. Rev. J. **8**(2), 55–68 (2000). Millennium Issue
8. Jiao, Z.X., Yu, B., Wu, S., Shang, Y.X.: An intelligent design method for actuation system architecture optimization for more electrical aircraft. Aerosp. Sci. Technol. (2019). <https://doi.org/10.1016/j.ast.2019.03.048>
9. Fu, J., Maré, J.C., Fu, Y.L.: Incremental modelling and simulation of power drive electronics and motor for flight control electro-mechanical actuators application. In: IEEE International Conference on Mechatronics and Automation, Beijing, China (2015). <https://doi.org/10.1109/ICMA.2015.7237676>
10. Fantinutto, R., Guglieri, G., Quagliotti, F.B.: Flight control system design and optimisation with a genetic algorithm. Aerosp. Sci. Technol. (2005). <https://doi.org/10.1016/j.ast.2004.09.003>



Modeling and Simulation of Redundancy Management of Electro-Hydrostatic Actuator

Jixin Song¹, Liming Yu¹, Xudong Han², and Jian Fu^{2(✉)}

¹ School of Automation Science and Electrical Engineering, Beihang University,
Beijing 100191, China

² School of Mechanical Engineering and Automation, Beihang University,
Beijing 100191, China
fujian@buaa.edu.cn

Abstract. Electro-hydrostatic actuators are widely used in the aerospace field, they are an important part of aircraft flight controls. This paper presents the redundancy management model of electro-hydrostatic actuators. First, the failure mode of the inverter is analyzed, then the structure layout of the three-redundancy inverter is proposed, and the fault monitoring is realized by a comparative monitoring method. Finally, the redundancy management model is built in Matlab/Simulink. The simulation results show that the redundancy management model can accurately monitor and effectively isolate the inverter faults, which verifies the feasibility and reliability of the model.

Keywords: EHA · Fault · Redundancy management · Inverter · Comparative monitoring

1 Introduction

With the increasing development of more electric and all electric aircraft, the airborne system is developing towards Power-By-Wire (PBW). PBW also relies on its own advantages such as high efficiency, high reliability, and easy maintenance to solve the problems of oil leakage, large volume, and large weight in traditional Hydraulic Servo Actuator (HSA). At present, Electro Hydrostatic Actuator (EHA) and Electro-Mechanical Actuator (EMA) are widely used in aerospace field. The difference between the two is the transmission method. EHA uses an electric motor to drive a hydraulic pump to provide power output, while EMA uses an electric motor to connect a mechanical gear device to provide power output [1–3].

Redundancy technology uses the same or similar redundancy units to form a highly reliable system, which plays a huge role in improving the reliability and safety of the system. The redundancy management is to complete the system status monitoring, fault detection, and redundancy switching while ensuring the normal operation of the redundancy system. With current technical conditions,

the reliability of components is limited, so redundancy technology is needed to improve the reliability potential of the system. The effect of redundancy technology is beyond the scope of merely increasing the reliability of components [4].

At present, the inverter fault diagnosis method can be divided into two situations: fault diagnosis based on traditional methods and fault diagnosis based on intelligent algorithms. Traditional methods process and analyze collected physical and chemical phenomena and detected signals and data to detect faults; while intelligent algorithms include neural grids, expert system and fuzzy algorithms, and so on. Reference [5], it conducted a simulation analysis on the fault diagnosis of the three-phase inverter in the Matlab/Simulink platform, in which the wavelet analysis method was selected to realize the fault diagnosis of the inverter, and a fourth bridge arm was added on the original three-phase inverter to realize inverter reconstruction. Reference [6], it proposed a four-switch topology of an electromagnetic double salient motor inverter. When the open circuit fault occurs in the power inverter, three bidirectional switches are used to isolate the fault circuit and form a four switch three-phase structure. The above two papers only studied the fault tolerance technology of single bridge arm.

This paper introduces a new redundancy management structure of an EHA. Compared with the single-channel EHA, the number of inverters is added to form the triple redundancy structure of the motor driver electronics. The redundancy management model of EHA is built on the Simulink platform. The faults detection and isolation of triple redundancy are completed by comparative monitoring, and the fault tolerance ability of the system is improved. Finally, typical faults are injected into the model to verify the feasibility of the redundancy model.

2 System Description

2.1 Introduction of the EHA

Study and compare various EHA redundancy schemes, and design the EHA redundancy management structure of this paper. The redundancy system in this paper mainly consists of a control unit, three inverters, a Permanent Magnet Synchronous Motor (PMSM), hydraulic system, sensors, and voting modules. The servo motor is driven by the driver, and the output torque drives the hydraulic pump to produce high-pressure oil to drive the movement of the hydraulic cylinder block to drive the load movement. The position sensor in the cylinder body feedbacks the position signal to the controller to form a closed loop of servo control and make the system move according to the predetermined input command [1]. The schematic diagram of the EHA system is shown in Fig. 1.

2.2 Introduction of the Redundancy Scheme

The reliability of single redundancy inverter is too low. In order to improve reliability, triple redundancy or quadruple redundancy is often used [4]. The

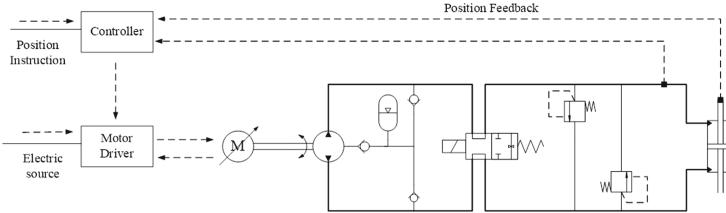


Fig. 1. Schematic diagram of EHA system

common redundancy layouts are mainly divided into a parallel system and series system. According to Eqs. 1 and 2, the reliability of parallel system is better than that of series system.

The reliability equation of the parallel system is

$$R(t) = 1 - \prod_{i=1}^n [1 - R_i(t)] \quad (1)$$

The reliability equation of the series system is

$$R(t) = \prod_{i=1}^n R_i(t) \quad (2)$$

where, $R_i(t)$ represents the reliability of each component, and all components are independent of each other.

Generally, the more redundancy, the higher the reliability can result in the cost and complexity will be higher. The double redundancy structure is simple, low cost, and low complexity, but it is challenging to monitor faults; the quadruple redundancy has high reliability and convenient fault monitoring methods, but there are complex structure, high cost, difficult maintenance. To balance the pros and cons, the triple redundancy structure can adopt the comparative monitoring to accurately monitor and effectively isolate the fault, with high reliability, medium cost and easy maintenance.

Therefore, the redundancy layout design in this paper is that the triple redundancy inverters are arranged in parallel and work in a hot standby mode to maximize the potential of system reliability.

2.3 Introduction of the Control Scheme

The control scheme of the whole system is vector control based on Proportional Integral (PI) regulator. For three-phase PMSM, vector control is more convenient, mainly including current loop control, speed loop control and SVPWM control algorithm: current loop control is used to control the speed of the motor so that it can achieve both speed regulation and speed stability; speed loop control is used to speed up the dynamic adjustment process of the system so

that the motor stator current is better close to the given current vector [7]. The outermost loop of the whole system is the position loop control, which makes the displacement of the piston rod of the hydraulic cylinder close to the given displacement output. Three groups of drivers drive the servo motor to move according to the output voltage and current of the voting module. The control scheme is shown in Fig. 2.

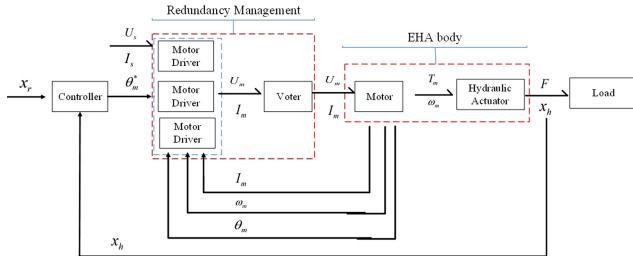


Fig. 2. Block diagram of system power transmission

3 System Redundancy Modeling

3.1 EHA Body

Electric Motor. The motor model uses a PMSM. The input current is PWM three-phase sine wave, which has excellent speed regulation performance and realizes the transformation from electrical energy to mechanical energy. The voltage equation and electromagnetic torque equation of PMSM are shown in Eqs. 3 and 4.

$$u_q = R i_q + L_q \frac{di_q}{dt} + \Psi_f n_p \omega_r \quad (3)$$

$$T_e = \frac{3}{2} n_p \Psi_f i_q = k_t i_q \quad (4)$$

Where u_q is the voltage of the q-axis stator; i_q is the current of the q-axis stator; L_q is the inductance of the q-axis stator; n_p is the number of pole pairs of the motor; ω_r is the speed of the motor rotor; Ψ_f is the equivalent flux linkage of the rotor magnetic field.

Hydraulic System. The hydraulic system model mainly includes hydraulic pumps and hydraulic cylinders, both of which are built by the transfer functions derived from their dynamic performance. The flow equations of hydraulic pumps and hydraulic cylinders, and the force balance equations of actuators are shown in Eqs. 5, 6 and 7.

$$Q_f = qn - C_{tp}p_1 + C_{ip}p_2 \quad (5)$$

$$Q_f = A \frac{dx_t}{dt} + \frac{V}{4E_y} \frac{dp_f}{dt} + C_{tg}p_f \quad (6)$$

$$Ap_f = m_t \frac{d^2x_t}{dt^2} + B_t \frac{dx_t}{dt} + K_t x_t + F_L \quad (7)$$

3.2 Controller

The controller adopts vector control scheme and uses a coordinate transformation equation to divide the stator current of the motor into excitation current and load current for separate control. The advantage is that the magnetic flow current and the torque current are completely decoupled to control the AC motor follow the control method of the DC motor. The position loop adopts Proportional(P) control, the speed loop adopts PI control, and the current loop adopts PI control [8].

3.3 Motor Driver

The control circuit and semiconductor components in the frequency conversion speed regulation system are the most prone to failure, which is the weak link of the system. The failure of the power converter in the drive system accounts for 82.5% of the total failures [5]. Therefore, it is necessary to improve the reliability of the system by designing redundancy structure.

Fault Analysis. The schematic diagram of the motor frequency control system is shown in Fig. 3.

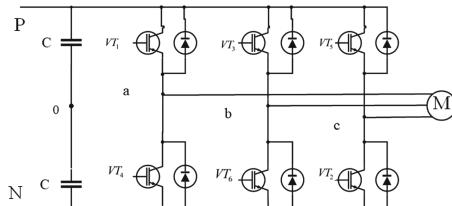


Fig. 3. Motor frequency control system

The possible faults of the motor frequency control system are as follows:

- 1) Fault 1: An open circuit occurs in a semiconductor device of a certain inverter bridge arm;

- 2) Fault 2: A short circuit occurs in a semiconductor device of a certain inverter bridge arm;
- 3) Fault 3: Two semiconductor devices of a certain inverter bridge arm are open at the same time;
- 4) Fault 4: Two semiconductor devices of a certain inverter bridge arm are short-circuited at the same time [9].

Fault 1 is mainly caused by the failure of the control circuit, leading to the failure of the semiconductor device drive; Fault 2 is mainly caused by the reverse breakdown of the power switch or the failure of the bridge insulation; the faults 3 and 4 are mostly caused by faults 1 and 2 cascading failure. The equivalent circuits of failure 1, 2 are shown in Fig. 4.

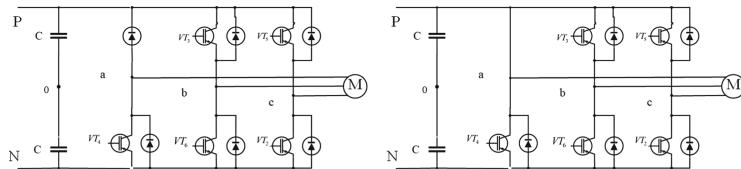


Fig. 4. Equivalent circuits of inverter open circuit fault and short circuit fault

Both of above faults will cause the inverter output phase voltage amplitude to decrease, which will affect the consequences of the motor and hydraulic actuator drive. Because the fault exists in the inverter bridge arm, it will also have a corresponding impact on other phases, making the robustness of the system significantly weakened [10].

Inverter Redundancy Management. The triple redundancy structure proposed in this paper works in a hot standby mode, and realizes the faults monitoring and faults isolation in comparative monitoring mode, the inverters redundancy layout and equivalent circuit are shown in Fig. 5.

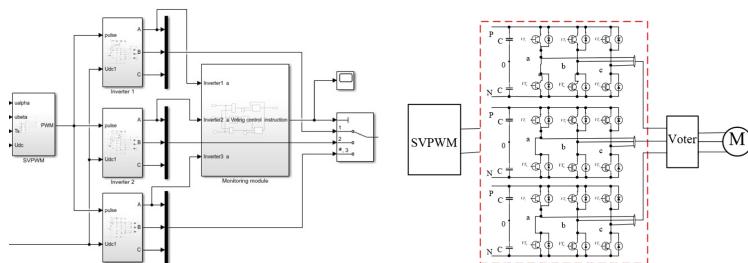


Fig. 5. Inverters redundancy layout and equivalent circuit

The monitoring voting module adopts a comparative monitoring method. When a bridge arm of the inverter fails, the three-phase voltage output by the inverter will be affected. Therefore, only one of the phase voltage waveforms can be collected as the acquisition waveform in the monitoring voting module. By default, inverter 1 is working normally, and the other two groups of inverters are in hot standby waiting for work; when inverter 1 fails, isolate the failed channel, inverter 2 starts to work, and inverter 3 is hot standby waiting for work. The judgment logic is shown in Table 1, and Fig. 6 is the redundancy monitoring voting module.

Table 1. Logic of monitoring voting

Judgment equation	Output inverter waveform
$A - B = 0$	Inverter 1
$A - B \neq 0; B - C = 0$	Inverter 2
$A - B \neq 0; B - C \neq 0; A - C = 0$	Inverter 1
$A - B \neq 0; B - C \neq 0; A - C \neq 0$	Invalid

Note: A, B and C represent phase a voltage of inverter 1, 2 and 3 respectively

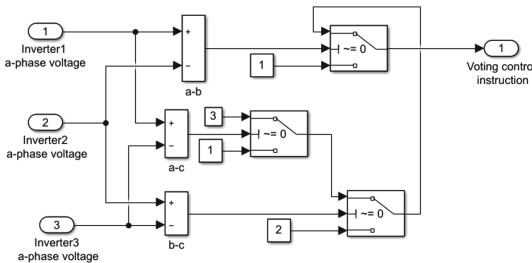


Fig. 6. Redundancy monitoring voting module

4 Simulation

The redundancy management model in this paper is built on the Simulink platform. The main modules include controller, driver, monitoring voting module, selector and EHA Body. The system model is shown in Fig. 7, and the main simulation parameters are shown in Table 2.

According to Fig. 7, system modeling is completed. The simulation time is 2 s. At 0.5 s, an open-circuit fault is injected into inverter 1, and at 1 s, a short-circuit fault is injected into inverter 2. Figure 8 shows the a-phase voltage waveforms

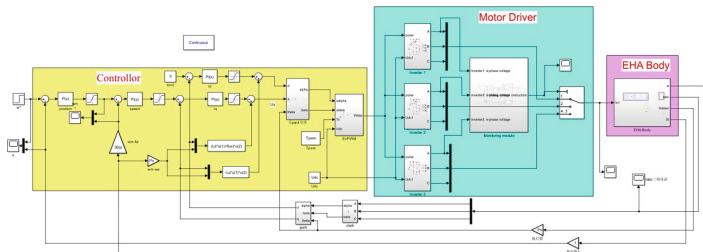


Fig. 7. EHA redundancy management model

Table 2. Main simulation parameters

Parameters	Values	Parameters	Values
DC voltage	311 V	Motor moment of inertia	$1.26110^{-4} \text{ kg}\cdot\text{m}^2$
Position input	0.05 m	Motor stator inductance	4.5610^{-2} H
Motor stator resistance	0.093Ω	Pump discharge coefficient	$1.84510^{-6} \text{ m}^3/\text{rad}$
Air load force	10 N	Acting area of piston rod	3.21310^{-3} m^2
Pole pairs of motor	3	Bulk modulus of hydraulic oil	1.54710^9 Pa

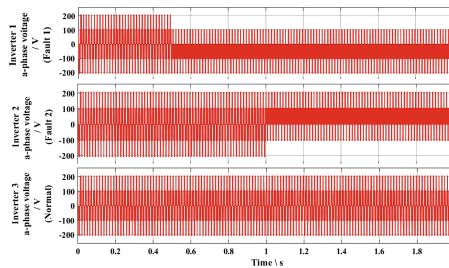


Fig. 8. Three inverters output a phase voltage waveform

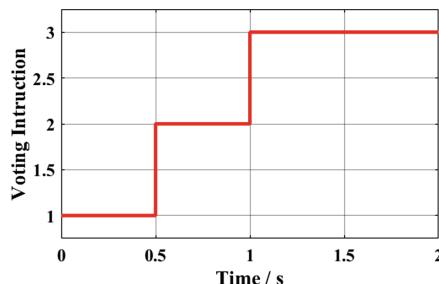


Fig. 9. The voting command

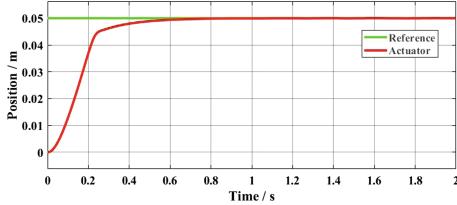


Fig. 10. The position output of the EHA system

of the three inverters; Fig. 9 shows the control input commands of the inverters redundancy voting module; Fig. 10 shows the position output of the EHA system.

It can be seen from Fig. 8 that when the inverter has an open circuit or short circuit fault, the output phase voltage amplitude decreases, which affects the drive of the motor and the hydraulic cylinder. It can be seen from Figs. 8 and 9 that when in 0.5 s, inverter 1 fails, the monitoring voting module detects the fault and isolates it, and the selector selects inverter 2 as the working inverter; when 1 s, inverter 2 fails, the monitoring voting module detects the fault and isolates it, and the selector selects inverter 3 as the working inverter. It can be seen from Fig. 10 that the whole servo movement process is almost not affected by inverters faults, and reaches a steady state quickly and smoothly. Compared with the references [5, 6], this method can realize multi-arm fault monitoring and isolation, and improve the system reliability. Therefore, the redundancy model is effective and feasible.

5 Conclusion

This paper describes and analyzes the redundancy management scheme of EHA. By comparing different redundancy layout schemes and analyzing inverter failure modes, an EHA redundancy model with triple-redundancy parallel and hot standby working mode is designed, and faults are monitored and isolated by comparative monitoring. The simulation results show that the model can effectively monitor the inverter faults and accurately isolate them. Although the faults injection is single, the monitoring voting module is designed for the common faults of inverters, which improves the system reliability and shows the practicality and feasibility of the model.

References

1. Tao, S.: Research on properties and control methods of electro hydraulic actuator in more electric aircraft. In: 3th China Aviation Science and Technology Conference Proceedings (Volume II), pp. 228–233. China Aviation Society (2017)
2. Fu, Y.L., Han, X., Yang, R., Fu, J., et al.: Review on design method of electro-hydrostatic actuator. *J. Beijing Univ. Aeronaut. Astronaut.* **43**(10), 1939–1952 (2017)

3. Liu, H., Yu, L., Zhang, Z., et al.: Fault diagnosis for SHA/EMA redundant system based on bond graph model. *J. Beijing Univ. Aeronaut. Astronaut.* **45**(04), 760–768 (2019)
4. Jin, H., Jiao, Z., Wang, S.P., Han, P.: Design and realization of reliability tri redundancy digital actor controller. *J. Beijing Univ. Aeronaut. Astronaut.* **32**(05), 548–552 (2006)
5. Song, J.S.: Research on fault diagnosis and fault tolerance control of three phase inverter. Jimei University (2012)
6. Wei, J.D., Zhou, B.: Fault tolerant strategies under open phase for doubly salient electro-magnet motor drives. In: Proceeding of International Conference on Electrical Machines and System, vol. 10, pp. 88–95 (2007)
7. Yang, Z.: Research on fault tolerant control strategy for dual three phase permanent magnet synchronous motor. Harbin Institute of Technology (2018)
8. Jia, Y.M.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003)
9. Huang, J.: Research on EMC and reliability in the motor drive system based on a three phase four leg inverter. Huazhong University of Science and Technology (2009)
10. Jia, Y.M.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000)



Global Exponential Rendezvous Control of Nonholonomic Unicycle Vehicles with Directed Communication Topology

Yixin Zhao¹, Bin Xing², Wei Huang¹, Shiyu Luo¹, Peng Li²,
and Wenjing Xie¹(✉)

¹ The School of Computer and Information Science, Southwest University,
Chongqing 400715, China
xiewenjing@swu.edu.cn

² Chongqing Innovation Center of Industrial Big-Data Co. Ltd.,
National Engineering Laboratory for Industrial Big-Data Application Technology,
Chongqing 400065, China

Abstract. This paper studies the full-state rendezvous control problem of multiple nonholonomic unicycle systems. A distributed state feedback time-varying control law is derived by exploiting model transformation, input-output feedback linearization technology, linear cooperation control theory and graph theory. Based on the carefully constructed output-like variable, the input-output feedback linearization approach overcomes the underactuated control challenge, makes the choice of controller parameters independent of initial states of vehicles, and allows the linear control theory applicable, rather than using the Lyapunov method and hence requiring the information flow undirected. Stability analysis proves that all the unicycle vehicles globally exponentially converge to a common fixed location and orientation, provided that the communication topology is directed and having a spanning tree. A numerical simulation is implemented for five nonholonomic unicycle vehicles, demonstrating the effectiveness of the proposed control scheme.

Keywords: Nonholonomic unicycle vehicles · Rendezvous control · Time-varying control

1 Introduction

Cooperative control of multiple networked agents has received much attention in recent years [1]. Typical group control objectives of agents include consensus [2], formation [3], flocking [4], rendezvous [5] and so on, which have been widely studied for many practical mechanical systems, e.g., aircrafts, ships, robots, vehicles, etc. As one research topic of this subject, cooperative control of nonholonomic unicycle vehicles has caught many researcher's eyes due to theoretical challenges and wide applications [6].

In the literature, cooperative control of unicycle vehicles mainly contains two cases: leader-following case [7] and leaderless case [8]. For the leader-following case, multiple unicycle vehicles are driven by control laws to track a prescribed reference trajectory with a desired team shape. For the latter non-leader case, unicycle vehicles are forced to rendezvous at/around a common value.

Leaderless *position* rendezvous/formation control problem of unicycle vehicles was firstly studied in the literature, e.g., see [8,9]. Leaderless *full-state* rendezvous/formation control problem of unicycle vehicles was then addressed in [5,10–12], where both position and orientation formation/rendezvous were considered. In [10], with the aid of nonsmooth Lyapunov theory and graph theory, a discontinuous control law was derived to guarantee the position and orientation rendezvous of unicycle systems, and additionally ensure the connectivity of graph, which was then extended to the case of position rendezvous with bounded control [13] and with collision avoidance [9]. Following the idea in [14], the author in [11] applied ternary and hybrid controllers that achieves the full-state consensus of unicycle undirected network with matched input disturbance. Both the position and the full-state consensus-based formation control of unicycle systems were solved in [12] by resorting to smooth time-varying persistently exciting controllers, which were given based on Lyapunov function and symmetric Laplacian matrix (i.e., undirected graph). In [5], the unicycle model was transformed into the low-order chained form, based on which backstepping technology, cascaded theory, and time-varying coordinate transformation were utilized to design a time-varying controller, guaranteeing the local generalized rendezvous of unicycle robots. On the other hand, it has been shown in the literature that the model of unicycle agents can be transformed into the canonical nonholonomic chained system, therefore the related leaderless rendezvous/consensus control of nonholonomic chained systems are also reviewed [5,15–19]. In [15], graph decomposition and input-to-state stability theory were combined to design a dynamic control law that realized semi-global output consensus control of nonholonomic chained systems. In [16], a discontinuous control law was derived based on recursive design method and the first order consensus theory, reaching the global finite-time consensus of nonholonomic systems with high-order chained structure. Based on graph-based Lyapunov function, smooth time-invariant static distributed controllers were developed in [17–19] to globally reach the asymptotic leaderless consensus control of chained systems.

Among the related works [5,10–12,15–19], only local results were obtained in [5,10,15], and the information flow was required to be undirected in [11,12,16–19] which was a litter strict. Thus, it is desirable to develop a new distributed control strategy to globally realize the rendezvous control of unicycle agents over directed communication topology with weaker requirements.

This paper investigates the full-state rendezvous control problem of networked nonholonomic unicycle systems. The model is firstly transformed into the nonholonomic chained form, and it is proved that the considered rendezvous control problem of original unicycle vehicles can be converted into the output consensus control problem of the transformed chained systems. The chained

system is divided into an integrator subsystem and a second-order nonlinear subsystem. For integrator subsystem, a time-varying distributed control law is designed for the first control input to make both the orientation rendezvous error and the control input globally exponentially converge to zero. For nonlinear subsystem, a new time-varying output-like variable is carefully constructed in such a way that the control problem can be simplified as the consensus control of some first-order integrator systems, and simultaneously that the convergence of internal state can be guaranteed. After that, a distributed control law is derived for nonlinear subsystem by combining the existing linear cooperative control theory and the input-output feedback linearization method.

The rest of this paper is arranged as follows. The problem formulation is firstly presented in Sect. 2, and the controller is then constructed in Sect. 3. Effectiveness of the proposed control strategy is illustrated in Simulation 4. Finally, the conclusion is included in Sect. 5.

2 Problem Formulation

In this paper, the considered vehicles are classic nonholonomic unicycles moving in a planar environment, and their kinematic models are represented by [5]

$$\dot{\bar{x}}_i = v_i \cos \theta_i, \quad \dot{\bar{y}}_i = v_i \sin \theta_i, \quad \dot{\theta}_i = \omega_i, \quad (1)$$

where $i \in \{1, 2, \dots, n\}$ is the index of vehicles, (\bar{x}_i, \bar{y}_i) is the position of vehicle i in the earth frame, θ_i denotes the orientation of vehicle i , and (v_i, ω_i) represent the linear and angular velocity control inputs of vehicle i , respectively.

Assume that the n vehicles with dynamics (1) constitute a communication network so that they can exchange (e.g., state) information with each other. The communication network is usually modeled as a directed graph $G(\mathcal{V}, \mathcal{A})$ [2], where $\mathcal{V} = \{1, 2, \dots, n\}$ is index/node set, and $\mathcal{A} = [a_{ij}] \in R^{n \times n}$ is adjacent matrix. Entry $a_{ij} > 0$ if system i can have access to system j , otherwise $a_{ij} = 0$. Self-loop is not allowed, that is, $a_{ii} = 0, i \in \mathcal{V}$. A graph has a spanning tree if there exists at least one node which has directed paths to all other nodes. The Laplacian matrix $L = [l_{ij}] \in R^{n \times n}$ associated with G is defined as $l_{ij} = -a_{ij}$, if $i \neq j$, and otherwise $l_{ii} = \sum_{j=1}^n a_{ij}$. Evidently, $L\mathbf{1} = 0$, where $\mathbf{1} = [1, \dots, 1]^T \in R^n$. This means that 0 is an eigenvalue of L with the right eigenvector $\mathbf{1}$. Particularly, 0 is a simple eigenvalue of L and other eigenvalues have positive real parts if G contains a spanning tree. Denote the left eigenvector of L associated with zero eigenvalue by $\bar{w} \in R^n$, that is, $\bar{w}^T L = 0$ and $\bar{w}^T \mathbf{1} = 1$.

Assumption 1. *The communication topology G of the group of unicycle agents is fixed, directed and having a spanning tree.*

The rendezvous control problem of unicycle vehicles is stated as: under Assumption 1, design a distributed control law (v_i, ω_i) such that

$$\lim_{t \rightarrow \infty} (\bar{x}_i - \bar{x}_j) = 0, \lim_{t \rightarrow \infty} (\bar{y}_i - \bar{y}_j) = 0, \lim_{t \rightarrow \infty} (\theta_i - \theta_j) = 0, \lim_{t \rightarrow \infty} (v_i, \omega_i) = 0, \forall i, j \in \mathcal{V}. \quad (2)$$

3 Controller Development

Step I: Model Transformation. Define the following transformations [17]:

$$\begin{aligned} x_{i1} &= \theta_i, x_{i2} = \bar{x}_i \sin \theta_i - \bar{y}_i \cos \theta_i, x_{i3} = \bar{x}_i \cos \theta_i + \bar{y}_i \sin \theta_i, \\ u_{i1} &= \omega_i, u_{i2} = v_i - \omega_i x_{i2}, \end{aligned} \quad (3)$$

which convert (1) into the nonholonomic chained form:

$$\dot{x}_{i1} = u_{i1}, \quad \dot{x}_{i2} = x_{i3} u_{i1}, \quad \dot{x}_{i3} = u_{i2}. \quad (4)$$

where (x_{i1}, x_{i2}, x_{i3}) and (u_{i1}, u_{i2}) are new state variables and control inputs.

The denotation ' $\xi(t) \xrightarrow{a} b$ ' will be used in the next Lemma, which means that the signal $\xi(t)$ is bounded and exponentially decays to b with the rate $a > 0$.

Lemma 1. *If*

$$x_{ik} - x_{jk} \xrightarrow{\sigma_k} 0, \quad k = 1, 3; \quad x_{i2} \xrightarrow{\sigma_2} 0, \quad u_{im} \xrightarrow{\mu_m} 0, \quad m = 1, 2 \quad (5)$$

then

$$\theta_i - \theta_j \xrightarrow{\sigma_1} 0, \quad \bar{x}_i - \bar{x}_j \xrightarrow{\sigma} 0, \quad \bar{y}_i - \bar{y}_j \xrightarrow{\sigma} 0, \quad v_i \xrightarrow{\mu} 0, \quad \omega_i \xrightarrow{\mu_1} 0, \quad (6)$$

where $(\sigma_1, \sigma_2, \sigma_3, \mu_1, \mu_2)$ are positive constants, $\underline{\sigma} = \min\{\sigma_2, \sigma_3\}$, and $\underline{\mu} = \min\{\sigma_2 + \mu_1, \mu_2\}$.

Proof. The proof is omitted due to space limitation.

Step II: Design u_{i1} such that $x_{i1} - x_{j1} \xrightarrow{\sigma_1} 0$ and $u_{i1} \xrightarrow{\mu_1} 0, \forall i, j \in \mathcal{V}$.

Lemma 2. *Under the following controller*

$$u_{i1} = - \sum_{j=1}^n a_{ij} (x_{i1} - x_{j1}) + k_1 e^{-\beta t}, \quad (7)$$

with $k_1 \neq 0$ and $\beta > 0$, if Assumption 1 holds, then the rendezvous error $x_{i1} - x_{j1}$ ($i, j \in \mathcal{V}$) and the control input u_{i1} globally exponentially converge to zero with the rate $\bar{\lambda}$, and additionally the common rendezvous point of x_{i1} is a fixed value θ^* , where

$$\theta^* = \bar{w}^T [x_{11}(0), x_{21}(0), \dots, x_{n1}(0)]^T + \frac{k_1}{\beta}, \quad \bar{\lambda} = \min \{ \operatorname{Re} [\lambda_2(L)], \beta \}, \quad (8)$$

and $\operatorname{Re} [\lambda_2(L)]$ denotes the real part of the nonzero eigenvalue of L with the smallest real part.

Proof. Define the vector $x_{*1} = [x_{11}, x_{21}, \dots, x_{n1}]^T$. Taking the derivative of x_{*1} and combining (4) and (7), we have $\dot{x}_{*1} = -Lx_{*1} + k_1 e^{-\beta t} \mathbf{1}$, which has the solution $x_{*1}(t) = e^{-Lt} x_{*1}(0) + \frac{k_1}{\beta} (1 - e^{-\beta t}) \mathbf{1}$, where $L\mathbf{1}$ is used. As the graph contains a spanning tree, it follows that $e^{-Lt} x_{*1}(0)$ tends to $\mathbf{1}\bar{w}^T x_{*1}(0)$ with the exponential decay rate $\text{Re}[\lambda_2(L)]$ [3]. By $e^{-Lt} x_{*1}(0) \xrightarrow{\text{Re}[\lambda_2(L)]} \mathbf{1}\bar{w}^T x_{*1}(0)$, one can see that the states x_{i1} globally exponentially converge to $\bar{w}^T x_{*1}(0) + \frac{k_1}{\beta} = \theta^*$ with the rate $\min\{\text{Re}[\lambda_2(L)], \beta\} = \bar{\lambda}$, that is,

$$x_{i1} \xrightarrow{\bar{\lambda}} \theta^*, \quad x_{i1} - x_{j1} \xrightarrow{\bar{\lambda}} 0. \quad (9)$$

Define the control input vector: $u_{*1} = [u_{11}, u_{21}, \dots, u_{n1}]^T = -Le^{-Lt} x_{*1}(0) + k_1 e^{-\beta t} \mathbf{1}$, which satisfies $u_{i1} \xrightarrow{\bar{\lambda}} 0$.

Step III: Design u_{i2} such that $x_{i2} \xrightarrow{\sigma_2} 0, x_{i3} - x_{j3} \xrightarrow{\sigma_3} 0, u_{i2} \xrightarrow{\mu_2} 0, \forall i, j \in \mathcal{V}$. In the following, we turn to the control of the subsystem:

$$\dot{x}_{i2} = x_{i3} u_{i1}, \quad \dot{x}_{i3} = u_{i2}. \quad (10)$$

Construct the new output-like variable for (10):

$$s_i = k_2 e^{\beta t} x_{i2} + x_{i3}, \quad (11)$$

where k_2 is a coefficient. Taking the time derivative of (11) yields:

$$\dot{s}_i = k_2 \beta e^{\beta t} x_{i2} + k_2 e^{\beta t} x_{i3} \left[- \sum_{j=1}^n a_{ij} (x_{i1} - x_{j1}) + k_1 e^{-\beta t} \right] + u_{i2} \triangleq \bar{u}_{i2}, \quad (12)$$

where \bar{u}_{i2} is the new control input replacing u_{i2} .

Lemma 3. Consider the system (9) and the controller u_{i1} in (7). Suppose that the conditions of Lemma 2 hold, and

$$0 < \beta < \min \{ \text{Re}[\lambda_2(L)], k_1 k_2 \}, \quad s_i \xrightarrow{\text{Re}[\lambda_2(L)]} s^*, \quad \forall i \in \mathcal{V},$$

then

$$x_{i2} \xrightarrow{\beta} 0, \quad x_{i3} \xrightarrow{\bar{\beta}} \frac{\beta}{\beta - k_1 k_2} s^*, \quad x_{i3} - x_{j3} \xrightarrow{\bar{\beta}} 0, \quad \forall i, j \in \mathcal{V},$$

where s^* is a constant, and $\bar{\beta} = \min\{\text{Re}[\lambda_2(L)], k_1 k_2\} - \beta$.

Proof. (a) *Model transformation.* Let us define the error $\bar{s}_i = s_i - s^*$. Corresponding to the output \bar{s}_i , the internal state \bar{x}_{i2} is defined as

$$\bar{x}_{i2} = e^{\beta t} x_{i2}, \quad (13)$$

whose derivative is calculated as

$$\dot{\tilde{x}}_{i2} = (\beta - k_1 k_2) \bar{x}_{i2} + k_1 s_i - e^{\beta t} (s_i - k_2 \bar{x}_{i2}) \sum_{j=1}^n a_{ij} (x_{i1} - x_{j1}). \quad (14)$$

To move the equilibrium of (14) to origin, let

$$\tilde{x}_{i2} = \bar{x}_{i2} + \frac{k_1}{\beta - k_1 k_2} s^*, \quad (15)$$

then (14) is converted as $\dot{\tilde{x}}_{i2} = (\beta - k_1 k_2) \tilde{x}_{i2} + k_1 \bar{s}_i - e^{\beta t} [\bar{s}_i + s^* - k_2 (\tilde{x}_{i2} - \frac{k_1}{\beta - k_1 k_2} s^*)] \sum_{j=1}^n a_{ij} (x_{i1} - x_{j1})$. Let

$$\Lambda_i(t) = e^{\beta t} \sum_{j=1}^n a_{ij} (x_{i1} - x_{j1}), \quad (16)$$

then the original subsystem (10) can be rewritten as

$$\begin{cases} \dot{\bar{s}}_i = \bar{u}_{i2}, \\ \dot{\tilde{x}}_{i2} = [(\beta - k_1 k_2) + k_2 \Lambda_i(t)] \tilde{x}_{i2} + [k_1 - \Lambda_i(t)] \bar{s}_i - \frac{\beta}{\beta - k_1 k_2} s^* \Lambda_i(t), \end{cases} \quad (17)$$

where $\Lambda_i(t)$ can be referred to as external signal governed/cascaded by the former subsystem $\dot{x}_{i1} = u_{i1} = - \sum_{j=1}^n a_{ij} (x_{i1} - x_{j1})$, \bar{s}_i is viewed as output variable, and \tilde{x}_{i2} is internal state.

(b) *Convergence of $\Lambda_i(t)$.* Under the controller u_{i1} in (7), we have known that $Lx_{*1}(t) \xrightarrow{\text{Re}[\lambda_2(L)]} 0$ and hence $L_i x_{*1}(t) = \sum_{j=1}^n a_{ij} (x_{i1} - x_{j1}) \xrightarrow{\text{Re}[\lambda_2(L)]} 0$, where L_i denotes the i th row of L . As $\beta < \text{Re}[\lambda_2(L)]$, $\Lambda_i(t)$ in (16) satisfies $\Lambda_i(t) \xrightarrow{\text{Re}[\lambda_2(L)] - \beta} 0$, guaranteeing system (17) and therefore (14) well defined.

(c) *Stability of zero dynamics of (17).* It is known that the output variable \bar{s}_i in (17) is exponentially convergent to zero. When $\bar{s}_i = 0$, the zero dynamics is obtained as $\dot{\tilde{x}}_{i2} = [(\beta - k_1 k_2) + k_2 \Lambda_i(t)] \tilde{x}_{i2} - \frac{\beta}{\beta - k_1 k_2} s^* \Lambda_i(t)$, which is a stable linear system disturbed by bounded convergent terms. The convergence rate of $\Lambda_i(t)$ is $\text{Re}[\lambda_2(L)] - \beta$. It follows from the theory on linear systems that \tilde{x}_{i2} vanishes with the exponential rate $\min\{k_1 k_2 - \beta, \text{Re}[\lambda_2(L)] - \beta\} = \bar{\beta}$, that is, $\tilde{x}_{i2} \xrightarrow{\bar{\beta}} 0$. Moreover, if considering $\bar{s}_i \xrightarrow{\text{Re}[\lambda_2(L)]} 0$ for \tilde{x}_{i2} , $\tilde{x}_{i2} \xrightarrow{\bar{\beta}} 0$ can also be concluded.

(c) *Final results.* By (11), (13) and (15), one gets $x_{i2} \xrightarrow{\beta} 0, x_{i3} \xrightarrow{\bar{\beta}} 0, x_{i3} - x_{j3} \xrightarrow{\bar{\beta}} 0$. The proof is completed.

Lemma 3 tells that, the remaining control task is reduced to the consensus control problem of the first-order integrators $\dot{s}_i = \bar{u}_{i2}$. For such s_i -subsystem, we design [2]

$$\bar{u}_{i2} = - \sum_{j=1}^n a_{ij} (s_i - s_j). \quad (18)$$

Lemma 4. Under Assumption 1 and (18), s_i globally exponentially converges to $s^* = \bar{w}_l^T [s_1(0), s_2(0), \dots, s_n(0)]^T$ with the rate $\text{Re}[\lambda_2(L)]$.

Proof. Substituting (18) into $\dot{s}_i = \bar{u}_{i2}$, we have $\dot{s}_i = -\sum_{j=1}^n a_{ij} (s_i - s_j)$. Let $s = [s_1, s_2, \dots, s_n]^T$, and the above system can be rewritten as $\dot{s} = -Ls$. For such system, the results in Lemma 4 have been proved in the literature [20].

Theorem 1. Suppose that Assumption 1 holds and $0 < \beta < \min\{\text{Re}[\lambda_2(L)], k_1 k_2\}$, then the proposed control law (7) and (18) guarantees the group of unicycle vehicles globally exponentially achieve the rendezvous control objective (2), more precisely, they globally exponentially rendezvous at the fixed location (x^*, y^*) and the fixed orientation θ^* , and the velocity control inputs (v_i, ω_i) exponentially converge to zero, where

$$x^* = \frac{\beta}{\beta - k_1 k_2} \bar{w}^T s(0) \cos \theta^*, \quad y^* = \frac{\beta}{\beta - k_1 k_2} \bar{w}^T s(0) \sin \theta^* \quad (19)$$

Proof. As Assumption 1 holds and $0 < \beta < \min\{\text{Re}[\lambda_2(L)], k_1 k_2\}$, it can be concluded from Lemmas 1–4 that $x_{i1} \xrightarrow{\sigma_1} \theta^*$, $x_{i2} \xrightarrow{\sigma_2} 0$, $x_{i3} \xrightarrow{\sigma_3} \frac{\beta}{\beta - k_1 k_2} \bar{w}^T s(0)$. One has $\bar{x}_i \xrightarrow{\sigma} \frac{\beta}{\beta - k_1 k_2} \bar{w}^T s(0) \cos \theta^*$, $\bar{y}_i \xrightarrow{\sigma} \frac{\beta}{\beta - k_1 k_2} \bar{w}^T s(0) \sin \theta^*$, $\theta_i \xrightarrow{\sigma_1} \theta^*$.

4 Simulation

In this section, we present a several simulation example to illustrate the effectiveness of the proposed control scheme. Suppose that there are five agents connected by the directed ring topology with $a_{ij} = 0.5$ or 0, leading to $\text{Re}[\lambda_2(L)] = 0.3455$, $\bar{w} = 0.2[1, 1, 1, 1, 1]^T$.

Initial positions and headings of unicycle vehicles are set as $(\bar{x}_1, \bar{y}_1, \theta_1)(0) = (-3, -7, -\pi)$, $(\bar{x}_2, \bar{y}_2, \theta_2)(0) = (-5, 3, -\pi)$, $(\bar{x}_3, \bar{y}_3, \theta_3)(0) = (5, 5, -\pi)$, $(\bar{x}_4, \bar{y}_4, \theta_4)(0) = (5, 4, \pi/2)$, $(\bar{x}_5, \bar{y}_5, \theta_5)(0) = (1, -7, -\pi)$.

To make controller coefficients meet $0 < \beta < \min\{\text{Re}[\lambda_2(L)], k_1 k_2\}$, let $k_1 = 0.7$, $k_2 = 0.5$, $\beta = 0.2$. Under the selected initial states and control coefficients, the rendezvous location and the orientation can be computed as $(x^*, y^*) = (-0.4622, -1.6706)$, $\theta^* = 1.3009$.

Carrying out the control strategy to each nonholonomic unicycle vehicle, the simulation results are obtained in Fig. 1. It can be observed from Fig. 1 that all the five vehicles globally exponentially rendezvous at the common location (x^*, y^*) , both the position and orientation rendezvous are realized. The simulation figures demonstrate the effectiveness of the results in Theorem 1.

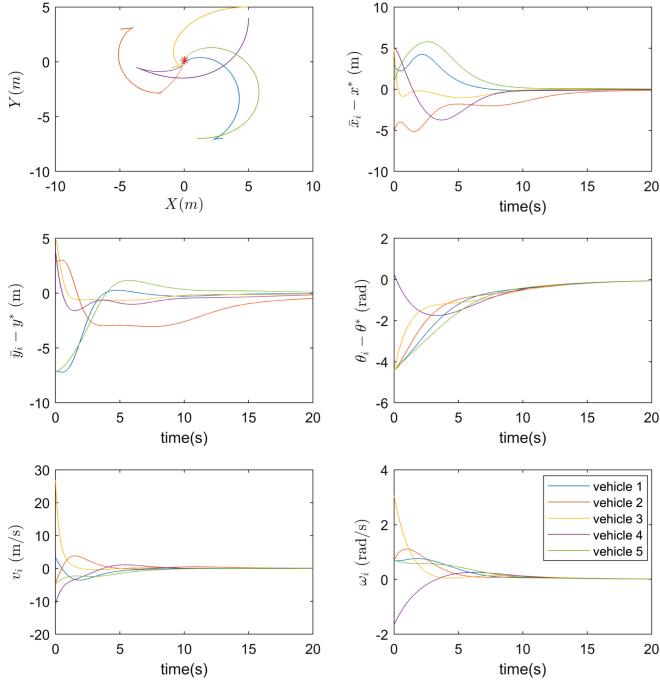


Fig. 1. Simulation results of the proposed control scheme.

5 Conclusion

In this paper, the full-state rendezvous control problem of unicycle vehicles is solved by proposing a distributed time-varying control law. Supposing that the communication graph contains a spanning tree, the proposed controller is capable of globally exponentially driving the unicycle vehicles to rendezvous at a common fixed position and orientation. The controller design is based on the combination of input-output feedback linearization technology, linear cooperation control theory, time-varying control skill and graph theory. Extending the proposed control scheme to leader-following case may be one future research work.

Funding. This work was supported by Fundamental Research Funds for the Central Universities (No. XDKJ2020C036), and Development and Application Demonstration of Intelligent Light Pole Internet of Things System of Chongqing Science and Technology Commission (No. cstc2017zdcy-zdyfX0054).

References

1. Zhu, B., Xie, L., Han, D., Meng, X., Teo, R.: A survey on recent progress in control of swarm systems. *Sci. China Inf. Sci.* **60**(7), 070201 (2017)
2. Olfati-Saber, R., Murray, R.M.: Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control* **49**(9), 1520–1533 (2004)
3. Dong, W., Farrell, J.A.: Formation control of multiple underactuated surface vessels. *IET Control Theory Appl.* **2**(12), 1077–1085 (2008)
4. Zou, Y., Wen, C., Guan, M.: Distributed adaptive control for distance-based formation and flocking control of multi-agent systems. *IET Control Theory Appl.* **13**(6), 878–885 (2019)
5. Cao, K., Jiang, B., Yue, D.: Rendezvous of multiple nonholonomic unicycles-based on backstepping. *Int. J. Control.* **91**(6), 1271–1283 (2018)
6. Huang, J., Chen, J., Fang, H., Dou, L.: An overview of recent progress in high-order nonholonomic chained system control and distributed coordination. *J. Control Decis.* **2**(1), 64–85 (2015)
7. Defoort, M., Demesure, G., Zuo, Z., Polyakov, A., Djemai, M.: Fixed-time stabilisation and consensus of non-holonomic systems. *IET Control Theory Appl.* **10**(18), 2497–2505 (2016)
8. Xie, W., Ma, B.: Position centroid rendezvous and centroid formation of multiple unicycle agents. *IET Control Theory Appl.* **8**(17), 2055–2061 (2014)
9. Li, P., Shengyuan, X., Chen, W., Wei, Y., Zhang, Z.: A connectivity preserving rendezvous for unicycle agents with heterogenous input disturbances. *J. Franklin Inst.* **355**(10), 4248–4267 (2018)
10. Dimarogonas, D.V., Kyriakopoulos, K.J.: On the rendezvous problem for multiple nonholonomic agents. *IEEE Trans. Autom. Control* **52**(5), 916–922 (2007)
11. Jafarian, M.: Robust consensus of unicycles using ternary and hybrid controllers. *Int. J. Robust Nonlinear Control* **27**(17), 4013–4034 (2017)
12. Maghenem, M., Bautista, A., Nuño, E., Loría, A., Panteley, E.: Consensus of multi-agent systems with nonholonomic restrictions via Lyapunov's direct method. *IEEE Control Syst. Lett.* **3**(2), 344–349 (2018)
13. Ajorlou, A., Aghdam, A.G.: Connectivity preservation in nonholonomic multi-agent systems: a bounded distributed control strategy. *IEEE Trans. Autom. Control* **58**(9), 2366–2371 (2013)
14. Yu, J., LaValle, S.M., Liberzon, D.: Rendezvous without coordinates. *IEEE Trans. Autom. Control* **57**(2), 421–434 (2011)
15. Yaojin, X., Tian, Y.-P., Chen, Y.Q.: Output consensus for multiple non-holonomic systems under directed communication topology. *Int. J. Syst. Sci.* **46**(3), 451–463 (2015)
16. Haibo, D., Wen, G., Cheng, Y., He, Y., Jia, R.: Distributed finite-time cooperative control of multiple high-order nonholonomic mobile robots. *IEEE Trans. Neural Netw. Learn. Syst.* **28**(12), 2998–3006 (2016)
17. Zhai, G., Takeda, J., Imae, J., Kobayashi, T.: Towards consensus in networked non-holonomic systems. *IET Control Theory Appl.* **4**(10), 2212–2218 (2010)
18. Xie, W., Ma, B.: Smooth time-invariant control for leaderless consensus of networked nonholonomic systems. *Int. J. Adv. Robot. Syst.* **14**(6), 1–9 (2017)

19. Xie, W., Ma, B., Huang, W., Zhao, Y., Fernando, T.: Asymptotic leaderless consensus control of nonholonomic chained systems. In: 2019 Chinese Control Conference (CCC), pp. 715–720. IEEE (2019)
20. Xie, W., Ma, B., Fernando, T., Iu, H.H.-C.: A new formation control of multiple underactuated surface vessels. *Int. J. Control.* **91**(5), 1011–1022 (2018)



Spectral Criterion for Stability of Mean-Field Stochastic Periodic Systems

Yuechen Cui and Hongji Ma^(✉)

College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China
ma_math@163.com

Abstract. This paper addresses the problems of stability and stabilizability for discrete-time mean-field stochastic systems with periodic coefficients. In terms of an orthogonal decomposition of system state, a monodromy operator is introduced for the considered dynamics. Based on the spectral distribution of monodromy operator, operator-spectral criteria are presented for asymptotic stability and weak stability, respectively. Further, by a group of coupled difference linear matrix inequalities (LMIs), Lyapunov-type stability criteria are derived for mean-field stochastic periodic systems. In addition, sufficient conditions are also obtained for both asymptotic stabilizability and regional stabilizability in terms of difference LMIs.

Keywords: Mean-field systems · Monodromy operator · Spectrum · Stability

1 Introduction

Mean-field stochastic systems, which are also referred to as McKEAN-Vlasov stochastic equations, have received considerable research attention in the past decades. Different from the classic stochastic systems, the expectation terms of system state and control input are taken into the evolution of dynamical models to characterize the negligible effect resulting from mutual interactions among individual agents/or multiple decision-makers. The study of mean-field systems dates back at least to [1] and has been applied in many areas ranging from finance, economics, smart grids, to social science and control/game theory. For the mentioned topics, interested readers may refer to [2–6] as well as the references therein.

Stability is one of the essential performances of a dynamic system and plays a key role to the exploration of control design. For stochastic continuous/discrete systems without mean-field term, great many results have been contributed to the stability analysis in different senses, such as stochastic stability in probability [7], asymptotic mean square stability [8], and exponential mean square stability [9]. Recently, a stability criterion has been provided for mean-field stochastic discrete-time linear systems in [10]. Up to now, almost all the available works concerned

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 562–570, 2021.

https://doi.org/10.1007/978-981-15-8458-9_60

with the mean-field stochastic systems are confined within constant coefficients, while few result is found for mean-field systems with time-varying coefficients. This paper devotes to investigating the stability of discrete-time mean-field stochastic linear systems with periodic coefficients and state/control-dependent noise. On one hand, periodicity is one of the most common phenomena in the real world, especially social and economical fields. So, it is necessary to study the mean-field systems with periodic coefficients. On the other hand, time-varying coefficients and the heterogeneity of different subsystems give rise to specific technical challenges. Hence, this work is never an trivial extension of the research on constant case. To some extent, this study aims to extend the results of [11] to the mean-field models. In order to work out the monodromy operator for the considered dynamics, an orthogonal decomposition of the system state is utilized. By means of spectral distribution of the introduced monodromy operator, we obtain spectral criteria of asymptotic mean square stability and weak stability. Besides, Lyapunov-type criteria are also presented by a group of coupled difference LMIs with periodic coefficients, which further lead to LMI-based conditions that guarantee the stabilizability of considered systems in different senses.

Notations. $\mathcal{C}(\mathcal{R})$: the set of all complex (real) numbers; $\mathcal{C}^n(\mathcal{R}^n)$: n -dimensional complex (real) vector space with the usual Euclidean norm $\|\cdot\|$; $\mathcal{C}^{n \times m}(\mathcal{R}^{n \times m})$: the space of all $n \times m$ complex (real) matrices with the operator norm $\|\cdot\|_2$; $\mathcal{S}_n(\mathcal{S}_n^+)$: the set of all $n \times n$ (positive definite) symmetric matrices with possible complex entries; $A > 0 (\geq 0)$: A is a symmetric positive (semi-) definite matrix; A' : the transpose of a matrix (vector) A ; $A^{-1}(A^+)$: the inverse (Moore-Penrose pseudoinverse) of a matrix A ; I_n : the $n \times n$ identity matrix; $\mathbb{Z}_+ := \{0, 1, \dots\}$ and $\mathbb{Z}_{1+} := \{1, 2, \dots\}$.

2 Main Results

On a complete probability space (Ω, \mathcal{F}, P) , we consider the following discrete-time stochastic mean-field systems with multiplicative noise and periodic coefficients:

$$\begin{cases} x_{t+1} = A_t x_t + \bar{A}_t E x_t + B_t u_t + \bar{B}_t E u_t \\ \quad + [C_t x_t + \bar{C}_t E x_t + D_t u_t + \bar{D}_t E u_t] w_t, \\ x_{t_0} = \xi \in \mathcal{R}^n, \quad t \in \mathbb{Z}_+, \end{cases} \quad (1)$$

where the involved coefficients are all real matrix-valued sequences with a common period $\theta \in N$ and have suitable dimensions, e.g., $A_{t+\theta} = A_t$, and $A_t \in \mathcal{R}^{n \times n}$. In (1), $x_t \in \mathcal{R}^n$ is the system state and $u_t \in \mathcal{R}^{n_u}$ stands for the control input. $\{w_t \in \mathcal{R} | t \in \mathbb{Z}_+\}$ is a wide-sense stationary process with second-order moment and the natural filtration $\mathcal{G}_t = \sigma\{w_s | s = 0, \dots, t\}$. Moreover, $E w_t = 0$ and $E w_t w_s = \delta_{ts}$, where $\delta_{(\cdot)}$ is the Kronecker functional. The initial state ξ is an n -dimensional random variable independent of the stochastic process $\{w_t \in \mathcal{R} | t \in \mathbb{Z}_+\}$ and $E\|\xi\|^2 < \infty$. Denote by \mathcal{F}_t the minimal σ -algebra consisting of \mathcal{G}_t and $\sigma(\xi)$, i.e., $\mathcal{F}_t = \mathcal{F}_t \vee \sigma(\xi)$.

First of all, we give some basic definitions on the stability and stabilizability.

Definition 1. The following discrete-time mean-field stochastic system:

$$\begin{cases} x_{t+1} = A_t x_t + \bar{A}_t E x_t + (C_t x_t + \bar{C}_t E x_t) w_t, \\ x_{t_0} = \xi \in \mathcal{R}^n, \quad t \in \mathbb{Z}_+, \end{cases} \quad (2)$$

(or $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ for short) is called asymptotically mean square stable (AMSS), if the solution x_t satisfies $\lim_{t \rightarrow +\infty} E \|x_t\|^2 = 0$ for any initial value $x_{t_0} = \xi \in \mathcal{R}^n$.

Definition 2. The discrete-time mean-field stochastic system (1) is said to be stabilizable (in the mean square sense) if there exists a pair of θ -periodic matrices $(K_t, \bar{K}_t) \in \mathcal{R}^{n_u \times n} \times \mathcal{R}^{n_u \times n}$ such that the corresponding closed-loop system:

$$\begin{cases} x_{t+1} = (A_t + B_t K_t) x_t + [\bar{A}_t + B_t \bar{K}_t + \bar{B}_t (K_t + \bar{K}_t)] E x_t \\ \quad + \{(C_t + D_t K_t) x_t + [\bar{C}_t + D_t \bar{K}_t + \bar{D}_t (K_t + \bar{K}_t)] E x_t\} w_t, \\ x_{t_0} = \xi \in \mathcal{R}^n, \quad t \in \mathbb{Z}_+, \end{cases} \quad (3)$$

is AMSS for any $\xi \in \mathcal{R}^n$. Moreover, $u_t = K_t x_t + \bar{K}_t E x_t$ is called a stabilizing feedback controller.

Definition 3. $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is called weakly stable (in the mean square sense) if for any $\varepsilon > 0$ and $t_0 \geq 0$, there always exists a $\rho > 0$ such that $E \|x_t\|^2 < \varepsilon$ whenever $E \|x_{t_0}\|^2 < \rho$ and $t \geq t_0 \geq 0$.

To investigate the stability of system (2), we take expectations on both sides of system (2) and get

$$\begin{cases} E x_{t+1} = (A_t + \bar{A}_t) E x_t, \\ E x_{t_0} = E \xi \in \mathcal{R}^n, \quad t \in \mathbb{Z}_+. \end{cases} \quad (4)$$

Subtracting (4) from (2), we arrive at the following equation immediately:

$$\begin{cases} x_{t+1} - E x_{t+1} = A_t(x_t - E x_t) + [C_t(x_t - E x_t) + (C_t + \bar{C}_t)E x_t] w_t, \\ x_{t_0} - E x_{t_0} = \xi - E \xi \in \mathcal{R}^n, \quad t \in \mathbb{Z}_+. \end{cases} \quad (5)$$

From the orthogonal decomposition of $x_t = E x_t + (x_t - E x_t)$, it can be seen that $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is AMSS if and only if (4) and (5) are AMSS simultaneously. Bearing this observation in mind, we introduce a linear operator \mathcal{L}_t associated to (4) and (5) as follows:

$$\mathcal{L}_t(X) = \mathbf{A}_t X \mathbf{A}'_t + \mathbf{C}_t X \mathbf{C}'_t + \hat{\mathbf{C}}_t X \hat{\mathbf{C}}'_t, \quad X \in \mathbb{S}_{2n}, \quad (6)$$

where \mathbb{S}_{2n} is a real subset of \mathcal{S}_{2n} :

$$\mathbb{S}_{2n} = \left\{ \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \middle| X_1, X_2 \in \mathcal{S}_n \cap \mathcal{R}^{n \times n} \right\}, \quad (7)$$

and the coefficients are given by

$$\mathbf{A}_t = \begin{bmatrix} A_t & 0 \\ 0 & A_t + \bar{A}_t \end{bmatrix}, \quad \mathbf{C}_t = \begin{bmatrix} C_t & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{\mathbf{C}}_t = \begin{bmatrix} 0 & C_t + \bar{C}_t \\ 0 & 0 \end{bmatrix}. \quad (8)$$

Hence, $\mathcal{L}_{t+\theta} = \mathcal{L}_t$. Moreover, it is easy to verify that \mathcal{L}_t is a linear positive operator, i.e., if $X \in \mathbb{S}_{2n}^+$, then $\mathcal{L}_t(X) \in \mathbb{S}_{2n}^+$, where \mathbb{S}_{2n}^+ denotes the set consisting of all nonnegative definite matrices of \mathbb{S}_{2n} .

Given $X, Y \in \mathbb{S}_{2n}$, define the inner product $\langle X, Y \rangle = \text{Tr}(X^*Y)$. Therefore, the adjoint operator of \mathcal{L}_t is expressed by

$$\mathcal{L}_t^*(X) = \mathbf{A}'_t X \mathbf{A}_t + \mathbf{C}'_t X \mathbf{C}_t + \hat{\mathbf{C}}'_t X \hat{\mathbf{C}}_t, \quad X \in \mathbb{S}_{2n}, \quad (9)$$

where the coefficients are the same as (8). Clearly, \mathcal{L}_t^* is also a linear θ -periodic positive operator.

In the sequel, the linear evolution operator of \mathcal{L}_t is denoted by $\mathcal{T}_{t,s} = \mathcal{L}_{t-1}\mathcal{L}_{t-2}\cdots\mathcal{L}_s$ for $t > s \geq 0$ and $\mathcal{T}_{t,t} = \mathcal{I}$ (the identity operator). Moreover, the spectral set of $\mathcal{T}_{t,s}$ is denoted by $\Lambda(\mathcal{T}_{t,s}) = \{\lambda | \mathcal{T}_{t,s}(X) = \lambda X, \lambda \in \mathcal{C}, 0 \neq X \in \mathcal{S}_{2n}\}$. From the periodicity of \mathcal{L}_t , we deduce that $\mathcal{T}_{t+\theta,s+\theta} = \mathcal{T}_{t,s}$. Particularly, we will write $\mathcal{T}_{t+\theta,t}$ as \mathcal{T}_t^θ for short.

Definition 4. \mathcal{T}_t^θ is called the monodromy operator of $[A_t, \bar{A}_t; C_t, \bar{C}_t]$.

In the subsequent analysis, we will need the following two linear operators.

$$\begin{aligned} \psi : Y \in \mathcal{C}^{2n \times 2n} &\mapsto (y_{11} \cdots y_{1,2n} \ y_{21} \cdots y_{2,2n} \cdots \ y_{2n,1} \cdots y_{2n,2n})' \in \mathcal{C}^{4n^2}, \\ \varphi : X \in \mathcal{S}_{2n} &\mapsto (x_{11} \cdots x_{1,2n} \ x_{22} \cdots x_{2,2n} \cdots \\ &\quad \cdots \ x_{2n-1,2n-1} \ x_{2n-1,2n} \ x_{2n,2n})' \in \mathcal{C}^{n(2n+1)}. \end{aligned}$$

According to [11] and [12], the linear operators ψ and φ are invertible in the sense that for any pair $(X, Y) \in \mathcal{S}_{2n} \times \mathcal{C}^{2n \times 2n}$, there always holds

$$(\varphi^{-1}(\varphi(X)), \psi^{-1}(\psi(Y))) = (X, Y) \in \mathcal{S}_{2n} \times \mathcal{C}^{2n \times 2n}.$$

Moreover, for any $X \in \mathcal{S}_{2n}$,

$$\psi(X) = G\varphi(X), \quad (10)$$

where $G \in \mathcal{R}^{4n^2 \times n(2n+1)}$ is of full-column rank with elements g_{kl} specified by

$$g_{kl} = \begin{cases} 1, & \text{if } (k, l) = \left(2n(i-1)+j, (2n-\frac{\check{m}}{2})(\check{m}-1)+\hat{m}\right) \\ 0, & \text{otherwise,} \end{cases}$$

where $\hat{m} = \max\{i, j\}$ and $\check{m} = \min\{i, j\}$ for $i, j = 1, \dots, 2n$.

By use of Kronecker product, (6) can be represented as:

$$\psi(\mathcal{L}_t(X)) = (\mathbf{A}_t \otimes \mathbf{A}_t + \mathbf{C}_t \otimes \mathbf{C}_t + \hat{\mathbf{C}}_t \otimes \hat{\mathbf{C}}_t)\psi(X) \quad (11)$$

and

$$\varphi(\mathcal{L}_t(X)) = L_t\varphi(X). \quad (12)$$

In (12), L_t is the induced matrix of \mathcal{L}_t with the following expression:

$$L_t = G^+(\mathbf{A}_t \otimes \mathbf{A}_t + \mathbf{C}_t \otimes \mathbf{C}_t + \hat{\mathbf{C}}_t \otimes \hat{\mathbf{C}}_t)G, \quad (13)$$

where $G^+ = (G'G)^{-1}G'$. Hence, $\{L_t\}_{t \in \mathbb{Z}_+}$ is a θ -periodic matrix sequence.

By induction on (12), the induced matrix of $T_{t,s}$ is given by

$$\Phi_{t,s} = L_{t-1}L_{t-2}\cdots L_s, \quad t > s \geq 0.$$

Denote by Φ_t^θ the induced matrix of the monodromy operator T_t^θ , then we have $\Phi_{t+\theta}^\theta = \Phi_t^\theta$.

Theorem 1. $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is AMSS if and only if $\Lambda(T_t^\theta) \subset \mathcal{B}(0, 1)$, where $\mathcal{B}(0, 1) = \{\lambda | \lambda \in \mathcal{C}, |\lambda| \leq 1\}$; $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is weakly stable if and only if $\Lambda(T_t^\theta) \subset \mathcal{B}(0, 1) := \{\lambda | \lambda \in \mathcal{C}, |\lambda| \leq 1\}$, and any $\lambda_i \in \Lambda(T_t^\theta)$ with $|\lambda_i| = 1$ must be the simple root of the monodromy operator.

Proof. Let $X_t = E(x_t - Ex_t)(x_t - Ex_t)'$ and $\bar{X}_t = (Ex_t)(Ex_t)'$, then the decomposition of $x_t = Ex_t + (x_t - Ex_t)$ implies that $\lim_{t \rightarrow \infty} E\|x_t\|^2 = 0$ if and only if $\lim_{t \rightarrow \infty} X_t = 0$ and $\lim_{t \rightarrow \infty} \bar{X}_t = 0$. By combining (4) with (5), it can be computed that X_t and \bar{X}_t satisfy

$$\begin{cases} X_{t+1} = A_t X_t A_t' + C_t X_t C_t' + (C_t + \bar{C}_t) \bar{X}_t (C_t + \bar{C}_t)', \\ X_{t_0} = E(\xi - E\xi)(\xi - E\xi)', \quad t \in \mathbb{Z}_+, \end{cases} \quad (14)$$

and

$$\begin{cases} \bar{X}_{t+1} = (A_t + \bar{A}_t) \bar{X}_t (A_t + \bar{A}_t)', \\ \bar{X}_{t_0} = (E\xi)(E\xi)', \quad t \in \mathbb{Z}_+, \end{cases} \quad (15)$$

respectively. Let

$$\mathbf{X}_t = \begin{bmatrix} X_t & 0 \\ 0 & \bar{X}_t \end{bmatrix}, \quad (16)$$

then (14) and (15) yield that \mathbf{X} satisfies the following difference equation:

$$\begin{cases} \mathbf{X}_{t+1} = \mathcal{L}_t(\mathbf{X}_t), \quad t \in \mathbb{Z}_+, \\ \mathbf{X}_{t_0} = \begin{bmatrix} E(\xi - E\xi)(\xi - E\xi)' & 0 \\ 0 & (E\xi)(E\xi)' \end{bmatrix}. \end{cases} \quad (17)$$

By the formula (12), (17) can be equivalently represented as the following discrete-time θ -periodic linear system:

$$\begin{cases} \varphi(\mathbf{X}_{t+1}) = L_t \varphi(\mathbf{X}_t), \quad t \in \mathbb{Z}_+, \\ \varphi(\mathbf{X}_{t_0}) = \varphi \left(\begin{bmatrix} E(\xi - E\xi)(\xi - E\xi)' & 0 \\ 0 & (E\xi)(E\xi)' \end{bmatrix} \right), \end{cases} \quad (18)$$

where $L_t \in \mathcal{R}^{n(2n+1) \times n(2n+1)}$ is determined by (13). Reminding the structure of \mathbf{X}_t , we have

$$\lim_{t \rightarrow \infty} E\|x_t\|^2 = 0 \Leftrightarrow \lim_{t \rightarrow \infty} \mathbf{X}_t = 0 \Leftrightarrow \lim_{t \rightarrow \infty} \varphi(\mathbf{X}_t) = 0. \quad (19)$$

By Proposition 3.3 [13], (18) is stable if and only if the spectral set of associated monodromy matrix Φ_t^θ satisfies $\Lambda(\Phi_t^\theta) \subset \mathcal{B}(0, 1)$. Moreover, if $T_t^\theta(X) = \lambda X$ with

$0 \neq X \in \mathcal{S}_{2n}$, then $\Phi_t^\theta \varphi(X) = \lambda \varphi(X)$ with $0 \neq \varphi(X) \in \mathcal{C}^{n(2n+1)}$, which implies $\Lambda(\mathcal{T}_t^\theta) \subseteq \Lambda(\Phi_t^\theta)$. Further, since the domain of \mathcal{T}_t^θ is isomorphic to that of Φ_t^θ , \mathcal{T}_t^θ has the same number of eigenvalues as Φ_t^θ if taking multiple eigenvalues into account. Thus, $\Lambda(\mathcal{T}_t^\theta) = \Lambda(\Phi_t^\theta)$. In conclusion, $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is AMSS if and only if $\Lambda(\mathcal{T}_t^\theta) \subset \mathcal{B}(0, 1)$. Again, by Proposition 3.3 [13] and similar arguments, the statement of weak stability can be obtained. This ends the proof.

Remark 1. When applying Theorem 1 to verify the stability of system (1), it may be more convenient to compute the spectrum of Φ_t^θ instead of \mathcal{T}_t^θ . However, Theorem 1 supplies a spectral characterization of stability that is fundamental to the study of stabilizability and to develop the Lyapunov-type criteria.

As an immediate extension of Theorem 1, we can obtain the following spectral criterion about stabilizability.

Corollary 1. *The discrete-time mean-field stochastic system (1) is stabilizable if and only if there exists a pair of θ -periodic feedback gains $(K_t, \bar{K}_t) \in \mathcal{R}^{n_u \times n} \times \mathcal{R}^{n_u \times n}$ such that the monodromy operator of system (3), denoted by $\bar{\mathcal{T}}_t^\theta$, satisfies $\Lambda(\bar{\mathcal{T}}_t^\theta) \subset \mathcal{B}(0, 1)$.*

Next, we will present some Lyapunov-type stability results for AMSS and weak stability of discrete-time mean-field stochastic periodic systems (2).

Theorem 2. (i) $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is AMSS if and only if there is a pair of θ -periodic real positive definite matrix sequences $(P_t^1, P_t^2) \in \mathcal{S}_n \times \mathcal{S}_n$ such that

$$\begin{cases} A_t' P_{t+1}^1 A_t + C_t' P_{t+1}^1 C_t - P_t^1 < 0, & t \in \mathbb{Z}_+, \\ (A_t + \bar{A}_t)' P_{t+1}^2 (A_t + \bar{A}_t) + (C_t + \bar{C}_t)' P_{t+1}^1 (C_t + \bar{C}_t) - P_t^2 < 0; \end{cases} \quad (20)$$

(ii) $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is AMSS if and only if there is a pair of θ -periodic real positive definite matrix sequences $(P_t^1, P_t^2) \in \mathcal{S}_n \times \mathcal{S}_n$ such that

$$\begin{cases} A_t P_t^1 A_t' + C_t P_t^1 C_t' - P_{t+1}^1 < 0, & t \in \mathbb{Z}_+, \\ (A_t + \bar{A}_t) P_t^2 (A_t + \bar{A}_t)' + (C_t + \bar{C}_t) P_t^1 (C_t + \bar{C}_t)' - P_{t+1}^2 < 0; \end{cases} \quad (21)$$

(iii) $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is weakly stable if there is a pair of θ -periodic real positive definite matrix sequences $(P_t^1, P_t^2) \in \mathcal{S}_n \times \mathcal{S}_n$ such that

$$\begin{cases} A_t' P_{t+1}^1 A_t + C_t' P_{t+1}^1 C_t - P_t^1 \leq 0, & t \in \mathbb{Z}_+, \\ (A_t + \bar{A}_t)' P_{t+1}^2 (A_t + \bar{A}_t) + (C_t + \bar{C}_t)' P_{t+1}^1 (C_t + \bar{C}_t) - P_t^2 \leq 0. \end{cases} \quad (22)$$

Proof. According to (19), it can be deduced that $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is AMSS if and only if the discrete-time linear system (17) is asymptotically stable. Note that for linear periodic homogeneous systems, exponential stability is equivalent to asymptotic stability. Thus, by Theorems 2.4(vii) and 2.5(ii) [8], (17) is asymptotically stable if and only if there exists a θ -periodic positive matrix sequence $\{P_t\}_{t \in \mathbb{Z}_+}$ satisfying

$$\mathcal{L}_t^*(P_{t+1}) - P_t < 0, \quad P_t \in \mathbb{S}_{2n}^+. \quad (23)$$

According to (7), we can set $P_t = \begin{bmatrix} P_t^1 & 0 \\ 0 & P_t^2 \end{bmatrix}$ without loss of generality. By calculating (23), we come to the statement (i). By the same line of [11], the rest of proof can be shown and details are omitted here.

Remark 2. By Theorem 2, it can be deduced that if $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is AMSS, then the following periodic Lyapunov equation has a θ -periodic real positive definite solution:

$$A'_t P_{t+1}^1 A_t + C'_t P_{t+1}^1 C_t - P_t^1 < 0, \quad t \in \mathbb{Z}_+, \quad (24)$$

which implies that the following stochastic periodic system is AMSS:

$$\begin{cases} x_{t+1} = A_t x_t + C_t x_t w_t, \\ x_0 = \xi \in \mathcal{R}^n, \quad t \in \mathbb{Z}_+. \end{cases} \quad (25)$$

Obviously, the converse implication does not hold since the solvability of (24) can not guarantee that of (20). Besides, it is natural to infer that the condition of statement (iii) should be not only sufficient but also necessary for $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ to be weakly stable. At present, how to validate this conjecture still remains open. According to [11], if (22) has a pair of θ -periodic real positive definite solution, then (25) is also weakly stable.

Remark 3. In the proof of Theorem 2, if we denote the left-hand side of (23) as

$$\mathcal{L}_t^*(P_{t+1}) - P_t = -Q_t, \quad P_t \in \mathbb{S}_{2n}^+,$$

then $Q_t > 0$ and $Q_t \in \mathbb{S}_{2n}^+$. By simple rearrangements, it is realized that $[A_t, \bar{A}_t; C_t, \bar{C}_t]$ is AMSS if and only if there is a pair of n -dimensional real positive definite matrices (Q_t^1, Q_t^2) such that the following Lyapunov-type equation admits a pair of θ -periodic real positive definite solutions:

$$\begin{cases} P_t^1 = A'_t P_{t+1}^1 A_t + C'_t P_{t+1}^1 C_t + Q_t^1, \\ P_t^2 = (A_t + \bar{A}_t)' P_{t+1}^2 (A_t + \bar{A}_t) + (C_t + \bar{C}_t)' P_{t+1}^1 (C_t + \bar{C}_t) + Q_t^2, \end{cases} \quad t \in \mathbb{Z}_+, \quad (26)$$

where $(P_t^1, P_t^2) \in \mathcal{S}_n \times \mathcal{S}_n$.

By Theorem 2, it is easy to derive a necessary and sufficient condition for the stabilization of considered systems. To this end, we introduce the following notations:

$$\tilde{A}_t = A_t + \bar{A}_t, \quad \tilde{B}_t = B_t + \bar{B}_t, \quad \tilde{C}_t = C_t + \bar{C}_t, \quad \tilde{D}_t = D_t + \bar{D}_t.$$

Corollary 2. *The discrete-time mean-field stochastic system (1) is stabilizable if and only if the following difference linear matrix inequalities (LMIs) are feasible with a θ -periodic quadruple (X_t, Y_t, Z_t, W_t) on $t \in [0, \theta - 1]$:*

$$\begin{bmatrix} X_t & X_t A'_t + Z_t B'_t & X_t C'_t + Z_t D'_t \\ * & X_{t+1} & 0 \\ * & * & X_{t+1} \end{bmatrix} > 0, \quad (27)$$

$$\begin{bmatrix} Y_t & Y_t \tilde{A}'_t + W_t \tilde{B}'_t & Y_t \tilde{C}'_t + W_t \tilde{D}'_t \\ * & Y_{t+1} & 0 \\ * & * & Y_{t+1} \end{bmatrix} > 0, \quad (28)$$

where X_t and Y_t are both n -dimensional real positive matrices, $Z_t \in \mathcal{R}^{n \times n_u}$, and $W_t \in \mathcal{R}^{n \times n_u}$. Furthermore, the stabilizing feedback gains are given by $K_t = Z_t' X_t^{-1}$ and $\bar{K}_t = W_t' Y_t^{-1} - Z_t' X_t^{-1}$.

Definition 5. Given $0 < \gamma \leq 1$, discrete-time mean-field stochastic system (1) is called $\mathcal{B}(0, \gamma)$ -stabilizable, if there is a pair of θ -periodic feedback gains $(K_t, \bar{K}_t) \in \mathcal{R}^{n_u \times n} \times \mathcal{R}^{n_u \times n}$ such that the monodromy operator \tilde{T}_t^θ of system (3) satisfies $\Lambda(\tilde{T}_t^\theta) \subset \mathcal{B}(0, \gamma) := \{\lambda | |\lambda| < \gamma, \lambda \in \mathcal{C}\}$.

In view of Corollary 1, it is clear that when $\gamma = 1$, the property of $\mathcal{B}(0, \gamma)$ -stabilization reduces to asymptotic mean square stabilization defined by Definition 2. Moreover, the monodromy operator \tilde{T}_t^θ of system (3) satisfies $\Lambda(\tilde{T}_t^\theta) \subset \mathcal{B}(0, \gamma)$ if and only if the following discrete-time mean-field stochastic system is AMSS:

$$\begin{cases} x_{t+1} = \frac{1}{\sqrt[\theta]{\gamma}} \{(A_t + B_t K_t)x_t + [\bar{A}_t + B_t \bar{K}_t \\ \quad + \bar{B}_t(K_t + \bar{K}_t)]Ex_t + \{(C_t + D_t K_t)x_t \\ \quad + [\bar{C}_t + D_t \bar{K}_t + \bar{D}_t(K_t + \bar{K}_t)]Ex_t\}w_t\}, \\ x_{t_0} = \xi \in \mathcal{R}^n, t \in \mathbb{Z}_+. \end{cases} \quad (29)$$

Similar to Corollary 2, we can obtain an LMI-based criterion for $\mathcal{B}(0, \gamma)$ -stabilization. The proof is omitted for brevity.

Theorem 3. The discrete-time mean-field stochastic system (1) is $\mathcal{B}(0, \gamma)$ -stabilizable if and only if the following coupled difference LMIs are feasible with a θ -periodic quadruple (X_t, Y_t, Z_t, W_t) on $t \in [0, \theta - 1]$:

$$\begin{bmatrix} X_t & \frac{1}{\sqrt[\theta]{\gamma}}(X_t A'_t + Z_t B'_t) & \frac{1}{\sqrt[\theta]{\gamma}}(X_t C'_t + Z_t D'_t) \\ * & X_{t+1} & 0 \\ * & * & X_{t+1} \end{bmatrix} > 0,$$

$$\begin{bmatrix} Y_t & \frac{1}{\sqrt[\theta]{\gamma}}(Y_t \tilde{A}'_t + W_t \tilde{B}'_t) & \frac{1}{\sqrt[\theta]{\gamma}}(Y_t \tilde{C}'_t + W_t \tilde{D}'_t) \\ * & Y_{t+1} & 0 \\ * & * & Y_{t+1} \end{bmatrix} > 0,$$

where X_t and Y_t are both n -dimensional real positive matrices, and $(Z_t, W_t) \in \mathcal{R}^{n \times n_u} \times \mathcal{R}^{n \times n_u}$.

Remark 4. It is notable that there is no essential difficulty to deal with the general regional stability and stabilization problem for discrete-time mean-field stochastic periodic systems by use of the preceding approach, which is left for the interested readers.

3 Conclusion

In this paper, we have studied the stability of mean-field stochastic periodic systems. Stability criteria are presented in terms of the spectral distribution of monodromy operator. In addition, Lyapunov-type stability criteria are obtained for the considered systems. By the Lyapunov criteria, LMI-based conditions are

established for the stabilizability in three senses. It is expected that the current study will have useful applications in the studies of related control issues, such as optimal and mixed H_2/H_∞ control of infinite-horizon cases, which deserve a further exploration.

Acknowledgement. This work was supported by the Natural Science Foundation of Shandong Province (ZR2016FM16), the SDUST Research Fund (No. 2015TDJH105), and Jing-Ying Project of Shandong University of Science and Technology.

References

1. McKean, H.P.: A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. U.S.A.* **56**, 1907–1911 (1966)
2. Huang, M., Caines, P.E., Malhamé, R.P.: Social optima in mean field LQG control: centralized and decentralized strategies. *IEEE Trans. Autom. Control* **57**, 1736–1751 (2012)
3. Wang, B., Huang, M.: Mean field production output control with sticky prices: nash and social solutions. *Automatica* **100**, 90–98 (2019)
4. Bauso, D., Tembine, H., Basar, T.: Opinion dynamics in social networks through mean-field games. *SIAM J. Control Optim.* **54**, 3225–3257 (2016)
5. Ni, Y.H., Elliott, R., Li, X.: Discrete-time mean-field stochastic linear-quadratic optimal control problems, II: infinite horizon case. *Automatica* **57**, 65–77 (2015)
6. Wang, G.C., Zhang, C.H., Zhang, W.: Stochastic maximum principle for mean-field type optimal control under partial information. *IEEE Trans. Autom. Control* **59**, 522–528 (2014)
7. Li, C., Chen, M., Lam, J., Mao, X.: On exponential almost sure stability of random jump systems. *IEEE Trans. Autom. Control* **57**, 3064–3077 (2012)
8. Dragan, V., Morozan, T., Stoica, A.M.: Mathematical Methods in Robust Control of Discrete-Time Linear Stochastic Systems. Springer, New York (2010). <https://doi.org/10.1007/978-1-4419-0630-4>
9. Ma, H., Jia, Y.: Stability analysis for stochastic differential equations with infinite Markovian switchings. *J. Math. Anal. Appl.* **435**, 593–605 (2016)
10. Ma, L.M., Zhang, W., Zhao, Y.: Study on stability and stabilizability of discrete-time mean-field stochastic systems. *J. Franklin Inst.* **356**, 2153–2171 (2019)
11. Ma, H., Hou, T., Zhang, W.: Stability and structural properties of stochastic periodic systems: an operator-spectral approach. In: Proceedings of 2016 American Control Conference, pp. 3880–3885, Boston, USA (2016)
12. Zhang, W., Chen, B.S.: \mathcal{H} -representation and applications to generalized Lyapunov equations and linear stochastic systems. *IEEE Trans. Autom. Control* **57**, 3009–3022 (2012)
13. Bittanti, S., Colaneri, P.: Periodic Systems: Filtering and Control. Springer, London (2009). <https://doi.org/10.1007/978-1-84800-911-0>



Research on Sliding Mode Control of Multi-differential Omnidirectional Mobile Platform

Huimin Zhang¹, Yue Ma^{1,2(✉)}, and Koanhee Cho¹

¹ School of Mechanical Engineering, Special Vehicle Lab,
Beijing Institute of Technology, Beijing 100081, China
yuema.bit@gmail.com

² Beijing Institute of Technology Chongqing Innovation Center,
Chongqing 401120, China

Abstract. Compared with traditional wheeled robots, omnidirectional mobile robots can achieve movement in any direction without changing the posture and have a center steering function, which is flexible and maneuverable. In order to solve the problem that the trajectory tracking effect is poor when there are uncertain parameters and unknown disturbances in the motion control of traditional omnidirectional mobile robots, this paper takes a four-wheel omnidirectional mobile platform as the research object and establishes its kinematic model. The motion control method based on sliding mode variable structure control is proposed and the trajectory tracking controller is designed. The simulation results show that the controller effectively improves the trajectory tracking accuracy during the movement of the omnidirectional mobile platform.

Keywords: Sliding mode control · Omnidirectional movement · Kinematic model

1 Overview of Omnidirectional Mobile Platform and Motion Control

In recent years, with the requirements of industrial automation and intelligence, the mobility and flexibility requirements of mobile operating robots have become higher and higher and omnidirectional mobile robots have begun to develop. Compared with the traditional wheeled robot, the omnidirectional mobile robot can achieve movement in any direction without changing the current posture and can complete the center steering at the same time. It has flexible movement and strong maneuverability and can move flexibly in various narrow working spaces. So it is very suitable for working in various complex indoor environments. The concept of omnidirectional movement was proposed relatively late and its development was relatively slow. Therefore, the related technologies of

omnidirectional mobile robots are not very mature. The motion control system is the basis for the entire omnidirectional mobile robot to realize its various functions. All motion-related actions of the robot are performed by the motion control system. So in a sense, the quality of motion control will directly affect the omnidirectional performance of the mobile robot system and is decisive for the full play of the unique motion advantages of the omnidirectional mobile robot. Research on motion control of omnidirectional mobile platforms has been carried out abroad for many years, but the influence of factors such as strong coupling and nonlinearity caused by its complicated mechanical structure has brought difficulty to robot motion control.

Due to the complex mechanical structure of the omnidirectional mobile robot, the research on its motion control is also started from a relatively simple kinematics level. The system kinematics model is established based on the input and output relationship and on this basis, the motion control method for the mobile robot is established. However, when the parameters of the kinematic model of the system cannot be obtained accurately or there are uncertain factors such as friction and coupling, the overall control effect of the robot is not satisfactory. To this end, some scholars apply adaptive methods and intelligent control methods to solve the problem. But these intelligent methods often require continuous learning to achieve the desired effect, resulting in poor real-time performance [1, 2].

Therefore, the researchers have turned their attention to the motion control with the uncertainties of the model. For the different types of uncertainty (parametric uncertainty and non-parametric uncertainty), the system models are different. So many different processing methods have been proposed, mainly including adaptive control, robust control, sliding model control and intelligent control methods.

In the research process, in order to realize the online estimation of the unknown parameters in the model and in order to improve the anti-interference ability, the online adjustment of the control parameters is required. Many researchers use adaptive control methods to achieve this goal. The adaptive control method can solve the uncertainty problems in some omnidirectional mobile platforms and robot motion control. In order to ensure that the adaptive control law can overcome the uncertainty through accurate online estimation of unknown parameters, the parameters are generally required to be slowly time-varying. But this is not completely guaranteed in the actual system. So it needs to be combined with other methods to meet the control requirements [3].

The controller designed based on the robust control theory, whether it is for the purpose of robust stability or suppression of interference, can be concluded that the feedback controller makes the closed-loop system stable and the H_{∞} norm of the closed-loop transfer function matrix is minimum or less than a certain value. Robust control mainly solves the problem of suppressing bounded uncertainty disturbances so that its norm is minimum or less than a certain value [4].

Sliding mode control is a control scheme that can effectively overcome the uncertain problems such as unbuilt dynamics of the model and external

interference. It is widely used in the motion control of mobile robots. Although the sliding mode control method can effectively overcome the influence of uncertain problems on the control system, due to its own chattering problem, it has added an additional difficulty to the controller design [5].

Although the above methods can achieve better control results for robot control in the presence of system model uncertainty and external interference, these methods generally require designers to have sufficient knowledge of the system's model knowledge and the design of the controller. A lot of complicated and rigorous derivation is required. When the application environment changes or the robot's mechanical structure changes, the control effect becomes poor due to the lack of adaptability of the dynamic model used. Therefore, the intelligent control (neural network control, fuzzy control, etc.) method that does not rely on the specific model of the system has been widely used in robot motion control [6–8]. The application of the intelligent control method is mainly reflected in two aspects. One is to use the powerful approximation ability of the neural network and the fuzzy control method to identify the model of the system online or offline. The second is online estimation of various unknown or difficult parameters such as friction, disturbance, unbuilt dynamics of the system. Although the method based on intelligent control can achieve very good motion control results in some cases, some control performances are also affected by some factors of these methods themselves, such as parameter setting, approximation accuracy, structural complexity and rule making, the expression form of the function, etc.

Through the analysis of the above-mentioned adaptive control, robust control, sliding mode control, intelligent control and other control methods, we can see that these control methods have more or less defects and their application range is only suitable for the control process. In some cases, modern research scholars increasingly integrate two or more of them to make up for their shortcomings, maximize their own advantages and design many hybrid controllers with different structures and perfect functions.

The omnidirectional mobile platform designed in this paper is shown in Fig. 1 and Fig. 2. The transmission diagram of the omnidirectional mobile platform is shown in Fig. 3. The platform adopts a driving scheme of a straight drive motor and four steering motors.

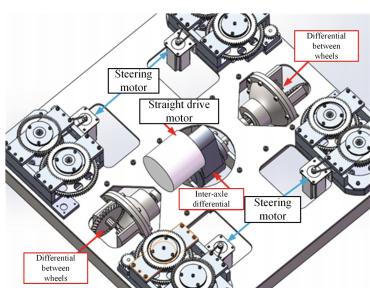


Fig. 1. 3D model of the omnidirectional platform



Fig. 2. The omnidirectional mobile platform

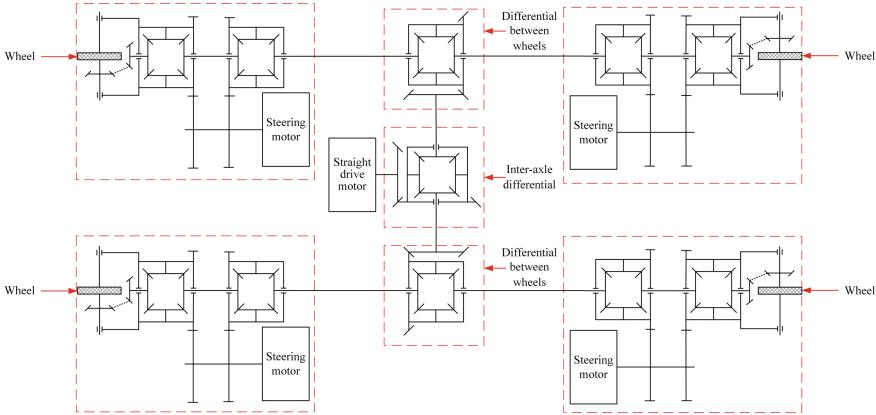


Fig. 3. Transmission diagram of omnidirectional mobile platform

The straight driving power is transmitted to the universal wheel mechanism through the inter-axle differential and the inter-wheel differential and finally drives wheels to roll. The four steering motors control the rotation of the four wheels. Based on this, this paper builds a kinematic model of the omnidirectional mobile platform and simulates its kinematic model through MATLAB.

In this paper, the four-wheel omnidirectional mobile platform is taken as the research object to study the problem of poor trajectory tracking accuracy when there are uncertain parameters and unknown disturbances in motion control. A trajectory tracking motion control method based on sliding mode variable structure control method is proposed. It solves the problems of uncoordinated motion and low trajectory tracking accuracy during the movement of the omnidirectional mobile robot.

2 Kinematic Modeling of Omnidirectional Mobile Platform

The kinematics analysis of mobile robots is to study the kinematic characteristics of mobile robots, without considering the force of the robot during the motion process. The kinematics model of the mobile robot is a mathematical model that reflects the relationship between the control input and motion output of the mobile robot. In this paper, the control input includes the speed n of the straight drive motor, the torque T of the straight drive motor, the rotation angle of the steering motor, the speed of the steering motor and the torque of the steering motor. The motion output includes the longitudinal speed, lateral speed, center of mass deflection angle and yaw rate of the omnidirectional platform.

As shown in Fig. 3, the omnidirectional platform uses a straight drive motor and four steering motors. The simplified kinematics model is shown in Fig. 4.

Suppose the global coordinate system is XOY , the vehicle track is B , the wheelbase is L , the steering angle of each wheel is θ_1 , θ_2 , θ_3 and θ_4 , the

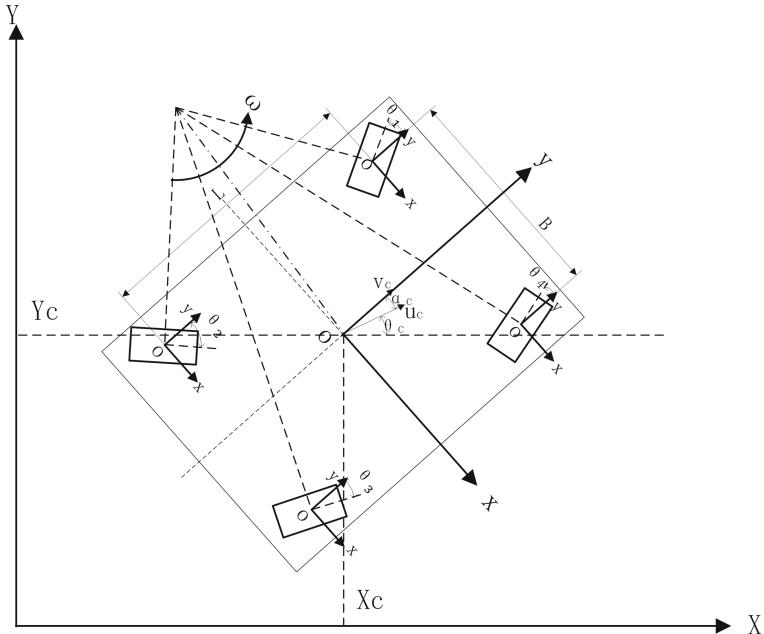


Fig. 4. Kinematics model of omnidirectional mobile platform

instantaneous speed of the vehicle is u and the angle between the instantaneous speed u and the y -axis of the vehicle coordinate system is α .

Define the vehicle coordinate system: the coordinate system is fixed to the vehicle body. The yz plane is the left-right symmetry plane of the vehicle body and the coordinate origin o coincides with the center of mass. When the vehicle is stationary on the horizontal road, the y -axis is parallel to the ground and points in the direction of the front of the vehicle, the z -axis passes the center of mass vertically and the x -axis passes the center of mass and points to the driver's right.

Define the tire coordinate system: suppose the wheel midplane perpendicular to the wheel rotation axis is the wheel plane and the coordinate origin o is the intersection of the intersection of the wheel plane and the ground plane and the projection line of the wheel rotation axis on the ground plane. The y -axis is the intersection of the wheel plane and the ground plane and points forward. The z -axis is perpendicular to the ground plane. The y -axis is perpendicular to the wheel plane and points to the right.

Set under the global coordinate system XOY , the horizontal and vertical coordinates of the center of mass of the vehicle are X_c and Y_c and the angle between the vehicle speed and the X axis of the global coordinate system is θ_c . The actual longitudinal speed and angular speed of the vehicle are respectively v_c and ω_c . The horizontal and vertical coordinates of the center of mass of the reference target vehicle are X_r and Y_r and the angle between the target

vehicle speed and the X axis of the global coordinate system is θ_r . The target longitudinal speed and angular speed of the vehicle are v_r and ω_r . Then

$$\begin{bmatrix} \dot{x}_c \\ \dot{y}_c \\ \dot{\theta}_c \end{bmatrix} = \begin{bmatrix} \cos(\theta_c - \alpha_c) & 0 \\ \sin(\theta_c - \alpha_c) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{v_c}{\cos \alpha_c} \\ \omega_c \end{bmatrix} \quad (1)$$

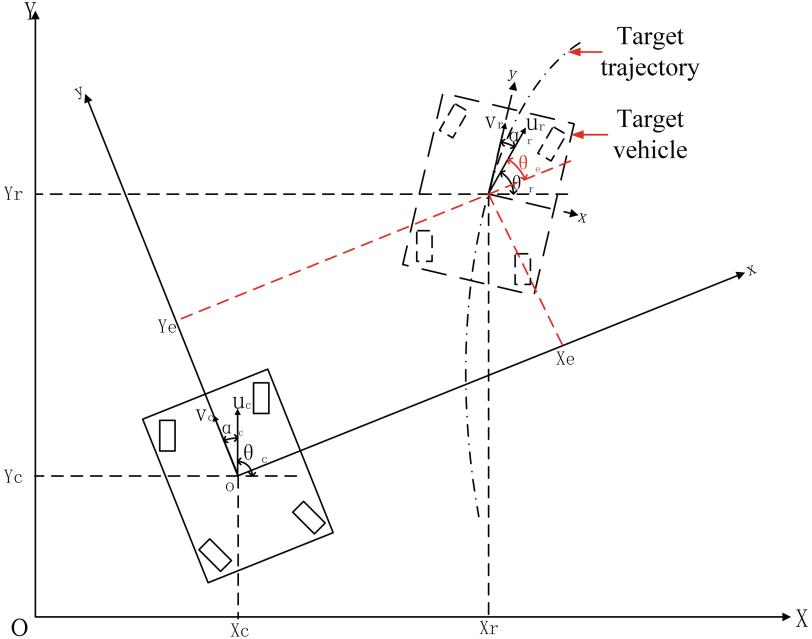


Fig. 5. Trajectory tracking model of omnidirectional mobile platform

As shown in Fig. 5, set under the vehicle coordinate system xoy , the horizontal and vertical coordinates of the reference vehicle center of mass are X_e and Y_e and the angle between the target vehicle speed and the x axis of the vehicle coordinate system is θ_e . From the coordinate transformation, the pose error of the platform in the vehicle coordinate system can be expressed as

$$\begin{bmatrix} x_e \\ y_e \\ \theta_e \end{bmatrix} = \begin{bmatrix} \cos \theta_c & \sin \theta_c & 0 \\ -\sin \theta_c & \cos \theta_c & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r - x_c \\ y_r - y_c \\ \theta_r - \theta_c \end{bmatrix} \quad (2)$$

Derived from Eq. 1 and Eq. 2, the differential equation of the pose error of the omnidirectional mobile platform is as follows

$$\begin{bmatrix} \dot{x}_e \\ \dot{y}_e \\ \dot{\theta}_e \end{bmatrix} = \begin{bmatrix} y_e \omega_c - u_c + u_r \cos \theta_e \\ -x_e \omega_c + u_r \sin \theta_e \\ \omega_r - \omega_c \end{bmatrix} \quad (3)$$

The relationship between the vehicle's actual longitudinal speed v_c , actual angular speed ω_c , center-of-mass slip angle α_c and the rotation angle of each wheel θ_m ($m = 1, 2, 3, 4$), vehicle speed u_c is

$$v_c = u_c \cos \alpha_c \quad (4)$$

$$\omega_c = \frac{u_c}{r_m \cos \theta_m + \frac{B}{2}} \quad (5)$$

$$\alpha_c = \arctan \frac{\frac{L}{2} - r_m \sin \theta_m}{\frac{B}{2} + r_m \cos \theta_m} \quad (6)$$

$r_m = \min(r_1, r_2, r_3, r_4)$, r is the distance from each wheel to the steering center.

According to Eq. 4, Eq. 5 and Eq. 6, the differential equation of the pose error of the omnidirectional mobile platform is further derived as follows

$$\begin{bmatrix} \dot{x}_e \\ \dot{y}_e \\ \dot{\theta}_e \end{bmatrix} = \begin{bmatrix} y_e \frac{u_c}{r_m \cos \theta_m + \frac{B}{2}} - u_c + u_r \cos \theta_e \\ -x_e \frac{u_c}{r_m \cos \theta_m + \frac{B}{2}} + u_r \sin \theta_e \\ \omega_r - \frac{u_c}{r_m \cos \theta_m + \frac{B}{2}} \end{bmatrix} \quad (7)$$

The structure diagram of the motion control system based on the omnidirectional mobile platform kinematics is shown in Fig. 6.

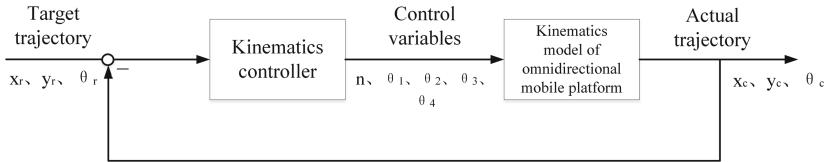


Fig. 6. The structure diagram of the motion control system

3 Sliding Mode Variable Structure Control Design

In view of the characteristics of nonlinear, strong coupling and multivariable characteristics of the omnidirectional mobile platform in this paper, it is difficult to establish an accurate mathematical model of the controlled mobile platform and the complexity of the external environment makes the tracking process uncertain. The trajectory of the mobile platform deviates from the ideal path. In this paper, a sliding mode variable structure controller based on the kinematic model is designed.

Sliding mode control is a control strategy of a variable structure control system. The fundamental difference from conventional control lies in the discontinuity of control, which is a switching characteristic that changes the system structure with time. Sliding mode variable structure control does not depend

on the precise mathematical model of the controlled object and has the advantages of fast response and insensitivity to parameters. It is suitable for the control of omnidirectional mobile platforms in complex environments. The expected motion of the omnidirectional mobile platform can be expressed as a function of spatial coordinates. Various states of the platform are time-varying, such as angular velocity, linear velocity, spatial position state, etc. The tracked trajectory not only has spatial position information, but also includes kinematics model and tracking trajectory function. Since several control laws in the sliding mode variable structure control method are set in advance, it is not affected by disturbances and does not depend on system parameters. So this method is more robust.

Designing the sliding mode variable structure controller includes the switching function and the sliding mode control law. Designing the switching function makes the sliding mode asymptotically stable and have good dynamic quality. The sliding mode control law is designed to satisfy the arrival condition and form a sliding mode area on the switching surface. The kinematic model of the omnidirectional mobile platform is a multi-input nonlinear system. The process of designing the sliding mode control switching function is as follows.

3.1 Switching Function

Let $V_y = \frac{1}{2}y_e^2$, when $\theta_e = -\arctan(v_r y_e)$,

$$\begin{aligned}\dot{V}_y &= y_e \dot{y}_e = y_e(-x_e \omega_c + v_r \sin \theta_e) \\ &= -y_e x_e \omega_c - v_r y_e \sin[\arctan(v_r y_e)]\end{aligned}\tag{8}$$

Where $v_r y_e \sin[\arctan(v_r y_e)] \geq 0$,

Let $f = x \sin(\arctan x)$, When $x > 0$, $\arctan x \in (0, \frac{\pi}{2})$, then $\sin(\arctan x) > 0$, $f > 0$. When $x < 0$, $\arctan x \in (-\frac{\pi}{2}, 0)$, then $\sin(\arctan x) < 0$, $f > 0$. So Eq. 8 is satisfied

$$\dot{V}_y \leq 0$$

When x_e converges to zero and θ_e converges to $-\arctan(v_r y_e)$, y_e converges to zero. Such a sliding mode surface ensures that the system has good global stability.

The switching function can be designed as

$$s = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} x_e \\ \theta_e + \arctan(v_r y_e) \end{bmatrix}\tag{9}$$

In order to ensure that any point in the state space has good approaching motion quality and reaches the switching surface within a limited time, sliding mode control is given by exponential approach law. The exponential approach law is as follows

$$\dot{s} = \begin{bmatrix} \dot{s}_1 \\ \dot{s}_2 \end{bmatrix} = \begin{bmatrix} -\epsilon_1 sgn s_1 - k_1 s_1 \\ -\epsilon_2 sgn s_2 - k_2 s_2 \end{bmatrix}\tag{10}$$

Among them, k_1 , k_2 , ϵ_1 and ϵ_2 are normal numbers. According to sliding mode control theory, the system satisfies the reachable conditions and is globally stable. Derivation of the former switching function combined with known conditions can be obtained

$$\begin{aligned}\dot{s} &= \begin{bmatrix} \dot{s}_1 \\ \dot{s}_2 \end{bmatrix} = \begin{bmatrix} \dot{s}_1 \\ \dot{\theta}_e + \frac{\partial \arctan(v_r y_e)}{\partial V_r} \dot{V}_r + \frac{\partial \arctan(v_r y_e)}{\partial y_e} \dot{y}_e \end{bmatrix} \\ &= \begin{bmatrix} y_e \omega_c - v_c + v_r \cos \theta_e \\ \omega_r - \omega_c + \frac{\partial \arctan(v_r y_e)}{\partial V_r} \dot{V}_r + \frac{\partial \arctan(v_r y_e)}{\partial y_e} (-x_e \omega_c + v_r \sin \theta_e) \end{bmatrix}\end{aligned}\quad (11)$$

3.2 Control Law

Derived from above, the control law can be obtained as follows

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} y_e \omega_c + v_r \cos \theta_e + \epsilon_1 sgn s_1 + k_1 s_1 \\ \omega_r + \frac{\partial \arctan(v_r y_e)}{\partial V_r} \dot{V}_r + \frac{\partial \arctan(v_r y_e)}{\partial y_e} (-x_e \omega_c + v_r \sin \theta_e) + \\ \epsilon_2 sgn s_2 + k_2 s_2 \end{bmatrix}\quad (12)$$

After obtaining v and ω , the control variables u , θ_1 , θ_2 , θ_3 and θ_4 can be determined according to the kinematic relationship

$$u = v \cos \alpha \quad (13)$$

$$\theta_1 = \begin{cases} -\arctan \frac{\frac{L}{2} + \frac{v}{\omega} \sin \alpha}{-\frac{B}{2} - \frac{v}{\omega} \cos \alpha}, & \frac{v}{\omega} \cos \alpha < -\frac{B}{2} \text{ or } \frac{v}{\omega} \cos \alpha > 0 \\ -\frac{\pi}{2} - \arctan \frac{\frac{B}{2} + \frac{v}{\omega} \cos \alpha}{\frac{L}{2} + \frac{v}{\omega} \sin \alpha}, & -\frac{B}{2} \leq \frac{v}{\omega} \cos \alpha \leq 0 \end{cases} \quad (14)$$

$$\theta_2 = \begin{cases} \arctan \frac{\frac{L}{2} - \frac{v}{\omega} \sin \alpha}{-\frac{B}{2} - \frac{v}{\omega} \cos \alpha}, & \frac{v}{\omega} \cos \alpha < -\frac{B}{2} \text{ or } \frac{v}{\omega} \cos \alpha > 0 \\ \frac{\pi}{2} + \arctan \frac{\frac{B}{2} - \frac{v}{\omega} \cos \alpha}{\frac{L}{2} - \frac{v}{\omega} \sin \alpha}, & -\frac{B}{2} \leq \frac{v}{\omega} \cos \alpha \leq 0 \end{cases} \quad (15)$$

$$\theta_3 = \begin{cases} -\arctan \frac{\frac{L}{2} - \frac{v}{\omega} \sin \alpha}{-\frac{B}{2} + \frac{v}{\omega} \cos \alpha}, & \frac{v}{\omega} \cos \alpha > \frac{B}{2} \text{ or } \frac{v}{\omega} \cos \alpha < 0 \\ -\frac{\pi}{2} - \arctan \frac{\frac{B}{2} - \frac{v}{\omega} \cos \alpha}{\frac{L}{2} - \frac{v}{\omega} \sin \alpha}, & 0 \leq \frac{v}{\omega} \cos \alpha \leq \frac{B}{2} \end{cases} \quad (16)$$

$$\theta_4 = \begin{cases} \arctan \frac{\frac{L}{2} + \frac{v}{\omega} \sin \alpha}{-\frac{B}{2} + \frac{v}{\omega} \cos \alpha}, & \frac{v}{\omega} \cos \alpha > \frac{B}{2} \text{ or } \frac{v}{\omega} \cos \alpha < 0 \\ \frac{\pi}{2} + \arctan \frac{\frac{B}{2} - \frac{v}{\omega} \cos \alpha}{\frac{L}{2} + \frac{v}{\omega} \sin \alpha}, & 0 \leq \frac{v}{\omega} \cos \alpha \leq \frac{B}{2} \end{cases} \quad (17)$$

In summary, the state of the system will reach the sliding mode within a limited time under the action of the sliding mode control law and the platform finally tracks the desired trajectory.

4 Simulation Results

In this paper, *MATLAB/Simulink* software is used. The circular trajectory and sinusoidal trajectory are used as the desired trajectory. The trajectory tracking control law obtained in this paper is simulated to verify the effectiveness of the control law. The expected trajectory in the trajectory tracking effect diagram is represented by a blue curve and the actual trajectory of the omnidirectional platform is represented by a brown curve.

4.1 Circular Trajectory Tracking Simulation

Expected trajectory: $x_r = 2\sin t$, $y_r = 2\cos t$, $\theta_r = -\frac{\pi}{2} - \frac{1}{2}t$

The initial pose of the actual mobile platform: $x_c = 0$, $y_c = 0$, $\theta_c = 0$

Controller parameters: $k_1 = 2$, $k_2 = 1$, $\epsilon_1 = 0.2$, $\epsilon_2 = 0.2$. The simulation result of circular trajectory tracking is shown in Fig. 7. The relative error of the simulation result is shown in Fig. 8.

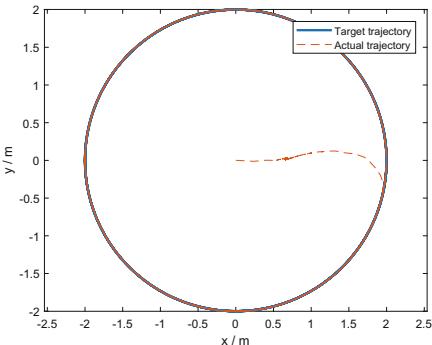


Fig. 7. Circular trajectory tracking simulation result

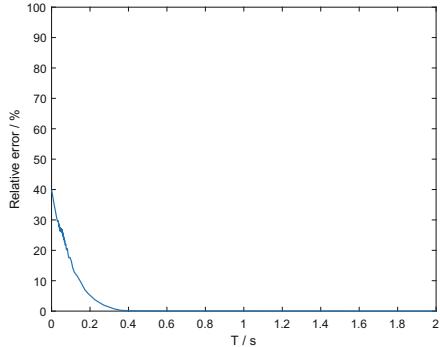


Fig. 8. Relative error graph of circular trajectory tracking simulation

It can be seen from Fig. 8 that under the action of the controller, the vehicle gradually approaches the target trajectory in 0.4 s. The vehicle tracks the target trajectory at 1.6 s and the relative error is kept within a relatively small range.

The simulation results of control variables v and ω are shown in Fig. 9 and Fig. 10.

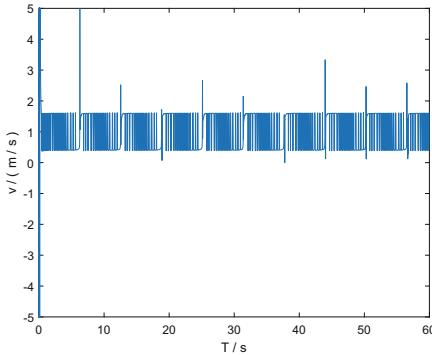


Fig. 9. Simulation result of control variable v

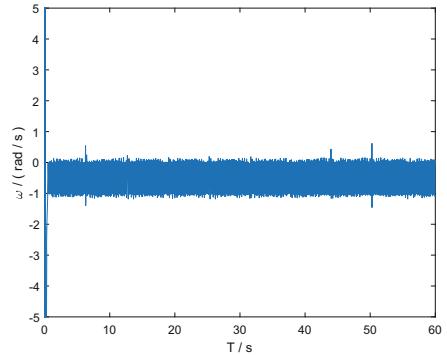


Fig. 10. Simulation result of control variable ω

4.2 Sinusoidal Trajectory Tracking Simulation

Expected trajectory: $x_r = 5 + t$, $y_r = 5 \sin \frac{\pi}{20}t$, $\theta_r = 0$

The initial pose of the actual mobile platform: $x_c = 0$, $y_c = 0$, $\theta_c = 0$

Controller parameters: $k_1 = 2$, $k_2 = 1$, $\epsilon_1 = 0.2$, $\epsilon_2 = 0.2$. The simulation result of sinusoidal trajectory tracking is shown in Fig. 11. The relative error of the simulation result is shown in Fig. 12.

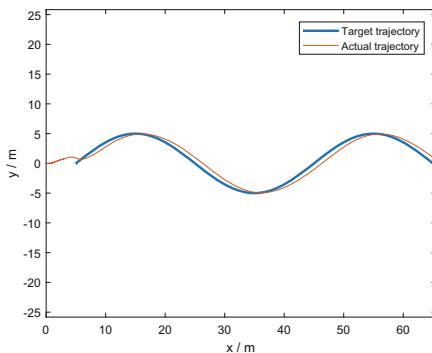


Fig. 11. Sinusoidal trajectory tracking simulation result

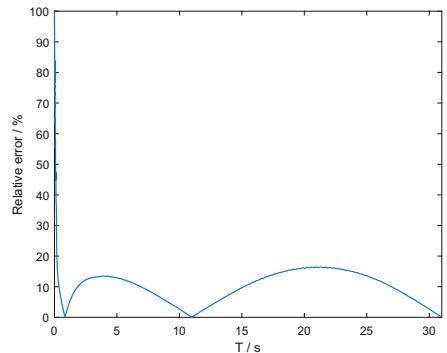


Fig. 12. Relative error graph of sinusoidal trajectory tracking simulation

It can be seen from Fig. 12 that under the action of the controller, the vehicle gradually approaches the target trajectory at 1 s and tracks the target trajectory at 1.8 s. After the vehicle reaches a steady state, the maximum relative error is about 16.6%.

The simulation results of control variables v and ω are shown in Fig. 13 and Fig. 14.

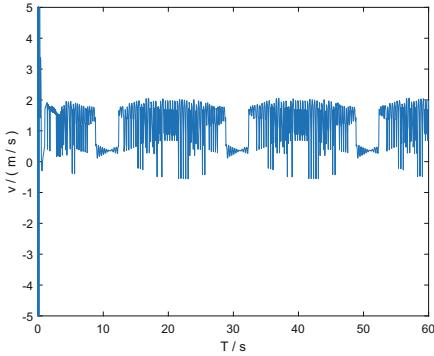


Fig. 13. Simulation result of control variable v

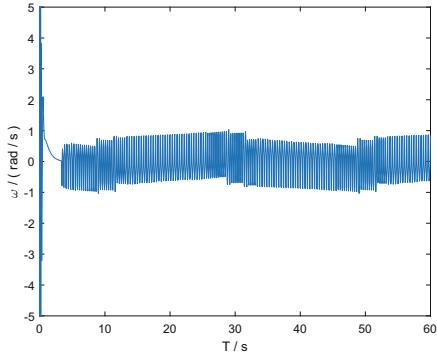


Fig. 14. Simulation result of control variable ω

It can be seen from simulations that for a multi-input multi-output and highly coupled nonlinear system such as the omnidirectional mobile platform, sliding mode variable structure control has a good control effect, which can keep the error of the system in a small range. It can be seen from simulation results that the kinematics controller is effective. The tracking error of the circular trajectory and the sinusoidal trajectory are both small. The trajectory tracking effect of the controller meets requirements. However, due to the characteristics of the sliding mode variable structure, the chatter problem exists in the controller output.

5 Conclusion

In this paper, the four-wheel omnidirectional mobile platform is taken as the research object to study the problem of poor trajectory tracking accuracy when there are uncertain parameters and unknown disturbances in motion control and its kinematic model is established. The trajectory tracking motion control method based on sliding mode variable structure control is proposed and the trajectory tracking controller is designed. The simulation results show that the controller effectively improves the trajectory tracking accuracy during the movement of the omnidirectional mobile platform. The trajectory tracking effect of the controller meets requirements.

For a multi-input multi-output and highly coupled nonlinear system such as the omnidirectional mobile platform, sliding mode variable structure control has a good control effect, which can keep the error of the system in a small range. It can be seen from simulation results that the kinematics controller is effective. However, due to the characteristics of the sliding mode variable structure, the chatter problem exists in the controller output. In the future research, continuous function can be used to replace the sign function in the sliding mode variable structure controller to reduce the chattering of the system. Future research can also combine other control methods to design hybrid controllers with different structures and perfect functions to make up for their shortcomings and maximize their advantages.

References

1. Huang, H.C., Tsai, C.C.: FPGA implementation of an embedded robust adaptive controller for autonomous omnidirectional mobile platform. *J. IEEE Trans. Ind. Electron.* **56**(5), 1604–1616 (2009). <https://doi.org/10.1109/TIE.2008.2009524>
2. Huang, H.C., Lu, S.: Intelligent motion control for four-wheeled omnidirectional mobile robots using ant colony optimization. *Appl. Artif. Intell.* **212**(3), 94–106 (2011)
3. Huang, J.T., Hung, T.V.: Singularity-free adaptive control for uncertain omnidirectional mobile robots. In: Proceedings of 9th Asian Control Conference, Istanbul, Turkey (2013)
4. Xu, D., Zhao, D., Yi, J., Tan, X.: Trajectory tracking control of omnidirectional wheeled mobile manipulators: robust neural network based sliding mode approach. *J. IEEE Trans. Syst. Man Cybern.* **39**(3), 788–799 (2009)
5. Liu, C.H.: Design of integral sliding-mode controller with finite-time convergence for omni-directional mobile robots. In: International Conference on Advanced Robotics Intelligent Systems, pp. 59–63 (2013)
6. Djebrahi, S., Benali, A., Abdessemed, F.: Modelling and control of an omnidirectional mobile manipulator. *J. Int. J. Appl. Math. Comput. Sci.* **22**(3), 601–616 (2012)
7. Ying, J., Bai, B., Wang, S., Ling, L.: Tracking control for omnidirectional lower limbs rehabilitation robot during centre-of-gravity shift. In: Proceedings of 29th Chinese Control Conference, Beijing, China (2010)
8. Xu, D., Zhao, D., Yi, J.: Motion control of omnidirectional mobile manipulators. *J. Chinese J. Mech. Eng.* **45**(1), 42–49 (2009)



Object Tracking Algorithm of UAV Based on Fast Kernel Correlation Filter

Xukuai Liu, Mingjian Sun^(✉), and Lingbo Meng

School of Information Science and Engineering, Harbin Institute of Technology (Weihai), Weihai 264209, China
sunmingjian@hit.edu.cn

Abstract. UAV (Unmanned Aerial Vehicle) is serving as a major platform for developing and testing the artificial intelligence technology. However, how to develop the technology of UAV visual object tracking encounters many numerous difficulties given that the further development of UAV technology is based on it. Firstly, based on the kernel correlation filtering algorithm, this paper uses multiple features to train the regressor respectively, and then fuses the feature map adaptively. Secondly, the template update strategy is changed according to the peak to sideline ratio of the feature map and the similarity between the templates. Next, through the verification on OTB100, the algorithm proposed in this paper is greatly improved compared with others seeing that tracking speed exceeds 30 fps, meeting the real-time requirements. Last but not least, the simulation system of UAV object tracking is built under the ROS (Robot Operating System) platform, and further verified the feasibility of the algorithm.

Keywords: UAV · Object tracking · Kernel correlation filter · Feature

1 Introduction

UAV has become a hot research area in various scientific research institutions and companies, because of high flexibility and maneuverability [1]. In addition, UAV can be well controlled by human beings, and often appears in various civil scenes, such as aerial photography, surveying and mapping, searching and rescuing. The successful execution of these tasks calls for the support of computer vision, especially visual object tracking, which can provide real-time location information of the target, and lay a foundation for UAV to carry out next advanced missions [2].

At present, visual object tracking in UAV faces many challenges, such as illumination, motion blur, etc. [3]. At the same time, UAV often needs to change the posture and adjust the angle of cloud platform during performing tasks, which often appear the occlusion of object caused by the change of the perspective. Therefore, how to solve these problems has become a key research direction in people's research.

Visual object tracking is mainly divided into filtering and deep learning. Deep learning object tracking algorithm generally needs a large number of offline data sets for model training, and also needs a large number of GPU devices to support, otherwise it cannot meet the real-time requirements. UAV generally has small space and high real-time requirements, so it is difficult to deploy deep learning object tracking algorithm. On the contrary, filtering object tracking algorithm is simple, feasible and fast, which can achieve high speed in CPU.

The filtering object tracking algorithm is modeled generally according to the information provided by the initial frame image, and then uses the mathematical model to find out the most similar region with the target in the subsequent frames. Musab et al. used camshift tracking algorithm on the AR.Drone platform for object tracking [4]. This method is fast to track, but may be not suitable for some occasions such as illumination changes in terms of the color histogram feature. Henriques et al. proposed a KCF (Kernel Correlation Filtering) algorithm, which uses cyclic matrix to increase the number of samples, and utilizes the multi-channel HOG (Histogram of Oriented Gradient) feature, which resulted in the tracking effect is better [5]. Chen et al. used the KCF tracking algorithm to track the target on the UAV platform [6]. The effect works well, but the accuracy needs to be improved because of its only a single feature.

The purpose of this paper is to design a high-precision and fast UAV visual object tracking system. Based on the KCF tracking algorithm, this paper utilized the HOG feature and CN (Color Names) feature to train the filter respectively, and adaptively fused the feature map of both. Meanwhile, this paper proposed a new template update strategy, peak to sidelobe ratio was used to measure the quality of tracking, and structural similarity algorithm was tested to describe the similarity between templates. Only the tracking state is well and certain differences between templates exists, the template can be updated. Last but not least, this paper made use of the object tracking benchmark and ROS simulation system to verify the algorithm.

2 Correlation Filter Object Tracking Algorithm

2.1 Kernel Correlation Filter

Tracker generally gets samples near the tracking result of the I_t frame, and trains a regressor by using those samples, which can obtain the response between the image and the template. In frame I_{t+1} , samples near the tracking result area of the I_t frame are being sampled, and the trained regression is used to determine the response of the samples. Meanwhile, the strongest response corresponds to the tracking result of the frame.

Filtering object tracking aims to build a linear regression function $f(\mathbf{z}) = \mathbf{w}^T \mathbf{z}$ to optimize the error function. In the KCF tracker, a linear regression $f(\mathbf{x}) = (w, \varphi(\mathbf{x}))$ is trained by rectangular area, which is a image block x with the size of $m \times n$ around the target as the center. At the same time, cyclic matrix is used to increase the number of samples and extract one-dimensional or multi-dimensional features in KCF algorithm. The error function is as follows:

$$\min_{\mathbf{w}} \sum_i (f(\mathbf{x}_i) - y_i)^2 + \lambda \|\mathbf{w}\|^2 \quad (1)$$

In formula (1), \mathbf{x}_i represents the sample, y_i is the target of regression, and $i \in \{0, \dots, m-1\} \times \{0, \dots, n-1\}$, the r serves as a parameter that controls the fitting situation. Using regularized least square classification model and kernel function, the model becomes:

$$\min_w \sum_{i=1} [(w, \varphi(\mathbf{x})) - y_i]^2 + \lambda \|w\|^2 \quad (2)$$

In formula (2), $\varphi(x)$ represents the mapping of the original input space to the Hilbert feature space. And the target solution can be expressed as follows:

$$\mathbf{w} = \sum_i \alpha_i \varphi(\mathbf{x}_i) \quad (3)$$

And the above question can be solved by using the kernel function $\varphi^T(\mathbf{x})\varphi(\mathbf{x}') = \kappa(\mathbf{x}, \mathbf{x}')$. By using the cyclic matrix and the discrete Fourier transform, the optimal solution for the kernel regular least squares is obtained as follows:

$$\boldsymbol{\alpha} = F^{-1} \left[\frac{F(\mathbf{y})}{F(\mathbf{k}^{xx}) + \lambda} \right] \quad (4)$$

In formula (4), F is the discrete Fourier transform, k^{xx} is the output of kernel k . Gauss kernel is used as kernel function, and the formula is:

$$k^{xx'} = \exp \left\{ -\frac{1}{\sigma^2} \left\{ \|x\|^2 + \|x'\|^2 - 2F^{-1}[F^*(x) \odot F(x')] \right\} \right\} \quad (5)$$

In a new frame, the location of the target is detected by acquiring the image block z of candidate window, the output response of the regression is as follows:

$$\hat{y} = F^{-1} \left[F(\mathbf{k}^{\hat{x}z}) \odot F(\hat{\boldsymbol{\alpha}}) \right] \quad (6)$$

In formula (6), \hat{x} and $\hat{\boldsymbol{\alpha}}$ respectively represent the target appearance template and the correlation filter template obtained by learning. The updating method is as follows:

$$\hat{\boldsymbol{\alpha}}_i = \beta \boldsymbol{\alpha}_i + (1 - \beta) \hat{\boldsymbol{\alpha}}_{i-1} \quad (7)$$

$$\hat{\mathbf{x}}_i = \beta \mathbf{x}_i + (1 - \beta) \hat{\mathbf{x}}_{i-1} \quad (8)$$

In formula (7–8), β stands for the update step size, and the location where \hat{y} is the maximum is the position of the target to be tracked.

Through the above description, the kernel correlation filtering algorithm has the following characteristics: 1) The dense sampling of the training sample is realized according to the cyclic matrix structure, and the solution of the

problem is transformed to the Fourier transform domain by using the cyclic matrix characteristics. Thus, the inverse of the matrix is avoided and the computation is greatly reduced. 2) Through the properties of cyclic matrix, the calculation of correlation filtering is extended to kernel space, and multi-channel feature calculation can be carried out without reducing the speed. Therefore, the kernel correlation filter has the advantages of flexible tracking framework and fast tracking speed. However, this kind of tracking algorithm still has some defects: The robustness of the algorithm needs to be improved because of only using a single feature. Besides, every frame is updated with the template, which may cause tracking drift when the tracking effect is not well. Based on the above shortcomings, this paper improves the kernel correlation filter algorithm to achieve a high-precision visual object tracking algorithm for UAV platform.

2.2 Multi Feature Fusion

It is becoming increasingly difficult to ignore the selected feature for object tracking. Better features can make the regression more responsive, which means better tracking results [7]. There are many kinds of features to describe objects, each of which has its own advantages and limitations. But if a single feature is used to describe the target, it can hardly be used in all situations. Therefore, a combination of features can be used to enhance the robustness of the system. It is noteworthy that the features used should not be too complex for high speed of tracking.

The HOG feature is a gradient direction histogram that calculates and statistics the local areas of the image to form the feature. The main idea of HOG is that the image and shape of local targets can be well described by the directional density distribution of edges or gradients. As HOG operates on local units of an image, it maintains good invariance in both geometric and optical distortion of the image, which can only happen in larger spatial domains. However, HOG feature has poor performance when it occurs in the motion blur of the target.

CN feature is to refine the 3 dimensions color feature into 11 dimensions through mapping function, and support PCA (Principal Component Analysis) to reduce the dimension to get adaptive color attributes. The tracking effect of this feature leaves much to be desired when the illumination changes, but it is better when the object is deformed or occluded. The results of KCF tracking algorithm by using HOG feature and CN feature are shown in Fig. 1:

Figure 1a shows the object tracking results in motion blur, and Fig. 1b displays the object tracking results when changed illumination happens. In Fig. 1, the green rectangle and the blue rectangle represent the results of KCF tracking algorithm using CN feature and HOG feature respectively. It can be known that the HOG feature has a certain robustness to the illumination change, while the CN feature has better performance in motion blur situation.

On the basis of each feature has its own characteristics, this paper puts forward an adaptive kernel correlation filtering algorithm which combines the features of HOG and CN. Different from the previous way of directly fusing the



Fig. 1. Comparison of tracking results using different features in KCF algorithm

characteristics of HOG and CN [8], this paper employs HOG and CN to train a regressor separately, and adaptively fuses the feature map obtained by regressor.

In the tracking process of frame i , assuming that the feature map obtained by the HOG feature is H_i , the feature map obtained by the CN feature is C_i . And in this paper, the quality of the feature map is measured by PSR (Peak to Sidelobe Ratio) [9]. In order to calculate the PSR, the correlation output g is divided into peak value (maximum value of feature map) and sidelobe (except the other pixels in the 11×11 window around the peak value of feature map). The calculation formula of PSR is as follows:

$$\frac{g_{\max} - \mu_{s1}}{\sigma_{s1}} \quad (9)$$

Where g_{\max} is the peak values of feature map, μ_{s1} and σ_{s1} are the mean and standard deviation of the sidelobe in feature map. It can be supposed that, the characteristic feature maps of HOG and CN are turned into h_i and c_i respectively by PSR calculated.

First, the output value of PSR is normalized to 0–1 by formula:

$$\hat{h}_i = \frac{h_i}{\max(h)} \quad (10)$$

$$\hat{c}_i = \frac{c_i}{\max(c)} \quad (11)$$

Assuming that the proportion of the feature map obtained by the fusion through the characteristics of HOG and CN is k and $(1 - k)$, it can be obtained by the following formula:

$$k = \frac{\hat{h}_i}{\hat{h}_i + \hat{c}_i} \quad (12)$$

In this paper, the *Basketball* sequence in OTB2015 [3] object tracking data set is used for testing, which is standard visual object tracking benchmark, and the normalized PSR value changes are shown in Fig. 2:

It can be seen from Fig. 2 that the two normalized values are between 0–1, which shows dynamic changes can be used as the basis of adaptive fusion.

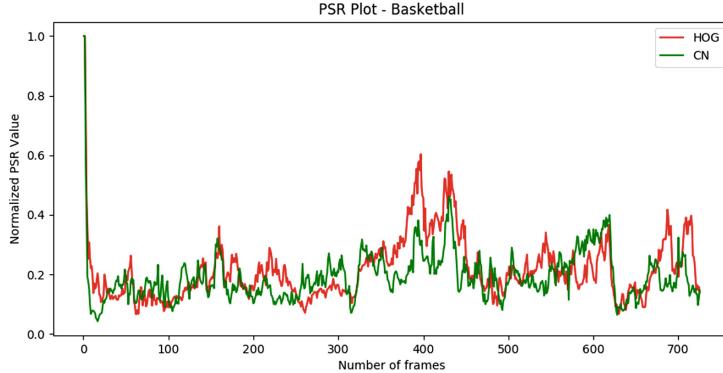


Fig. 2. The changes of normalized PSR value in *Basketball* sequence

3 Template Update Method

In the KCF tracking algorithm, the template update strategy is adopted for every frame. This method can adapt to the deformation of the target in a better way, but it also causes the following problems. Firstly, it will take more time to update the template every frame than every few frames, which causes the tracking speed to decrease. Secondly, when a large area of occlusion encounters or the tracking effect is not very satisfactory, template updating will cause tracking drift and it is difficult to get the correct new template.

In this paper, PSR and template similarity are used to decide whether to update the template. When the tracking effect is not well or a large area of occlusion occurs, which means the PSR value is lower than a certain threshold value, and the template should not be updated. And when there is a large similarity between the pre updated template and the previous template, it will also not be updated. What is more, updating similar templates makes little difference in tracking results and wastes time.

Structural similarity (SSIM) algorithm is taken to evaluate the quality of compression by measuring the similarity between pre-compressed and post-compressed images, which can also be used to measure the similarity between two templates. Comparison with brightness, contrast and structure are the unit constituting the SSIM algorithm. Brightness contrast refers to the average gray value of a picture as a brightness measurement, and the comparison function is as follows:

$$L(X, Y) = \frac{2u_X u_Y + C_1}{u_X^2 + u_Y^2 + C_1} \quad (13)$$

In SSIM algorithm, gray standard deviation is used as the basis of contrast measurement, and the comparison function is as follows:

$$C(X, Y) = \frac{2\sigma_X \sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2} \quad (14)$$

The structure contrast function is as follows:

$$S(X, Y) = \frac{\sigma_{XY} + C_3}{\sigma_X \sigma_Y + C_3} \quad (15)$$

In formula (12–14), u_X and u_Y represent the mean of image X and Y , σ_X and σ_Y represent the standard deviation of image X and Y , and σ_X^2 and σ_Y^2 represent the variance of image X and Y respectively. σ_{XY} represents the covariance between X and Y . In addition, C_1 , C_2 , and C_3 are constants to avoid denominator 0.

The final SSIM calculation is as follows:

$$\text{SSIM}(X, Y) = L(X, Y) * C(X, Y) * S(X_p Y_0) \quad (16)$$

The maximum value of SSIM function is no more than 1, and the larger the value, the more similar the two pictures are. The SSIM algorithm is used to measure the similarity of templates is shown in Fig. 3:

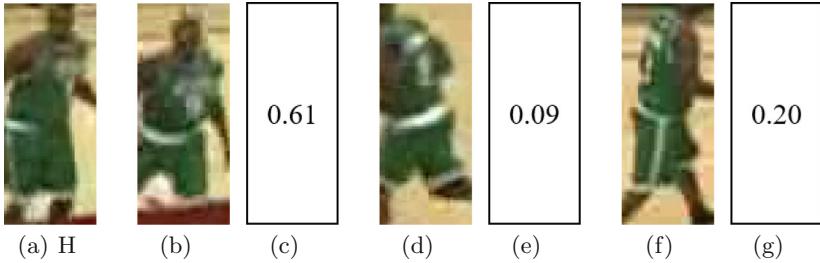


Fig. 3. Using SSIM algorithm to compare the similarity between templates

Figure 3(a) is the template of the benchmark, Fig. 3(b), Fig. 3(d) and Fig. 3(f) are the templates to be compared, Fig. 3(c), Fig. 3(e) and Fig. 3(g) show the similarity value between the template of the benchmark and the template to be compared by using SSIM algorithm. It can be seen that when the shapes of the two templates are similar, the value of SSIM is larger. On the contrary, it is smaller.

4 Experimental Validation Test

After the completion of tracking design, this paper exploited OTB100 to test, which includes 100 common object tracking videos. Tested algorithm are KCF_M (the algorithm proposed in this paper), KCF, CSK [3], MOSSE [9]. The evaluation indicators are the average accuracy and the average success rate used in the OTB data set. The comparison results are shown in Fig. 4:

In this paper, motion blur and occlusion are selected for analysis. It can be seen that the algorithm proposed in this paper is better than other classical filter-based tracking algorithms. From Fig. 4(a) and Fig. 4(c), it shows that in the

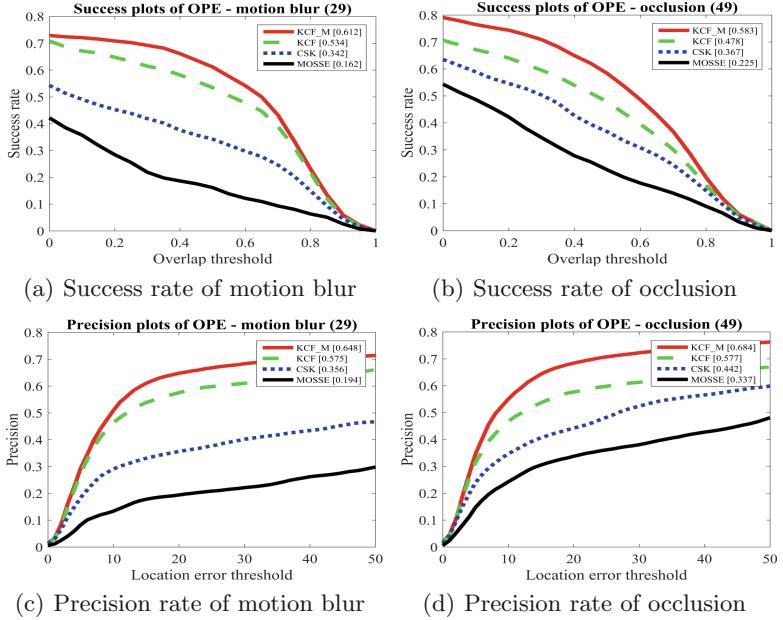


Fig. 4. Test results of OTB100 object tracking data set

motion blur section, the average success rate and average accuracy are increased by 14.6% and 12.7%, respectively, compared with the KCF algorithm. From Fig. 4(b) and Fig. 4(d), it is observed that in the occlusion part, the average success rate and average accuracy are increased by 22.0% and 18.5% respectively compared with the KCF algorithm. The above results show that the proposed algorithm has certain feasibility and advancement, and runs faster than 30FPS on the GTX1050 graphics card to meet the real-time requirements.

This paper established a simulation system of object tracking for UAV under the ROS platform, and tested the tracking algorithm. In this system, the UAV was modeled as four-rotator while flight control selected PX4 environment and tracked a four-wheeled car. The keyboard controlled the movement and turning of the car, and the quad-rotator could get image information from the camera on board. The tracking results are shown in Fig. 5:

It can be known that in the ROS simulation environment, the UAV can always track and move the car, and the car always stays in the center of vision of the UAV, which further proves the effectiveness of the algorithm in this paper.

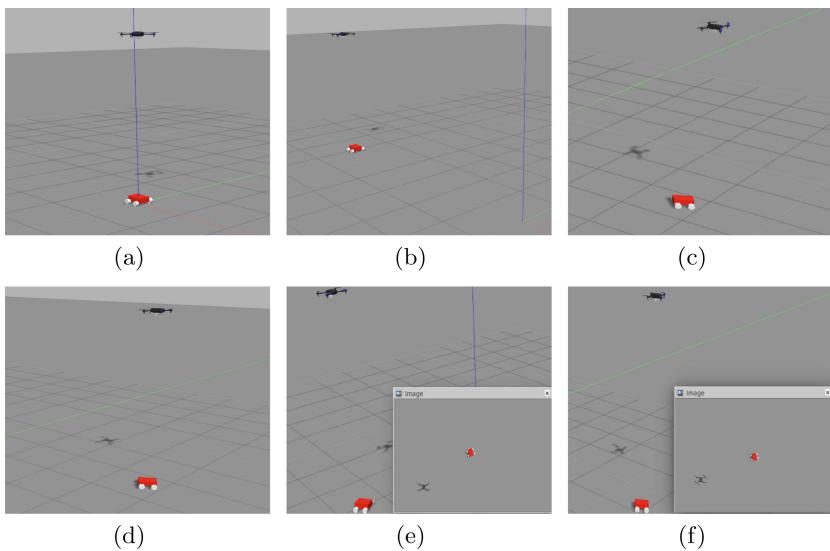


Fig. 5. Object tracking diagram in ROS simulation environment

5 Conclusions

- (1) This paper analyzes some characteristics of the kernel correlation filter tracking algorithm, and fuses the HOG features and CN features adaptively to make up for each other's shortcomings as to improve the object tracking accuracy and success rate.
- (2) Given that the template updating strategy of the general kernel correlation filtering algorithm is too simple, this paper combines the PSR index and similarity comparison in order to develop the template updating method to enhance the robustness of the system.
- (3) In this paper, we test the algorithm on the object tracking benchmark, the results indicate that the algorithm proposed in this paper is better than KCF in regards of motion blur and occlusion given an over 10% increase compared with KCF.
- (4) It can be concluded that the performance of algorithm proposed in this paper do better in deploying to UAV object tracking platform via establishing the simulation environment of UAV object tracking in ROS environment and testing the object tracking algorithm. As a result, this algorithm possesses good practical application value.

References

1. Sandino, J., Pegg, G., Gonzalez, F., et al.: Aerial mapping of forests affected by pathogens using UAVs, hyperspectral sensors, and artificial intelligence. *Sensors* **18**(4), 944 (2018)
2. Kanellakis, C., Nikolakopoulos, G.: Survey on computer vision for UAVs: current developments and trends. *J. Intell. Robot. Syst.* **87**(1), 141–168 (2017)
3. Wu, Y., Lim, J., Yang, M.H.: Object tracking benchmark. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(9), 1834–1848 (2015)
4. Coşkun, M., ünal, S.: Implementation of tracking of a moving object based on camshift approach with a UAV. *Procedia Technol.* **22**, 556–561 (2016)
5. Henriques, J.F., Caseiro, R., Martins, P., et al.: High-speed tracking with kernelized correlation filters. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(3), 583–596 (2014)
6. Cheng, H., Lin, L., Zheng, Z., et al.: An autonomous vision-based target tracking system for rotorcraft unmanned aerial vehicles. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1732–1738. IEEE (2017)
7. Wang, N., Shi, J., Yeung, D.Y., et al.: Understanding and diagnosing visual tracking systems. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3101–3109 (2015)
8. Li, Y., Zhu, J.: A scale adaptive kernel correlation filter tracker with feature integration. In: European Conference on Computer Vision, pp. 254–265. Springer, Cham (2014)
9. Bolme, D.S., Beveridge, J.R., Draper, B.A., et al.: Visual object tracking using adaptive correlation filters. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2544–2550. IEEE (2010)
10. Henriques, J.F., Caseiro, R., Martins, P., et al.: Exploiting the circulant structure of tracking-by-detection with kernels. In: European Conference on Computer Vision, pp. 702–715. Springer, Heidelberg (2012)



Develop Trajectory Tracking Controller for 4WS Vehicle

Koanhee Cho¹, Yue Ma^{1,2(✉)}, and Huimin Zhang¹

¹ Special Vehicle Lab, School of Mechanical Engineering,
Beijing Institute of Technology, Beijing 100081, China

yuema.bit@gmail.com

² Beijing Institute of Technology Chongqing Innovation Center,
Chongqing 401120, China

Abstract. This paper applies two trajectory tracking controllers for four-wheel independent steering system (4WS) vehicle. For increase the tracking performance those two controller use vehicle position and orientation as a controller input. And the control output is vehicle linear velocity and vehicle angular velocity. And each wheels steering angel and velocity is calculated from vehicle linear velocity and vehicle angular velocity. For the zero-sideslip steering maneuver, first calculate the instantaneous center of rotation (ICR) position, and using the ICR position the wheel steering angle is calculated. It makes the vehicle can put the ICR position in everywhere on the plant. And it means the vehicle is capable of carrying out all moves on a plane. The conformity determination is made using MATLAB/Simulink software.

Keywords: U-model · Backstepping method · Four-wheel independent steering vehicle · Trajectory tracking

1 Introduction

Since the rapid industrialization of the 20th century, the development of transport and transport means is accelerating. Vehicles using a variety of wheels (two to four, six, etc.) are currently being developed and practical, along with active research on mobile robots. Automatic Guided Vehicle (AGV) is a portable robot that follows path as like a marked long line or wire on the floor, or navigation using radio wave, vision camera, magnet or laser. And solve the transport heavy obstacle in factor or warehouse. For AGV there are some important part like mapping, path following and driving. In this paper for increase the performance of driving, will use four-wheel independent steering vehicle. Four-wheel independent steering (4WS) is a system what can increase vehicle stability while maneuvering at high speed, or to decrease turning radius at low speed. Theoretically a 4WS vehicle can put the instantaneous center of rotation(ICR) where ever on the plan. The idea of 4WS system is comes from 1907 Japan. For a vehicle for passenger the 4WS vehicle is not that attractive, but for the unmanned ground

vehicle(UGV) it has advantages in mobility. There was also many research for 4WS vehicle, for controller robust decoupling [1], PID [2], IMC controller [3], fuzzy selector controller [4], sliding-mode controller [5,6], Backstepping method controller [9], and Backstepping-sliding mode controller [10] are developed. Those controllers used to use single track vehicle model when they define vehicle steering angle. But single track vehicle model has problem when the ICR on vehicle coordinate X-axis, it cannot describe some vehicle motion. The wheel steering angle are 90° is same when the vehicle just move Y-axis direction on vehicle coordinate. And those controllers used to use steering angle for control input. [9] use vehicle position, orientation and sideslip angle as a control input, for increase the trajectory tracking performance. To solve the above problem, this paper proposes two controllers. The first one is U-model controller, and second one is Backstepping method controller and two controller use vehicle position and orientation value for controller variable. And the wheel steering angel designed based on 4WS vehicle instead of using single track vehicle model.

2 System Modeling

2.1 Basic Terminologies

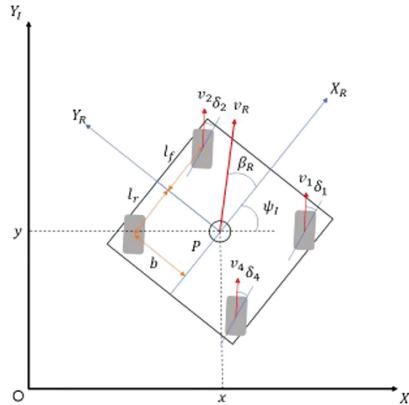


Fig. 1. Basic terminologies of 4WS vehicle

The vehicle basic terminologies as follow:

OX_IY_I : Global coordinate

PX_RY_R : Vehicle local coordinate

P : The center of gravity

v_R : Vehicle linear velocity in vehicle local coordinate

β_R : Vehicle sideslip angle

ψ_I : The heading angle of the vehicle.

Wheel number: From the front right wheel count up clockwise.

v_i : linear velocity of wheels CG

δ_i : The steering angle between $X_R - axis$ and wheel velocity

l_f : Distance between $Y_R - axis$ and front wheels

l_r : Distance between $Y_R - axis$ and rear wheels

b : Distance between $X_R - axis$ and each wheels

3 Controller Design

3.1 U-Model Controller Design

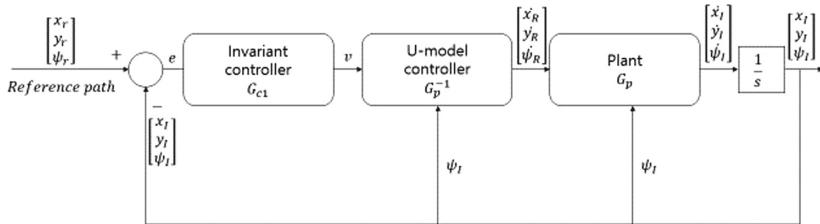


Fig. 2. Block diagram of U-model controller

The block diagram of U-model controller is shown in Fig. 2. For design U-model controller have to design two part, first one is U-block what is inverse of the plant. Second one is invariant controller G_{c1} .

The plant is same with the vehicle kinematic model:

$$\begin{bmatrix} \dot{x}_I \\ \dot{y}_I \\ \dot{\psi}_I \end{bmatrix} = \begin{bmatrix} \cos\psi_I & -\sin\psi_I & 0 \\ \sin\psi_I & \cos\psi_I & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x}_R \\ \dot{y}_R \\ \dot{\psi}_R \end{bmatrix} = \begin{bmatrix} \cos\psi_I & -\sin\psi_I & 0 \\ \sin\psi_I & \cos\psi_I & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_R \cos\beta_R \\ v_R \sin\beta_R \\ \dot{\psi}_R \end{bmatrix}$$

U-model controller defined inverse plant. So, in this case U-model controller as follow:

$$\begin{bmatrix} \dot{x}_R \\ \dot{y}_R \\ \dot{\psi}_R \end{bmatrix} = \begin{bmatrix} \cos\psi_I & -\sin\psi_I & 0 \\ \sin\psi_I & \cos\psi_I & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \dot{x}_I \\ \dot{y}_I \\ \dot{\psi}_I \end{bmatrix} = \begin{bmatrix} \cos\psi_I & \sin\psi_I & 0 \\ -\sin\psi_I & \cos\psi_I & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x}_I \\ \dot{y}_I \\ \dot{\psi}_I \end{bmatrix} \quad (1)$$

For invariant controller, PID controller is used. Each proportional, integral and differential gains are as follows:

$$\begin{aligned} k_P &= 20 \\ k_I &= 0.02 \\ k_D &= 0.001 \end{aligned}$$

Those gains are found empirical method. And the filter coefficient is 100.

For the stability of U-model controller, in perfect match, $G_p^{-1}G_p = 1$, it is achieved by specifying the close loop transfer matrix G stable, and the PID controller G_{c1} is also stable, so the U-model control is stable as well.

3.2 Backstepping Method Controller Design

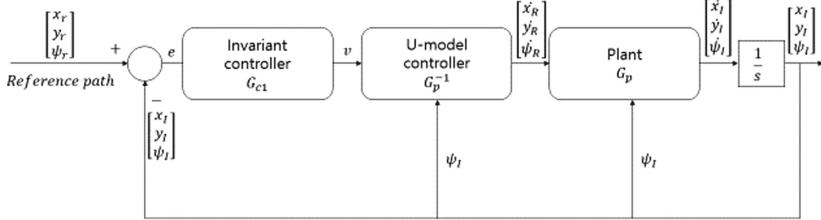


Fig. 3. Block diagram of Backstepping method controller

The block diagram of Backstepping method controller is shown in Fig. 3. The tracking error vector e can be defined on vehicle local coordinate as follow:

$$e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} \cos\psi_I & \sin\psi_I & 0 \\ -\sin\psi_I & \cos\psi_I & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_I - x_r \\ y_I - y_r \\ \psi_I - \psi_r \end{bmatrix} \quad (2)$$

To minimize these errors, a backstepping controller is designed. Using backstepping control method, a Lyapunov function is chosen as follow:

$$V = \frac{1}{2}e_1^2 + \frac{1}{2}e_2^2 + \frac{1}{2}e_3^2$$

The time derivative of chosen Lyapunov function is:

$$\dot{V} = e_1\dot{e}_1 + e_2\dot{e}_2 + e_3\dot{e}_3 \quad (3)$$

Time derivative of the tracking error vector is:

$$\dot{e} = \begin{bmatrix} \dot{e}_1 \\ \dot{e}_2 \\ \dot{e}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & e_2 \\ 0 & 1 & -e_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x}_I \\ \dot{y}_I \\ \dot{\psi}_I \end{bmatrix} - \begin{bmatrix} \cos\psi_I & \sin\psi_I & 0 \\ -\sin\psi_I & \cos\psi_I & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x}_r \\ \dot{y}_r \\ \dot{\psi}_r \end{bmatrix} \quad (4)$$

Substituting Eq. 4 in Eq. 3, obtain follow equation:

$$\dot{V} = e_1(\dot{x}_I - \dot{x}_r \cos\psi_I - \dot{y}_r \sin\psi_I) + e_2(\dot{y}_I + \dot{x}_r \sin\psi_I - \dot{y}_r \cos\psi_I) + e_3(\dot{\psi}_I - \dot{\psi}_r) \quad (5)$$

When the time derivative of Lyapunove function is smaller then 0 ($\dot{V} \leq 0$), the system is stable. So, in this case the control law is chosen as follows:

$$U = \begin{bmatrix} \dot{x}_I \\ \dot{y}_I \\ \dot{\psi}_I \end{bmatrix} = \begin{bmatrix} -k_1 e_1 + \dot{x}_r \cos\psi_I + \dot{y}_r \sin\psi_I \\ -k_2 e_2 - \dot{x}_r \sin\psi_I + \dot{y}_r \cos\psi_I \\ -k_3 e_3 + \dot{\psi}_r \end{bmatrix} \quad (6)$$

Substituting Eq. 6 in Eq. 5:

$$\dot{V} = -k_1 e_1^2 - k_2 e_2^2 - k_3 e_3^2$$

So, when the k_1, k_2 and k_3 are bigger then 0, the system is stable. Here $k_i = 1$.

Finally, the control law of backstepping method controller can describe Eq. 6

4 Simulation and the Results

The reference inputs are as follows:

From 0 s to 2 s the vehicle moves $+X_I$ direction with 0.5 m/s^2 acceleration, till 8 s the vehicle moves $+X_I$ direction with 1 m/s velocity, till 10 s the vehicle moves $+X_I$ direction with -0.5 m/s^2 acceleration, next 1 s the vehicle stop for steering to move $+Y_I$ direction, till 13 s the vehicle moves $+Y_I$ direction with 0.5 m/s^2 acceleration, till 20 s the vehicle moves $+Y_I$ direction with 1 m/s velocity, and final 10 s the vehicle moves zero-sideslip steering maneuver with 1 m/s velocity and -1 rad/s angular velocity.

The initial condition of each variable is as follows:

Initial wheel steering angle	$\delta_i(0)$	0 rad
Initial wheel linear velocity	$v_i(0)$	0 m/s
Initial sideslip angle	$\beta_R(0)$	0 rad
Initial vehicle orientation angle	$\psi_I(0)$	0 rad
Distance between $Y_R - axis$ and front wheels	l_f	0.5 m
Distance between $Y_R - axis$ and rear wheels	l_r	0.5 m
Distance between $X_R - axis$ and each wheels	b	0.5 m

4.1 U-Model Controller Simulation Result

Figure 4 shows the reference $X_I - axis$, $Y_I - axis$ and ψ trajectories. The controller result is good, as Fig. 5 show the biggest errors of each value is 0.05 m in X-axis, 0.05 m in Y-axis, and 0.05 rad in vehicle orientation angle.

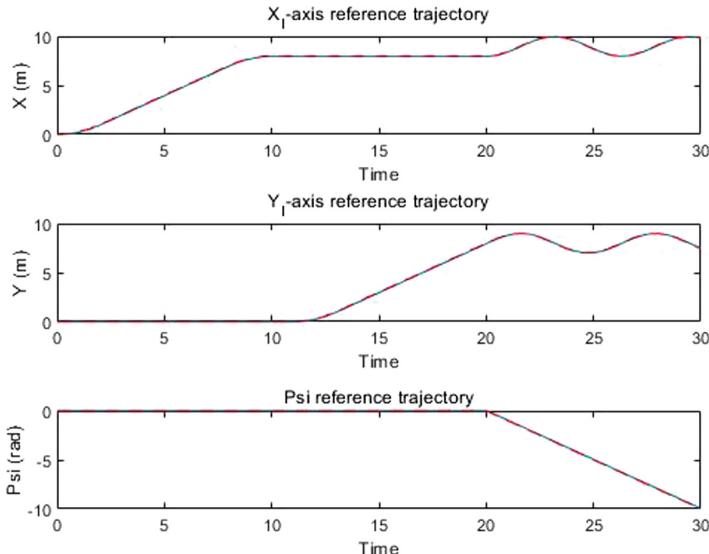


Fig. 4. U-model controller reference trajectory result

4.2 Backstepping Method Controller Simulation Result

Figure 6 shows the reference $X_I - axis$, $Y_I - axis$ and ψ trajectories. The controller result is good, as Fig. 7 show the errors are too small. It can be regards as 0.

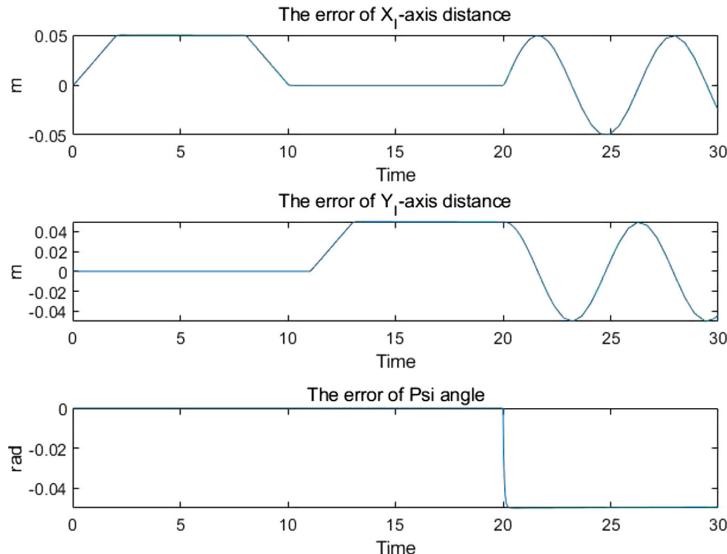


Fig. 5. U-model controller trajectory errors result

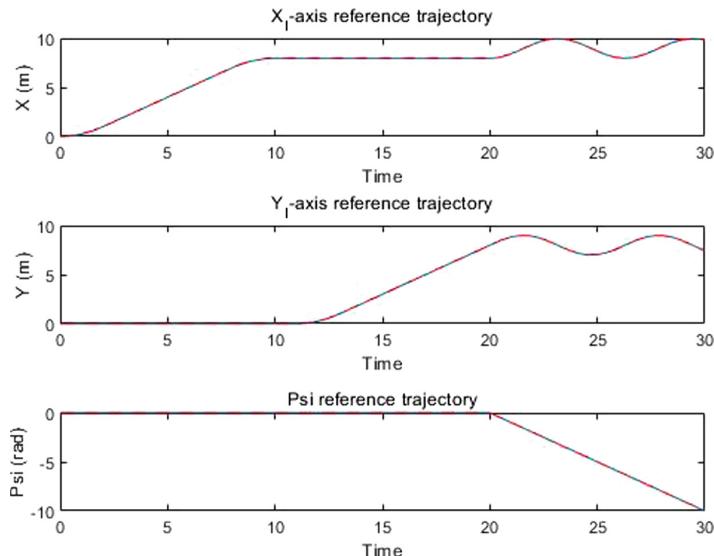


Fig. 6. Backstepping method controller reference trajectory result

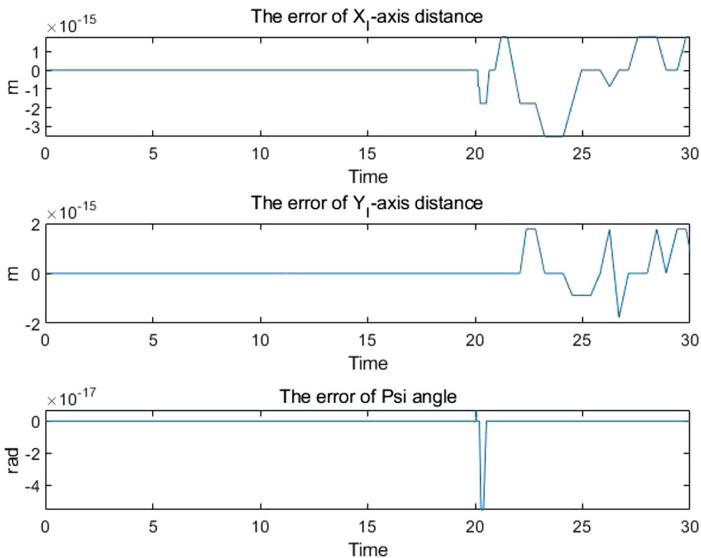


Fig. 7. Backstepping method controller trajectory errors result

5 Conclusions

This paper presented two 4WS vehicle controller for track reference trajectories. A 4WS vehicle kinematic modeling is designed based on real 4WS vehicle structure for let the vehicle able rotation with Y_R velocity. Based on kinematic modeling, two controller designed. The first controller is U-model controller, and second one is Backstepping method controller based on Lyapunov stability. Those controller simulation using MATLAB/Simulink program to verify the effectiveness and the performance of the 4WS vehicle. The trajectory has both steering maneuvers, the first one is zero-sideslip steering maneuver what the sideslip angle is not changed during the vehicle moves. And the second steering maneuver is parallel steering maneuver what the four wheel move same direction and same velocity. Those two controller tracked reference path very well. The U-model controller errors are bigger then Backstepping method controller, the U-model controller errors are 0.05 m in $X_I - axis$, 0.05 m in $Y_I - axis$, and 0.05 rad in vehicle orientation angle. It was shown that the designed two controllers can make the 4WS vehicle track the reference path using zero-sideslip steering maneuver and parallel steering maneuver very well.

References

1. Ackermann, J.: Robust decoupling, ideal steering dynamics and yaw stabilization of 4WS cars. *Automatica* **30**(11), 1761–1768 (1994). <https://doi.org/10.1109/icmsao.2013.6552547>

2. Lam, T.L., Qian, H., Xu, Y.: Behavior-based steering control for four wheel independent steering vehicle. In: 2008 IEEE International Conference on Robotics and Biomimetics (2009). <https://doi.org/10.1109/robio.2009.4913059>
3. Canale, M., Fagiano, L.: Stability control of 4WS vehicles using robust IMC techniques. *Veh. Syst. Dyn.* **46**(11), 991–1011 (2008). <https://doi.org/10.1080/00423110701790723>
4. Yang, Z., Wang, Z., Su, W., Zhang, J.: Multi-mode control method based on fuzzy selector in the four wheel steering control system. In: IEEE ICCA 2010 (2010). <https://doi.org/10.1109/icca.2010.5524179>
5. Solea, R., Filipescu, A., Minzu, V., Filipescu, S.: Sliding-mode trajectory-tracking control for a four-wheel-steering vehicle. In: IEEE ICCA 2010 (2010). <https://doi.org/10.1109/icca.2010.5524422>
6. Hamzah, N., Aripin, M.K., Sam, Y.M., Selamat, H., Ismail, M.F.: Yaw stability improvement for four-wheel active steering vehicle using sliding mode control. In: 2012 IEEE 8th International Colloquium on Signal Processing and Its Applications (2012). <https://doi.org/10.1109/cspa.2012.6194704>
7. Selekwa, M.F., Nistler, J.R.: Path tracking control of four wheel independently steered ground robotic vehicles. In: IEEE Conference on Decision and Control and European Control Conference (2011). <https://doi.org/10.1109/cdc.2011.6160677>
8. Amdouni, I., Jeddi, N., El Amraoui, L.: Optimal control approach developed to four-wheel active steering vehicles. In: 2013 5th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO) (2013). <https://doi.org/10.1109/icmsao.2013.6552547>
9. Setiawan, Y.D., Nguyen, T.H., Pratama, P.S., Kim, H.K., Kim, S.B.: Path tracking controller design of four wheel independent steering automatic guided vehicle. *Int. J. Control Autom. Syst.* **14**(6), 1550–1560 (2016). <https://doi.org/10.1007/s12555-015-0216-7>
10. Tu, X., Gai, J., Tang, L.: Robust navigation control of a 4WD/4WS agricultural robotic vehicle. *Comput. Electron. Agric.* **164**, 104892 (2019). <https://doi.org/10.1016/j.compag.2019.104892>
11. Zhu, Q.M., Guo, L.Z.: A pole placement controller for non-linear dynamic plants. *Proc. Inst. Mech. Eng. Part I: J. Syst. Control Eng.* **216**(6), 467–476 (2002). <https://doi.org/10.1177/095965180221600603>
12. Zhu, Q.M., Zhao, D.Y., Zhang, J.: A general U-block model-based design procedure for nonlinear polynomial control systems. *Int. J. Syst. Sci.* **47**(14), 3465–3475 (2015). <https://doi.org/10.1080/00207721.2015.1086930>
13. Zhu, Q., Zhang, W., Na, J., Sun, B.: U-model based control design framework for continuous-time systems. In: 2019 Chinese Control Conference (CCC) (2019). <https://doi.org/10.23919/chicc.2019.8866624>
14. Geng, X., Zhu, Q., Liu, T., Na, J.: U-model based predictive control for nonlinear processes with input delay. *J. Process Control* **75**, 156–170 (2019). <https://doi.org/10.1016/j.jprocont.2018.12.002>
15. Zhu, Q., Weicun, Z., Zhang, J., Sun, B.: U-neural network-enhanced control of nonlinear dynamic systems. *Neurocomputing* (2019). <https://doi.org/10.1016/j.neucom.2019.04.008>



Design of Fatigue Driving Detection Algorithm Based on Image Processing

Shangzheng Liu^{1(✉)}, Yalei Wu², Qunpo Liu^{2(✉)}, and Qinghui Zhu¹

¹ Nanyang Institute of Technology,
Nanyang 473004, Henan Province, People's Republic of China
lszfrank@126.com

² School of Electrical Engineering and Automation, Henan Polytechnic University,
Jiaozuo 454000, Henan Province, People's Republic of China
lqpy@hpu.edu.cn

Abstract. One of the primary causes of traffic accidents is drowsy of the drivers involved. A warning system about drowsy status (fatigue or drowsiness) of the driver, helping to limit the traffic accidents caused by falling asleep behind the wheel by determining the status of the eyes combined with the status of mouth. In this paper, we present a novel approach for determining the facial landmarks. Besides, we proposed an algorithm to detect and identify the status of eye and mouth. Testing results confirms the effectiveness and feasibility of the proposed algorithm.

Keywords: Fatigue detection · Fatigue driving · Image processing

1 Introduction

Driving safety is influenced by two key factors: 1) the mental states of the driver 2) the external environment. Various psychological conditions, e.g., fatigue, distraction and motion sickness can affect driving safety.

Fatigue driving has a close relation with traffic accidents such as collision, personal injury and so on. Investigations show that up to 37% of all vehicle fatalities involve a fatigue driving or drowsy driving in the United States [1,2]. In Europe, drowsy driving contributes to one fourth to one third of road accidents [3]. Some researchers found that fatigue driving has been estimated to be involved in 2% to 23% of all crashes [4].

The degradation in driving performance because of fatigue accounts for a small, but significant, percent of highway accidents [5,6]. The growing number of accidents caused by fatigue driving has been becoming a serious societal problem in recent years. Therefore, it is of great significance to detect fatigue driving and to develop an efficient intervention for driving mental fatigue not only to reduce the social cost of traffic safety but also to improve the health of drivers and passengers.

Fatigue driving detection measures include subjective evaluation method and objective evaluation method. The objective method can be categorized into: 1) evaluation of biomedical signals including pulse rate, EEG, changes of head position, eye-closure rate, and eyelid movement; 2) evaluation of driver–vehicle data, including steering angle, throttle/brake input, and speed. The major drawback of the techniques based on biomedical signals is that they require directly placing sensors on the driver’s body. This will make the driver feel uncomfortable. Methods based on driver-vehicle data are prone to be affected by driving habits of drivers as well as size and shape of vehicles.

2 Face Landmark Detection Based on Classification and Regression Tree

2.1 Ensemble of Regression Trees

Ensemble of regression trees (ERT) uses gradient boosting regression tree to detect face landmark positions. It constructs a gradient boosting decision tree (GBDT). Using residual, facial landmarks gradually return to the real shape. GBDT is a cascade regression tree. In GBDT, every leaf has a residual. When an input reaches to one node of regression tree, residual will be added to the input. This will rectify the regression. Finally, summation of residuals is feed to image. The landmarks of face are found.

Let $x_i \in R^2$ be the x,y-coordinates of the i th facial landmark in an image I . Then the vector $C = (X_1^T, X_2^T, \dots, X_n^T)^T \in R^{2n}$ denotes the coordinates of all the n facial landmarks in I . $\hat{C}^{(t)}$ denotes current estimate of C . Regressor $r_t(.,.)$ is added to the current shape estimate $\hat{C}^{(t)}$ to improve the estimate:

$$\hat{C}^{(t+1)} = \hat{C}^{(t)} + r_t(I, \hat{C}^{(t)}) \quad (1)$$

For the second regressor, gradient tree boosting is used to obtain a series of regression tree. Assume we have training data $(I_1, C_1), \dots, (I_n, C_n)$ where each I_i denotes a face image and C^i its shape vector. Let $\hat{C}_i^{(0)}$ is the initial shape, $\Delta\hat{C}_i^{(0)} = C^i - \hat{C}_i^{(0)}$ denotes the residual between real shape and initial shape of sample [7,8].

Algorithm 1 Learning r_t in the cascade

Step 1: Training data is $\{(I_i, \hat{C}_i^{(t)}, \Delta C_i^{(t)})\}_{i=1}^N, v$ is learning rate, $0 < v < 1$. Initialise

$$f_0(I, \hat{C}^{(t)}) = \arg \min \sum \|\hat{C}_i^{(t)} - r\|, \quad (2)$$

Here $f_0(I, \hat{C}^{(t)})$ is a regression tree.

Step 2: for $k = 1, \dots, K$, set for $i = 1, \dots, N$

$$r_{ik} = \Delta C_i^{(t)} - f_{k-1}(I_i, \hat{C}_i^{(t)}) \quad (3)$$

Update $f_k(I_i, \hat{C}_i^{(t)})$ as the following

$$f_k(I_i, \hat{C}_i^{(t)}) = f_{(k-1)}(I_i, \hat{C}_i^{(t)}) + vr_{ik} \quad (4)$$

Step 3: Output $r_t(I, \hat{C}^{(t)}) = f_k(I, \hat{C}(t))$. Inorder to avoid the over fitting, Shrinkage is introduced to GBDT. Shrinkage can significantly improve the generalization of GBDT, and improve precision of detection data. It can be expressed as follows.

$$F_{(m+1)}(x) = F_{(m)}(x) + v \cdot \gamma_{(m+1)} \cdot h_{(m+1)}(x) \quad (5)$$

Here v is learning rate, $v \in [0, 1]$. Generally, v is small, accordingly learning rate is slow. v is usually set to 0.1.

Extract Landmark. Using ERT to extract face landmark as follows.

Step 1: Collect sample data, label landmark.

Step 2: Based on labeled sample, generate xml file which save the coordinates of landmark.

Step 3: Feed the xml file and image to model, training model.

Step 4: Test the trained model. If the performance doesn't reach the expected, go to next.

Step 5: Adjust parameters, training the model.

Step 6: Choose the best model as output.

Choose Landmark. When training model, the total number of landmark is determined firstly, then landmark is labeled. Here the number of landmark is 68. Eyes, mouth, nose and shape of face are labeled as the following figure. Next, we will label the training data. After each image is labeled, each image is represented by a rectangle as (Fig. 1):

$$c = (a_1, a_2, \dots, a_n)^T, \quad (6)$$



Fig. 1. Facial landmark.

here $a_i = (x_i, y_i)^T$ denotes the coordinates of the i th landmark. Landmark should meet the followings:

1. Landmark should reflect the shape of face.
2. Landmark should bilateral symmetry within face.
3. The labeled landmark of training set should be consistency.

Settings of model are:

Tree depth: 4.

Cascade dept: 15.

Learning rate (nu): 0.1. Nu is regular terms.

3 Detection and Judgement of Fatigue

3.1 Judge Fatigue Based on Eye Dynamic Character

When eye is open, distance between top and bottom landmark of eye is large. Otherwise, eye is closed, the distance is small, as shown in Fig. 2. Using ERT, we can locate the landmark of eye. Then, using the landmark, we can obtain the width and length of eye. Next, we can calculate the ratio of width and length (EAR). EAR denotes the state of eye. When eye is open, EAR is large. Otherwise, EAR is small. EAR is calculated as following.

$$EAR = \frac{\|P_2 - P_6\| + \|P_3 - P_5\|}{2 \|P_1 - P_4\|} \quad (7)$$

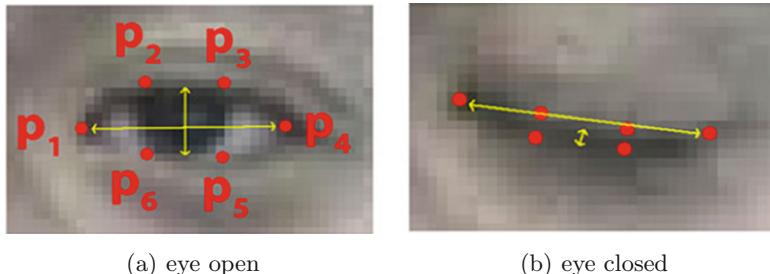


Fig. 2. Open and closed eye

Landmarks of eye are $P_1, P_2, P_3, P_4, P_5, P_6$, as shown in Fig. 2. $\|P_1 - P_4\|$ denotes length of eye. $\|P_2 - P_6\|$ and $\|P_3 - P_5\|$ denotes width of eye. Average of two eye's EAR conforms to reality. According to the analysis of eye's state, EAR is 0.3, when eye is open, when eye is closed, EAR is 0.15. Normally, EAR is about 0.17. Based on analysis, 0.23 is set as threshold to judge the state of eye. As EAR is less than 0.23, eye is closed. AS in normal state, human being

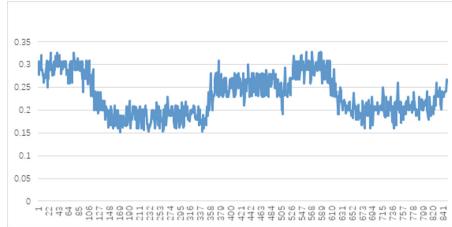


Fig. 3. EAR of eye.

will blink 15 every minute. And duration of blink is less than 1 s. When human driver is fatigue, duration of blink will prolong, number of blink also increases. And duration of blink will be greater than 1.5 s. Number of blink will increase to 20 times every minute. Based on these facts, threshold of blink is set to 20 times every minute. When the number of blink is greater than 20, we can judge the subject is in fatigue state (Fig. 3).

According to the forgoing analysis, we have the following rules to judge whether a driver is in fatigue or not.

1. Judge cycle of eye closure. When the ratio of eye closure and period is greater than 80%, subject is in fatigue state.
2. Judge number of blinking. When the number of blinking is greater than 20 in a period, subject is in fatigue state.

3.2 Judge Fatigue Based on Yawning

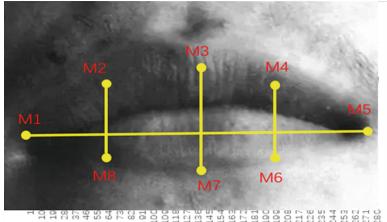
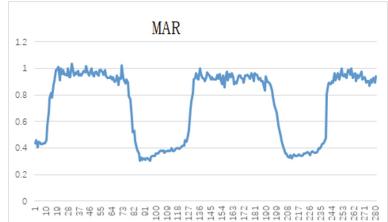
When drivers fatigue, they will have series of behavior reactions such as eyes closed or yawning. Therefore, driver fatigue can be calibrated by calculating the open degree of mouth (MAR). Like EAR, MAR can be calculated as the following equation.

$$MAR = \frac{\|M_2 - M_8\| + \|M_3 - M_7\| + \|M_4 - M_6\|}{3 \|M_1 - M_5\|}, \quad (8)$$

M Landmarks of eye are $M_1, M_2, M_3, M_4, M_5, M_6, M_7, M_8$, as shown in Fig. 4. $\|M_1 - M_5\|$ denotes length of mouth. $\|M_2 - M_8\|$, $\|M_3 - M_7\|$ and $\|M_4 - M_6\|$ denote width of mouth. When a driver in normal state, MAR is about 0.4. When a driver is yawning, MAR will increase to 1, as shown in Fig. 5. Based on analysis of simulation results, MAR threshold is set to 0.75, and the lasting time threshold is set to 2 s. The over all system diagram is shown in Fig. 6.

4 Experimental Results

The proposed algorithm is tested on 1800 videos, and there were 21236 faces in them. 1200 blinking images, 1200 yawning images and 3600 normal images are selected as testing data.

**Fig. 4.** Landmark of mouth**Fig. 5.** Simulation results of MAR

4.1 Detection of Facial Landmark

Table 1 shows the eye detection results in two conditions. Figure 7 shows the contours of eye wearing glasses and without glasses.

Table 1. Eye detection results

Total number	Leak number	Wrong number	Accuracy %
300	18	0	94

4.2 Detection of Blink

Table 2 shows detection of blink frequency and cycle. Table 2 shows as the frequency of blink is less than 20 times/minute, output of the proposed system is normal. When the frequency of blink is greater than 20 times/minute, output of the proposed system is fatigue as shown in Fig. 8. At the same time, alarm go off.

Table 2. Detection results of blink

Total number	Total time	Frequency (times/minute)	Testing result	Real state
13	60.55	13	Normal state	Normal state
22	26.6	49.62	Fatigue	Fatigue

4.3 Detection of Yawn

Table 3 shows simulation results of yawning. When the MAR is greater than 0.75 and the duration is greater than 2 s, subject is yawning in sleepy state as shown in Fig. 9.

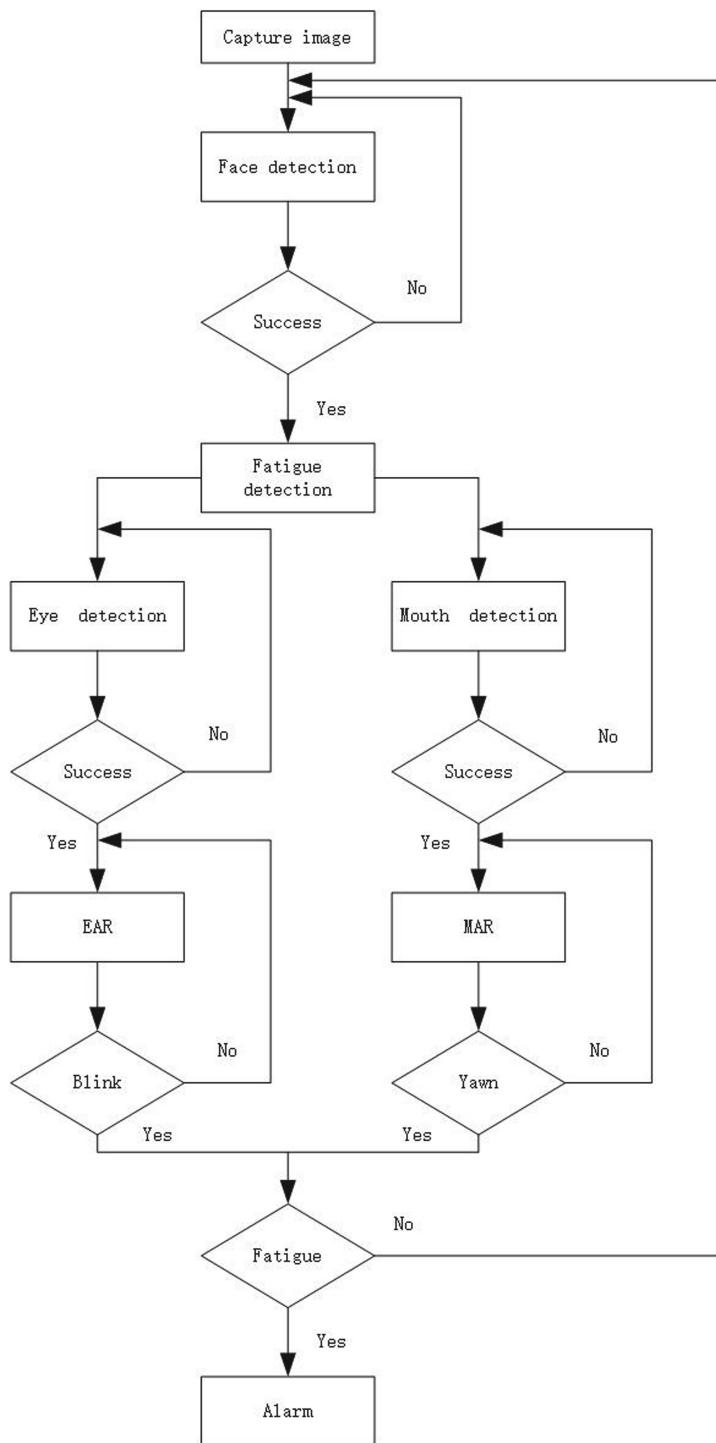


Fig. 6. System diagram



(a)withoutglasses (b) with glasses

Fig. 7. Contour of eye

(a) Normal State (b) Sleepy State

Fig. 8. Simulation results of blink**Table 3.** Detection results of yawn

MAR	Duration time	Judgement	Testing result	Real state
0.67	3.85	NO	Normal state	Normal state
1.02	3.55	YES	Fatigue	Fatigue



(a) Normal State (b) Sleepy State

Fig. 9. Simulation results of yawn

5 Conclusions

This paper designs a driver fatigue recognition algorithm based on eye state and mouth state. Firstly, we take GBDT to detect facial landmark positions. Then, we judge the eye state with the method based on EAR and average eyes blinking time. In order to improve the reliability, we also use MAR and average mouth open time to judge mouth state. The experimental results show the proposed algorithm can judge human eye state and mouth state accurately. As to self-adaption, drivers with smaller eyes is step further research content.

References

1. Summala, H., Mikkola, T.: Fatal accidents among car and truck drivers: effects of fatigue, age, and alcohol consumption. *Hum. Factors* **36**(2), 315–326 (1994). <https://doi.org/10.1177/001872089403600211>
2. Smith, P., Shah, M., da Vitoria Lobo, N.: Determining driver visual attention with one camera. *IEEE Trans. Intell. Transp. Syst.* **4**(4), 205–218 (2003). <https://doi.org/10.1109/TITS.2003.821342>
3. Kingshuk, M., Rakesh, K., Souvik, D.: Effective estimation of driver drowsiness based on eye status detection and analysis. In: 2014 International Conference on Devices, Circuits and Communications (ICDCCom) (2014). <https://doi.org/10.1109/ICDCCom.2014.7024717>
4. Zhitao, X., Zhiqiang, H., Lei, G., Jun, W., Yuelong, L.: Fatigue driving recognition network: fatigue driving recognition via convolutional neural network and long short-term memory units. *IET Intell. Transp. Syst.* **13**(9), 1410–1416 (2019). <https://doi.org/10.1049/iet-its.2018.5392>
5. Swapnali, K., Masoud, A., Ehsan, R.: A data-driven model to identify fatigue level based on the motion data from a smartphone. In: 2019 IEEE Western New York Image and Signal Processing Workshop (WNYISPW) (2019). <https://doi.org/10.1109/WNYIPW.2019.8923100>
6. Zhao, L., Duan, N., Yang, C.: Fatigue-driving recognition based on the state of the human eye and electroencephalographic signals. In: 2016 Chinese Control and Decision Conference (CCDC), pp. 1567–1572. IEEE Press, New York (2016). <https://doi.org/10.1109/HPDC.2001.945188>
7. Jia, Y.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000). <https://doi.org/10.1109/87.845885>
8. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003). <https://doi.org/10.1109/TAC.2003.815033>



Optimal Design for Active Vibration Control of Aero-engines

Ting Fang and Jiqiang Wang^(✉)

Jiangsu Province Key Laboratory of Aerospace Power Systems,
Nanjing University of Aeronautics and Astronautics, Nanjing 210016, Jiangsu, China
jiqiang.wang@nuaa.edu.cn

Abstract. Passive vibration isolation of aero-engines is often problematic associated with heavy weight and poor adaptability. The purpose of this paper is to disseminate an optimal geometric design approach with the following procedures: 1) install an actuator at the engine casing; 2) design an active vibration damping system; 3) establish a mathematical model of engine vibration transmission path, and 4) design a feedback control system with the control method of “geometric design” to reduce vibration. The active vibration control system model was built and simulated by Matlab/Simulink software. The results show that the proposed method can effectively improve the vibration performance at the casing.

Keywords: Aero-engine · Actuators · Active vibration control

1 Introduction

As a key component of the aircraft, the aero-engine is a complex thermal-mechanical-fluid machinery with high rotational speed. Therefore, vibration inevitably occurs during engine operation. Indeed, more than 90% of structural strength failures of engines are related to vibration [1]. In recent years, with the continuous development of China’s aerospace technology, the problem of engine vibration has gradually been paid attention to. The reduction of engine vibration level has always been an urgent problem to be solved in the field of engine technology. For a long time, the industry has widely adopted passive methods to suppress vibration. It does not require external energy input to provide control force. Instead, it uses structural damping and other passive technologies to absorb and consume vibration energy. However, they are lack of flexibility and henceforth unable to effectively adjust themselves in a changing environment, especially for vibration problems in the low frequency range [2–4]. Therefore, active vibration control has gradually become a research hotspot in the field of vibration.

At present, extensive research has been conducted in the field of active vibration control. The United States has effectively combined active vibration control

technology with noise monitoring systems and successfully applied it to the latest nuclear submarines, enabling submarines to effectively avoid sound detection [5]; Australia has also applied this technology in the military field. Assembling this technology into a submarine can effectively suppress the 6 degree-of-freedom vibration in the double-layer vibration isolation device [6]. Most of the domestic development of vibration control systems are based on traditional software such as MATLAB/Simulink or dSPACE. Liu Zongwei used engine speed as the initial reference signal and built an adaptive active noise control model based on Simulink toolbox for simulation [7]. Using dSPACE as a platform, Wang Shiyao built a double-layer vibration isolation model to conduct semi-physical simulation of active vibration control, and successfully verified the feasibility of hardware-in-the-loop research. Other advanced methods are also proposed for handing robustness for suppressing vibration, see and references therein [8–10].

The main contents of this paper are as follows:

1. the mathematical model of the vibration transmission system of aero-engine is built in Matlab/Simulink.
2. a new control algorithm-“Geometric Design Method” is used to control the vibration of the model, and a compare with the PID control is carried out to verify the effectiveness of the new method.

2 Aero-engine Vibration Model

2.1 Shaker Model

The shaker is mainly composed of a magnetic flux telescopic rod (GMM), a top rod, a support spring, a damper, a housing and a base. The working principle of the shaker is: alternating current flows into the coil to produce a changing electromagnetic field. The interaction of the magnetic field and the flux rod drives the ejector rod to vibrate back and forth, completing the complex electrical energy-magnetic field energy-mechanical energy mutual conversion process. The schematic diagram of the shaker system is shown in Fig. 1:

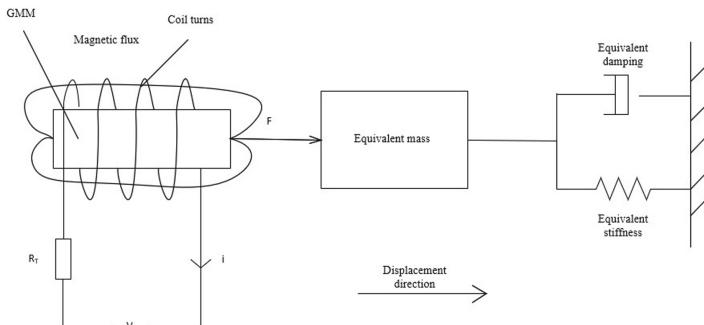


Fig. 1. Schematic diagram of the shake system

According to the working principle of the shaker system, its mathematical model is expressed as follows:

$$V = iR + L \frac{di}{dt} \quad (1)$$

where: V is the input voltage; R is the drive coil resistance; i is the current; L is the inductance; t is the time.

The shaker magnetomotive force:

$$F_1 = NI + \frac{x}{d_{33}} \quad (2)$$

where, x is the output displacement of the shaker; d_{33} is the axial dynamic magnetostriction coefficient of the telescopic rod.

The shaker magnetic flux:

$$\Phi = \frac{F_1}{R_T} = \frac{NI + x/d_{33}}{R_T} \quad (3)$$

where: R_T is the total reluctance of the shaker.

The stretching force of GMM shall satisfy the following relation:

$$F = \frac{\Phi}{d_{33}} \quad (4)$$

According to Newton's second law:

$$F = M_T \frac{d^2x}{dt^2} + C \frac{dx}{dt} + Kx \quad (5)$$

In the formula: GMM's equivalent stiffness coefficient $K = (AE)/l$

where A, E, L are the cross-sectional area, elastic modulus and length of the telescopic rod. Change (5) with Laplace to obtain the functional relationship between F and x:

$$G(S) = \frac{x(S)}{F(S)} = \frac{\omega_n^2/K}{S^2 + 2\zeta\omega_n S + \omega_n^2} \quad (6)$$

where: ω_n is the natural frequency, $\omega_n = \sqrt{\frac{K}{M_T}}$, $\zeta = \frac{C}{2\sqrt{KM_T}}$.

From Eqs. (1)–(6), the transfer function of the shaker can be obtained, so that the transfer function from current to displacement is:

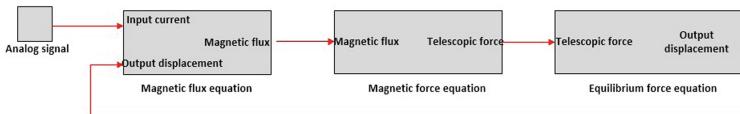
$$G_I(S) = \frac{x(S)}{I(S)} = \frac{Nd_{33}\omega_n^2}{KR_T d_{33}^2(S^2 + 2\zeta\omega_n S + \omega_n^2) - \omega_n^2} \quad (7)$$

According to the above formula, use Simulink to construct the block diagram and mathematical model of the system model of the shaker and the Parameter value table of shaker system is shown in Table 1:

In the follow-up control design of this article, two shakers were used respectively, with one shaker as a stimulus while the other as an actuator. The system model block diagram of the shaker is shown in Fig. 2:

Table 1. Parameter value table of shaker system

The parameters of the shaker system	Symbol	Value
The coil number of turns/(turns)	N	200
GMM piezomagnetic coefficient/(m/A)	d_{33}	0.00000001
Total reluctance/(H)	R_T	58000000
Equivalent stiffness coefficient/(N/m)	k	176000000
Natural frequency/(Hz)	$\omega_n/2\pi$	1670
Damping ratio/	ζ	0.26

**Fig. 2.** Shaker system model block diagram

2.2 Vibration Transmission Model

In this section, a simplified model of the vibration transmission path from the rotor to the hanger part of the installation section will be established based on physical principles. In order to clearly establish the vibration transmission path model, it will establish the “rotor unbalance signal (input) -inner casing-outer casing-installation section (sensor output)” system from the rotor to the installation section. In order to highlight the necessary structure of the vibration control system, the inner casing and the outer casing are simplified into two rigid bodies with mass-spring-dampers. This system is constructed according to the feedback principle of the control theory, and is composed of a rigid body, an acceleration sensor, a controller and an actuator. The working principle of the system is shown in Fig. 3.

In the picture, m_1 and m_2 are the quality of the inner and outer casings; k_1 and k_2 are the stiffness of the supporting springs; c_1 and c_2 are the damping coefficients; y_1 represents the displacement when the rotor vibrates, y_2 represents the displacement of the inner casing, and y_3 represents the displacement of the outer casing; F_t indicates the excitation force of the source; F_p indicates the main power generated by the actuator.

The vibration transmission path is composed of the above parts.

According to Newton's second law, the dynamic analysis of the vibration transmission system, the motion differential equation of the system is:

$$\begin{cases} m_2\ddot{y}_3 - c_2(\dot{y}_2 - \dot{y}_3) - k_2(y_2 - y_3) - F_p = 0 \\ m_1\ddot{y}_2 + c_2(\dot{y}_2 - \dot{y}_3) + k_2(y_2 - y_3) - c_1(\dot{y}_1 - \dot{y}_2) - k_1(y_1 - y_2) = 0 \\ k_1y_1 + c_1\dot{y}_1 = F_t \end{cases} \quad (8)$$

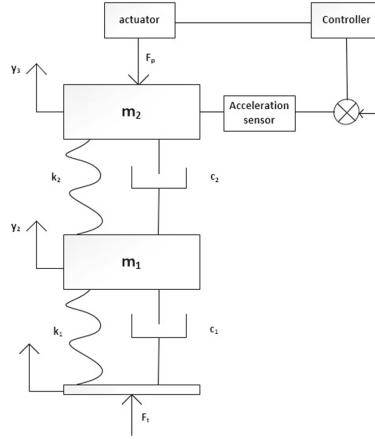


Fig. 3. Schematic diagram of active vibration control system

Laplace transform on Eq. (8) gives the following equation:

$$(s^2 \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} + s \begin{bmatrix} c_1 + c_2 - c_2 \\ -c_2 & c_2 \end{bmatrix} + \begin{bmatrix} k_1 + k_2 - k_2 \\ -k_2 & k_2 \end{bmatrix}) Y_{(s)} = F_{(s)} \quad (9)$$

In this way, the transfer relationship between the outer casing displacement $Y_{3(s)}$ and the main power $F_{p(s)}$.

The outer casing displacement $Y_{3(s)}$ and the excitation force $F_{p(s)}$ are obtained:

$$\begin{aligned} G_m(s) &= \frac{Y_{3(s)}}{F_{p(s)}} \\ &= \frac{m_2 s^2 + c_1 s + c_2 s + k_1 + k_2}{m_1 m_2 s^4 + (c_1 m_2 + c_2 m_1 + c_2 m_2) s^3 + (k_1 m_2 + k_2 m_1 + k_2 m_2 + c_1 c_2) s^2 + (c_1 k_2 + c_2 k_1) s + k_1 k_2} \end{aligned} \quad (10)$$

$$\begin{aligned} G_r(s) &= \frac{Y_{3(s)}}{F_t(s)} \\ &= \frac{c_2 s + k_2}{m_1 m_2 s^4 + (c_1 m_2 + c_2 m_1 + c_2 m_2) s^3 + (k_1 m_2 + k_2 m_1 + k_2 m_2 + c_1 c_2) s^2 + (c_1 k_2 + c_2 k_1) s + k_1 k_2} \end{aligned} \quad (11)$$

The parameters of the vibration transfer model are shown in Table 2:

3 Active Vibration Control System Design and Simulation

3.1 Geometric Design Algorithm

This article presents a new control algorithm-geometric design method. The selection range of the controller is reduced to the intersection of two circles by

Table 2. Relevant parameters of vibration transmission path

Active control system parameters	Symbol	Value
Inner casing quality/(kg)	m_1	546
Outer casing quality/(kg)	m_2	546
Stiffness of rotor-inner casing/(N/m)	k_1	50000
Stiffness of inner casing-outer casing/(N/m)	k_2	10000
Damping of rotor-inner casing/(N · s/m)	c_1	4000
Damping of inner casing-outer casing/(N · s/m)	c_2	1000

parameter transformation and substitution and using the combination of algebra and graphics to solve the problem of finding the optimal distribution using the form of graphical method. This provides a new way of thinking for the design of the controller.

For the following systems:

$$\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} u \\ d \end{bmatrix} \quad (12)$$

In the above formula, u represents the control input and y represents the control output. g is the transfer function obtained during modeling.

Design the controller k such that $u = ky$ and the simultaneous (12) formula, we can get:

$$\begin{cases} y = (1 + g_{11}k)^{-1}g_{12}d \\ z = g_{21}k(1 + g_{11}k)^{-1} + g_{22}d \end{cases} \quad (13)$$

Make a ratio between the output value of the controller and the input value, when the ratio is less than 1, the vibration damping effect can be achieved.

$$\begin{cases} T_y = (1 - g_{11}k)^{-1} \\ T_z = 1 + \frac{g_{21}k(1 - g_{11}k)^{-1}g_{12}}{g_{22}} \end{cases} \quad (14)$$

Define the two parameters α and G such that:

$$\begin{cases} \alpha + 1 = (1 - g_{11}k)^{-1} \\ G = \frac{g_{11}g_{22}}{g_{12}g_{21}} \end{cases} \quad (15)$$

From Eq. (14) (15) that:

$$\begin{cases} |T_y| = |\alpha + 1| \\ |T_z| = \left| \frac{\alpha}{G} + 1 \right| \end{cases} \quad (16)$$

To achieve the damping effect, the ratio of the two formulas in (16) should be less than 1:

$$\begin{cases} |\alpha + 1| < 1 \\ |\alpha + G| < |G| \end{cases} \quad (17)$$

α can be found by the above formula. Bring α into Eq. (15) to find the expression of the optimal controller K :

$$k = \frac{\alpha}{(\alpha + 1)g_{11}} \quad (18)$$

3.2 Control System Design

Using the “geometric design method” controller, the excitation force is transmitted to the outer casing through the inner casing to cause the outer casing vibration acceleration to be controlled. The structure diagram of the acceleration feedback control system composed of the conversion relationship between the signal generator, power amplifier, vibration exciter and acceleration sensor can be drawn, as shown in Fig. 4.

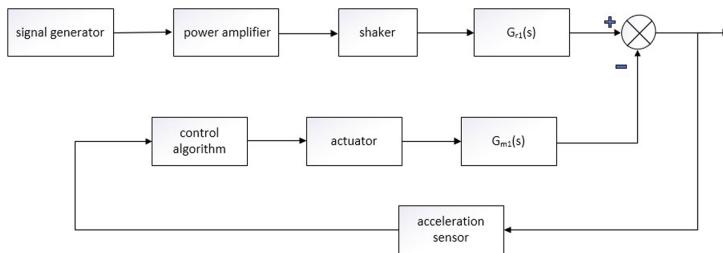


Fig. 4. Block diagram of the control system structure

$$G_{m1}(s) = \frac{m_2 s^4 + c_1 s^3 + c_2 s^2 + k_1 s^2 + k_2 s^2}{m_1 m_2 s^4 + (c_1 m_2 + c_2 m_1 + c_2 m_2) s^3 + (k_1 m_2 + k_2 m_1 + k_2 m_2 + c_1 c_2) s^2 + (c_1 k_2 + c_2 k_1) s + k_1 k_2} \quad (19)$$

$$G_{r1}(s) = \frac{c_2 s^3 + k_2 s^2}{m_1 m_2 s^4 + (c_1 m_2 + c_2 m_1 + c_2 m_2) s^3 + (k_1 m_2 + k_2 m_1 + k_2 m_2 + c_1 c_2) s^2 + (c_1 k_2 + c_2 k_1) s + k_1 k_2} \quad (20)$$

3.3 Simulation Results

During the *Matlab/Simulink*, bring relevant parameters into the system transfer function, and build a control system model, using PID controller and “geometric design” controller for simulation to verify the shock absorption of the active vibration control system, and analyze the simulation results.

It can be obtained from the simulation results in Fig. 5. When the system generates a sinusoidal signal with 10 N and frequency 5 Hz, the actuator can

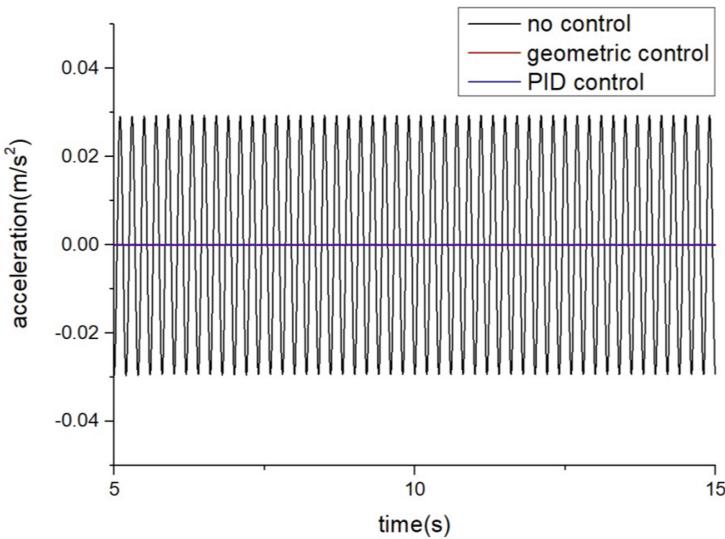


Fig. 5. Active vibration control simulation diagram

respond in time and the corresponding main power is generated to act on the surface of the outer casing, so that the final vibration acceleration of the outer casing almost converges to 0, which has a very good shock absorption effect.

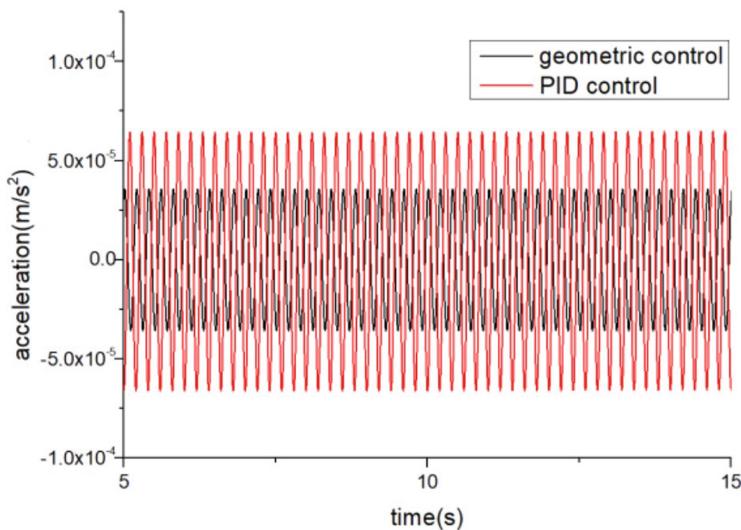


Fig. 6. Comparison of geometric control and PID control effect

It can be seen from the simulation results in Fig. 6 that under the action of the excitation force with the same frequency and the same amplitude, both the PID algorithm and the geometric algorithm can control the amplitude to 10^{-5} level.

4 Conclusion

In order to solve the problems of poor adaptability and insufficient control effect of aero-engine passive vibration isolation, this paper studies active vibration control based on force feedback active control technology. The contributions can be summarized as below:

- (1) Models have been built for both the shaker and the aero-engine vibration transmission path, and the system was simulated and analyzed. The simulation results show that the active vibration reduction can be achieved effectively by exerting a force on the casing.
- (2) This article has used a new control algorithm—"geometric design method," which is adopted to achieve the global vibration control problem through local limited control. Comparing with PID control, this algorithm is more direct and effective, and can achieve ideal control effect.

Acknowledgment. We are grateful for the financial support by the Central Military Commission Foundation to Strengthen Program Technology Fund (No. 2019-JCJQ-JJ-347); Major Special Basic Research Projects for Aviation Engines and Gas Turbines (No. 1002-DLJ19002); and Major Special Basic Research Projects for Green and Sustainable Manufacturing (No. 1002-DCB16001).

References

1. Li, Q., Hu, B., Xu, Z.: Aero-engine strength vibration testing technology, Beijing University of Aeronautics and Astronautics Press (1995)
2. Tokhi, O., Veres, S. (eds.): Active Sound and Vibration Control: Theory and Applications. IEE Control Engineering Series, vol. 62 (2002)
3. Benaroya, H.: Mechanical Vibration: Analysis, Uncertainties and Control, 2nd edn. Marcel Dekker, Inc. (2004)
4. Coleman, R., Remington, P.J.: Noise and vibration control engineering, chap 18. In: Active Control of Noise and Vibration. Willey, Hoboken (2006)
5. John, W.: Sensor and Actuator Networks Networks for Acoustic Signature Monitoring and Control, Undersea Defence Technology 1999, Nice, France (1999)
6. Li, X., Cazzolato, B.S., Hansen, C.H.: Active vibration control of an intermediate mass: vibration isolation in ships. In: Proceedings of the Annual Australian Acoustical Society Conference, Adelaide, Australia, 13–15 November 2002
7. Liu, Z.: Modeling analysis and adaptive active control of noise quality in vehicles. Jilin University (2007)
8. Wang, S.: Semi-physical simulation experiment of active vibration control system based on dSPACE. Harbin Engineering University (2013)

9. Jia, Y.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000)
10. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003)



Stabilization of Linear Systems by a Novel Event-Triggered Control with Time-Varying Threshold

Xinyu Jiang and Fei Hao^(✉)

The Seventh Research Division, School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing 100191, China
{1029672094,fhao}@buaa.edu.cn

Abstract. In this paper, the stabilization problem of linear time invariant systems by event-triggered control (ETC) with time-varying threshold is studied. A novel event-triggered condition with time-varying threshold of the measurement errors is proposed and further the corresponding integral-based triggering condition is constructed to ensure asymptotic stability of systems and avoid Zeno behavior. Several numerical examples are put forward to illuminate the feasibility and efficiency of the aforementioned methods.

Keywords: Linear system · Event-triggered control · Time-varying triggering conditions

1 Introduction

In general, the sampling of sensors and the computing of controller law are based on a time-triggered manner in the controlled systems. In this way, the control tasks are executed periodically [1]. However, to some extent, this time-triggered control method may waste control resources for unnecessary sampling. In order to exclude this shortcoming, ETC has absorbed increasing attention [2]. In this method, only when the triggering condition is violated, that is, the event is triggered, will sensors acquire the new sampling data and transfer them to the controller. Simultaneously, the control law is calculated and the control tasks will be implemented. Thus, ETC ensures control tasks are executed effectively [3].

There are numerous articles in reference to event-triggered control. In literature [4], state will be sampled and the control law will be updated only if the norm of the sampling error exceeds the current threshold. As for linear system, with external interference, the state feedback control based on the event-triggered mechanism is studied in [5]. For linear system, the output feedback control based on the event-triggered conditions and corresponding self-trigger conditions are studied in [6].

As for event-triggered control under an integral-based triggering condition, works for linear time-invariant systems [7], for non-linear systems [8] and for input-to-state stability of linear systems [9] are on the topic of it.

Note that, in all above-mentioned works, triggering conditions depend on the structural information of system, which may result in Zeno behavior with the parameters selected inappropriately [12]. Aiming to address this issue, we propose a triggering condition based on time-varying threshold for event-triggered control. Besides, considering the integral-based triggering strategy has a larger triggering interval [7] and deals with pulsed interference effectively, we further put forward the corresponding integral-based triggering condition with time-varying threshold in this paper.

2 Problem Formulation and Preliminaries

The maximum and minimum eigenvalues of the symmetric matrix C are denoted by $\lambda_M(C)$ and $\lambda_m(C)$. The transpose of a matrix J is denoted by J^T . As for vector ϕ , the 2-norm is denoted by $\|\phi\| =: \sqrt{\phi^T \phi}$. As to matrix J , the induced 2-norm is denoted by $\|J\| =: \sqrt{\lambda_M(J^T J)}$.

Given linear time-invariant system

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (1)$$

where $x \in \mathbb{R}^{n \times 1}$ is the state vector of system, and (A, B) is stabilizable. Therefore, there exists a matrix H such that $(A+BH)$ is a Hurwitz matrix. The state-feedback control law is designed as

$$u(t) = Hx(t_k) \quad t \in [t_k, t_{k+1}), \quad (2)$$

where t_k denotes the time instant when the k th event is triggered by event-triggering conditions. The measurement error is defined as follows

$$\epsilon(t) = x(t_k) - x(t) \quad t \in [t_k, t_{k+1}). \quad (3)$$

According to (1), (2) and (3), we have

$$\dot{x}(t) = (A + BH)x(t) + BH\epsilon(t) \quad t \in [t_k, t_{k+1}).$$

Denote that $\bar{A} = A + BH$ and $\bar{B} = BH$. We obtain

$$\dot{x}(t) = \bar{A}x(t) + \bar{B}\epsilon(t) \quad t \in [t_k, t_{k+1}). \quad (4)$$

Hence, this paper mainly aims at designing event-triggering conditions with time-varying threshold such that (4) is asymptotically stable and avert Zeno behavior.

3 Main Result

In this section, we will propose a novel triggering condition with time-varying threshold. Then, the asymptotic stability of the origin and the absence of Zeno behavior are proven. Furthermore, the corresponding integral-based triggering condition with time-varying threshold is put forward.

3.1 Novel Triggering Condition with Time-Varying Threshold

Given a novel triggering condition with time-varying threshold

$$t_{k+1} = \inf \{t > t_k \mid \|\epsilon(t)\|^2 = \xi(t)\}, \quad (5)$$

where the time-varying threshold function $\xi(t)$ is designed to satisfy: there exists a continuous function $\nu(t)$ satisfying $\nu(t) > 0$, $\dot{\nu}(t) \leq 0$ and $\lim_{t \rightarrow \infty} \nu(t) = 0$ such that the following equations hold

$$\xi(t) > 0, \quad \lim_{t \rightarrow \infty} \xi(t) = 0, \quad -\nu(t)\xi(t) \leq \dot{\xi}(t) \leq 0. \quad (6)$$

Theorem 1. *The origin of system (4) under the event condition (5) is asymptotically stable.*

Proof. Considering that $(A + BH)$ is Hurwitz, there exists positive definite matrices J, Q making the Lyapunov equation $J\bar{A} + \bar{A}^T J = -Q$ holds. Take the candidate Lyapunov function $V(x(t)) = x^T(t)Jx(t)$, then its derivative along the solution of (4) satisfies

$$\begin{aligned} \dot{V}(x(t)) &= \dot{x}^T(t)Jx(t) + x^T(t)J\dot{x}(t) \\ &\leq -\lambda_m(Q)\|x(t)\|^2 + 2\|x(t)\|\|J\bar{B}\|\|\epsilon(t)\| \\ &\leq -(1-b)\lambda_m(Q)\|x(t)\|^2 - b\lambda_m(Q)(\|x(t)\| - \frac{\|J\bar{B}\|\cdot\|\epsilon(t)\|}{b\lambda_m(Q)})^2 \\ &\quad + \frac{\|J\bar{B}\|^2 \cdot \|\epsilon(t)\|^2}{b\lambda_m(Q)} \\ &\leq -(1-b)\lambda_m(Q)\|x(t)\|^2 + \frac{\|J\bar{B}\|^2 \cdot \|\epsilon(t)\|^2}{b\lambda_m(Q)} \\ &\leq -(1-b)\frac{\lambda_m(Q)}{\lambda_M(J)}V(x(t)) + \frac{\|J\bar{B}\|^2}{b\lambda_m(Q)}\cdot\|\epsilon(t)\|^2, \end{aligned}$$

where parameter $b : 0 < b < 1$. Denote that $\varphi = (1-b)\frac{\lambda_m(Q)}{\lambda_M(J)}$ and $\psi = \frac{\|J\bar{B}\|^2}{b\lambda_m(Q)}$. It's easy to know $\varphi, \psi > 0$. Thus, we have

$$\dot{V}(x(t)) \leq -\varphi V(x(t)) + \psi\|\epsilon(t)\|^2. \quad (7)$$

According to Comparison Lemma [11] and (5), we have

$$\begin{aligned} V(x(t)) &\leq e^{-\varphi t}V(x(0)) + \int_0^t e^{-\varphi(t-\tau)}\psi\|\epsilon(\tau)\|^2 d\tau \\ &\leq e^{-\varphi t}V(x(0)) + \psi e^{-\varphi t} \int_0^t e^{\varphi\tau}\xi(\tau)d\tau. \end{aligned} \quad (8)$$

Let $g(t) = \psi e^{-\varphi t} \int_0^t e^{\varphi\tau}\xi(\tau)d\tau$. Apparently, $g(t)$ is continuous, $g(0) = 0$ and $g(t) \geq 0$ holds for $t \in [t_k, t_{k+1})$. Besides, it is easy to know that $\lim_{t \rightarrow \infty} g(t) = 0$. According

to the above mentioned discussion, we can obtain that $g(t)$ has the maximum value when $\dot{g}(t) = 0$. Thus, when $g(t)$ has the maximum value, we have $\varphi \int_0^t e^{\varphi\tau} \xi(\tau) d\tau = e^{\varphi t} \xi(t)$. Therefore, from (8), it is clear that

$$V(x(t)) \leq e^{-\varphi t} V(x(0)) + \frac{\psi}{\varphi} \xi(t). \quad (9)$$

Note that $V(x(t)) = x^T(t) J x(t)$ and $V(x(0)) = x^T(0) J x(0)$. Hence, one has

$$\|x(t)\|^2 \leq \frac{\lambda_M(J)}{\lambda_m(J)} e^{-\varphi t} \|x(0)\|^2 + \frac{\psi}{\varphi \lambda_m(J)} \xi(t). \quad (10)$$

From (6) and (10), we have $\lim_{t \rightarrow \infty} \|x(t)\|^2 = 0$. This is equivalent to $\lim_{t \rightarrow \infty} \|x(t)\| = 0$. For linear systems, the above formula can prove the asymptotic stability of the origin. Hence, the proof is completed.

Theorem 2. *The Zeno behavior does not occur in the system (4) under event condition (5).*

Proof. To prove the result, consider the derivative of $\frac{\|\epsilon(t)\|^2}{\xi(t)}$. Note that $\dot{\epsilon}(t) = -\dot{x}(t) = -\bar{A}x(t) - \bar{B}\epsilon(t)$. According to Young Inequality [10] and (6), we have

$$\begin{aligned} \frac{d}{dt} \cdot \frac{\|\epsilon(t)\|^2}{\xi(t)} &= \frac{2\epsilon^T(t)\dot{\epsilon}(t)\xi(t) - \dot{\xi}(t)\|\epsilon(t)\|^2}{\xi^2(t)} \\ &\leq \frac{2\epsilon^T(t)\dot{\epsilon}(t)\xi(t) + \nu(t)\xi(t)\|\epsilon(t)\|^2}{\xi^2(t)} \\ &\leq \frac{2\|\epsilon(t)\| \cdot (\|\bar{A}\| \cdot \|x(t)\| + \|\bar{B}\| \cdot \|\epsilon(t)\|) + \nu(t)\|\epsilon(t)\|^2}{\xi(t)} \\ &\leq \frac{\|\epsilon(t)\|^2}{\xi(t)} \cdot (2\|\bar{B}\| + \nu(t)) + \frac{\|\bar{A}\|}{\xi(t)} \cdot (\|\epsilon(t)\|^2 + \|x(t)\|^2) \\ &\leq \frac{\|\epsilon(t)\|^2}{\xi(t)} \cdot (2\|\bar{B}\| + \|\bar{A}\| + \nu(t_k)) + \frac{\|\bar{A}\|}{\lambda_m(J)} \cdot \frac{V(x(t))}{\xi(t)}. \end{aligned}$$

Denote that $N_1 = 2\|\bar{B}\| + \|\bar{A}\| + \nu(t_k)$, $N_2 = \frac{\|\bar{A}\|}{\lambda_m(J)}$. Apparently, $N_1, N_2 > 0$, then the above inequality can be written as

$$\frac{d}{dt} \cdot \frac{\|\epsilon(t)\|^2}{\xi(t)} \leq N_1 \cdot \frac{\|\epsilon(t)\|^2}{\xi(t)} + N_2 \cdot \frac{V(x(t))}{\xi(t)}. \quad (11)$$

The following content studies the boundedness of $\frac{V(x(t))}{\xi(t)}$ through its derivative. According to (5), (7) and (11), one has

$$\begin{aligned} \frac{d}{dt} \cdot \frac{V(x(t))}{\xi(t)} &\leq \frac{(-\varphi V(x(t)) + \psi\|\epsilon(t)\|^2)\xi(t) + \nu(t)\xi(t)V(x(t))}{\xi^2(t)} \\ &\leq \frac{(-\varphi V(x(t)) + \psi\xi(t))\xi(t) + \nu(t)\xi(t)V(x(t))}{\xi^2(t)} \\ &= \frac{V(x(t))}{\xi(t)} \cdot (-\varphi + \nu(t)) + \psi. \end{aligned} \quad (12)$$

The analysis of (12) is as follows

(1) Let $T = \max \{0, \arg \min \{\nu(t) - (1-a)\varphi < 0\}\}$, where parameter $a : 0 < a < 1$. Thus, $\nu(t) < (1-a)\varphi$ holds for all $t > T$. Therefore, from (12), the inequality $\frac{d}{dt} \cdot \frac{V(x(t))}{\xi(t)} \leq -a\varphi \cdot \frac{V(x(t))}{\xi(t)} + \psi$ holds. Since any $\frac{V(x(t))}{\xi(t)} > \frac{\psi}{a\varphi}$, then $\frac{d}{dt} \cdot \frac{V(x(t))}{\xi(t)} \leq 0$.

Therefore, for any $t > T$, $\frac{V(x(t))}{\xi(t)} \leq \frac{\psi}{a\varphi}$ holds.

(2) For all $t \in [0, T]$, since $\frac{V(x(t))}{\xi(t)}$ is continuous, there exists $t_1 \in [0, T]$ such that $\frac{V(x(t))}{\xi(t)} \leq \frac{V(x(t_1))}{\xi(t_1)}$ holds.

Let $N_3 = \max \left\{ \frac{\psi}{a\varphi}, \frac{V(x(t_1))}{\xi(t_1)} \right\}$. It is clear that $\frac{V(x(t))}{\xi(t)} \leq N_3$ holds for all $t \in [t_k, t_{k+1})$. Hence, according to (11), we have $\frac{d}{dt} \cdot \frac{\|\epsilon(t)\|^2}{\xi(t)} \leq N_1 \cdot \frac{\|\epsilon(t)\|^2}{\xi(t)} + N_2 N_3$. By Comparison Lemma [11], one has

$$\begin{aligned} \frac{\|\epsilon(t)\|^2}{\xi(t)} &\leq e^{N_1(t-t_k)} \frac{\|\epsilon(t_k)\|^2}{\xi(t_k)} + \int_{t_k}^t e^{N_1(t-\tau)} N_2 N_3 d\tau \\ &= \frac{N_2 N_3}{N_1} (e^{N_1(t-t_k)} - 1). \end{aligned}$$

From (5), when the control event is triggered, the inequality $\frac{N_2 N_3}{N_1} (e^{N_1(t-t_k)} - 1) \geq 1$ holds. Hence, we obtain that $t - t_k \geq \frac{1}{N_1} \cdot \ln \left(\frac{N_1}{N_2 N_3} + 1 \right)$, which proves the system is capable to avoid Zeno behavior. Hence, the proof is completed.

3.2 Integral-Based Triggering Condition with Time-Varying Threshold

Given an integral-based triggering condition with time-varying threshold

$$t_{k+1} = \inf \left\{ t > t_k \mid \int_{t_k}^t \|\epsilon(\tau)\|^2 d\tau = \zeta(t) \right\}, \quad (13)$$

where the time-varying threshold $\zeta(t)$ is designed to satisfy: there exists a positive constant d_I such that the following equations hold

$$\zeta(t) > 0, \quad \lim_{t \rightarrow \infty} \zeta(t) = 0, \quad -d_I \zeta(t) \leq \dot{\zeta}(t) \leq 0. \quad (14)$$

Theorem 3. *The origin of system (4) under the event condition (13) is asymptotically stable.*

Proof. Since (7) has nothing to do with the triggering conditions, we obtain

$$\dot{V}(x(t)) \leq -\varphi V(x(t)) + \psi \|\epsilon(t)\|^2.$$

By Comparison Lemma [11] and (13), it is clear that

$$\begin{aligned} V(x(t)) &\leq e^{-\varphi(t-t_k)} V(x(t_k)) + \int_{t_k}^t e^{-\varphi(t-\tau)} \psi \|\epsilon(\tau)\|^2 d\tau \\ &\leq e^{-\varphi(t-t_k)} V(x(t_k)) + \psi \int_{t_k}^t \|\epsilon(\tau)\|^2 d\tau \\ &\leq e^{-\varphi(t-t_k)} V(x(t_k)) + \psi \zeta(t). \end{aligned} \quad (15)$$

Note that $V(x(t)) = x^T(t)Jx(t)$ and $V(x(t_k)) = x^T(t_k)Jx(t_k)$, one has

$$\|x(t)\|^2 \leq \frac{\lambda_M(J)}{\lambda_m(J)} e^{-\varphi(t-t_k)} \|x(t_k)\|^2 + \frac{\psi}{\lambda_m(J)} \zeta(t).$$

Let $\varphi_{IJ} = \frac{\lambda_M(J)}{\lambda_m(J)}$ and $\psi_{IJ} = \frac{\psi}{\lambda_m(J)}$. It is easy to know $\varphi_{IJ}, \psi_{IJ} > 0$. Thus, the above inequality can be written as

$$\|x(t)\|^2 \leq \varphi_{IJ} e^{-\varphi(t-t_k)} \|x(t_k)\|^2 + \psi_{IJ} \zeta(t). \quad (16)$$

From (14) and (16), we obtain that $\lim_{t \rightarrow \infty} \|x(t)\|^2 = 0$. This is equivalent to $\lim_{t \rightarrow \infty} \|x(t)\| = 0$. For linear system, the above formula can prove the asymptotic stability of the origin. Hence, the proof is completed.

Theorem 4. *The Zeno behavior does not occur in system (4) under event condition (13).*

Proof. Let $\mu(t) = \int_{t_k}^t \|\epsilon(\tau)\|^2 d\tau$ $t \in [t_k, t_{k+1})$. To prove the result, consider the derivative of $\frac{\mu(t)}{\zeta(t)}$. From (14), we have

$$\begin{aligned} \frac{d}{dt} \cdot \frac{\mu(t)}{\zeta(t)} &= \frac{\dot{\mu}(t)\zeta(t) - \dot{\zeta}(t)\mu(t)}{\zeta^2(t)} \\ &\leq \frac{\dot{\mu}(t)\zeta(t) + d_I \zeta(t)\mu(t)}{\zeta^2(t)} \\ &= d_I \cdot \frac{\mu(t)}{\zeta(t)} + \frac{\|\epsilon(t)\|^2}{\zeta(t)}. \end{aligned}$$

Note that $\mu(t_k) = \int_{t_k}^{t_k} \|\epsilon(\tau)\|^2 d\tau = 0$, by Comparison Lemma [11], one has

$$\frac{\mu(t)}{\zeta(t)} \leq e^{d_I t} \cdot \int_{t_k}^t \frac{\|\epsilon(\tau)\|^2}{e^{d_I \tau} \cdot \zeta(\tau)} d\tau. \quad (17)$$

Let $g_I(t) = \frac{1}{e^{d_I t} \cdot \zeta(t)}$ $t \in [t_k, t_{k+1})$. According to (14), for all $t \in [t_k, t_{k+1})$, $\dot{g}_I(t) \leq \frac{1}{e^{d_I t} \cdot \zeta^2(t)} (d_I \zeta(t) - d_I \zeta(t))$ i.e. $\dot{g}_I(t) \leq 0$ holds, which shows that the function $g_I(t)$ decreases monotonically. Note that $g_I(t_k) = \frac{1}{e^{d_I t_k} \cdot \zeta(t_k)}$, from (17), we can obtain

$$\frac{\mu(t)}{\zeta(t)} \leq \frac{e^{d_I(t-t_k)}}{\zeta(t_k)} \cdot \int_{t_k}^t \|\epsilon(\tau)\|^2 d\tau. \quad (18)$$

Note that $\epsilon(t) = x(t_k) - x(t)$, from (14) and (16), it is easy to find that

$$\begin{aligned} \|\epsilon(t)\|^2 &\leq 2(\|x(t_k)\|^2 + \|x(t)\|^2) \\ &\leq 2(\varphi_{IJ} e^{-\varphi(t-t_k)} + 1) \|x(t_k)\|^2 + 2\psi_{IJ} \zeta(t) \\ &\leq 2(\varphi_{IJ} + 1) \|x(t_k)\|^2 + 2\psi_{IJ} \zeta(t_k) \quad t \in [t_k, t_{k+1}). \end{aligned} \quad (19)$$

According to (18) and (19), we have

$$\frac{\mu(t)}{\zeta(t)} \leq \frac{2(\varphi_{IJ} + 1)\|x(t_k)\|^2 + 2\psi_{IJ}\zeta(t_k)}{\zeta(t_k)} \cdot e^{d_I(t-t_k)} \cdot (t - t_k).$$

Denote that $a_I = \frac{2(\varphi_{IJ} + 1)\|x(t_k)\|^2 + 2\psi_{IJ}\zeta(t_k)}{\zeta(t_k)}$. It is obvious that $a_I > 0$. Then, the above inequality can be written as $\frac{\mu(t)}{\zeta(t)} \leq a_I e^{d_I(t-t_k)}(t - t_k)$. Let $h_I(t) = a_I e^{d_I(t-t_k)}(t - t_k) - 1$. According to (13), we can obtain that $h_I(t) \geq 0$ holds when the control event is triggered. Note that $h_I(t_k) = -1$, $\lim_{t \rightarrow \infty} h_I(t) = \infty$ and $h_I(t)$ is continuous and monotonously increasing. Thus, there exists $t_2 > t_k$ which makes $h_I(t_2) = 0$ hold. Thus, the triggering interval $t_2 - t_k > 0$ i.e. the system is capable to exclude the Zeno behavior. Hence, the proof is completed.

4 Simulation

In this section, several numerical examples about ETC under a novel triggering condition and the corresponding integral-based triggering condition with time-varying threshold are shown. Consider a linear time-invariant system $\begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} -5 \\ 2 \end{pmatrix} u(t)$. The initial state value is $(x_1(0), x_2(0))^T = (1, 1)^T$. The control law is $u(t) = (90 \quad 222)(x_1(t_k), x_2(t_k))^T$ for $t \in [t_k, t_{k+1})$. Given a novel time-varying triggering condition

$$t_{k+1} = \inf \{t > t_k \mid \|\epsilon(t)\|^2 = \xi_1(t)\}, \quad (20)$$

and the corresponding time-varying integral-based triggering condition

$$t_{k+1} = \inf \left\{ t > t_k \mid \int_{t_k}^t \|\epsilon(\tau)\|^2 d\tau = \zeta_1(t) \right\}, \quad (21)$$

where $\xi_1(t) = \zeta_1(t) = (\frac{1}{\ln(\alpha t + \beta)})^\gamma$ ($\alpha, \gamma > 0, \beta > 1$). It is easy to know that $\dot{\xi}_1(t) = -\frac{\gamma\alpha}{(\alpha t + \beta) \ln^\gamma(\alpha t + \beta)} \xi_1(t)$. Apparently, $\xi_1(t)$ meet the requirements shown in (6). Let $d_I = \frac{\gamma\alpha}{\beta \ln \beta}$. Hence, we have $\dot{\xi}_1(t) \geq -d_I \xi_1(t)$ which proves that $\xi_1(t)$ meet the requirements for integral-based triggering condition with time-varying threshold which are shown in (14). In the simulation results below, let $\alpha = 1, \beta = 100, \gamma = 3$. i.e. let

$$\xi_1(t) = \left(\frac{1}{\ln(t + 100)}\right)^3. \quad (22)$$

The abovementioned formulas are served as time-varying threshold in the following simulation.

In accordance with Fig. 1(a), we know that system based on (20) with time-varying threshold (22) is asymptotically stable. According to Fig. 1(b), we know

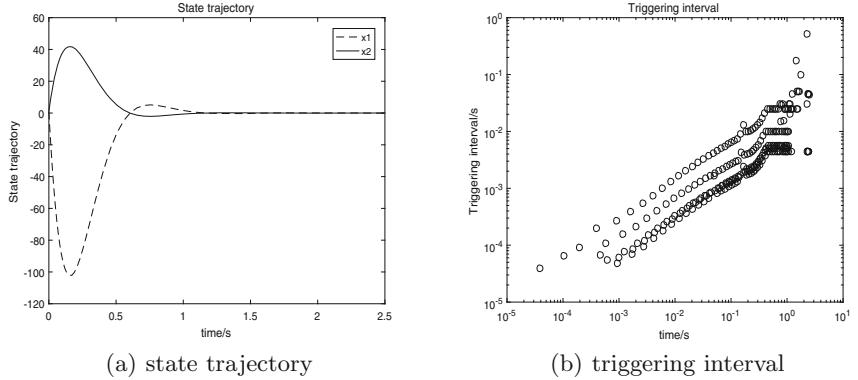


Fig. 1. Results for novel triggering condition (20) with time-varying threshold (22)

that the minimum time triggering interval is 3.86×10^{-5} s. Considering that the simulation step size is 1×10^{-5} s, therefore, it proves that Zeno behavior does not occur in the above system.

Figure 2(a) indicates that system based on integral-based triggering condition (21) with time-varying threshold (22) is asymptotically stable. On the basis of Fig. 2(b), the minimum triggering time interval is 0.003 s. Note that the simulation step size is 0.001 s, therefore, Zeno behavior does not occur in the above system. Through the comparison of Fig. 1 and Fig. 2, we discover that under the same time-varying threshold, the triggering numbers of integral-based triggering condition (21) is less than that of the triggering condition (20). Nevertheless, the control performance of system with integral-based triggering condition (21) is slightly inferior to the triggering condition (20).

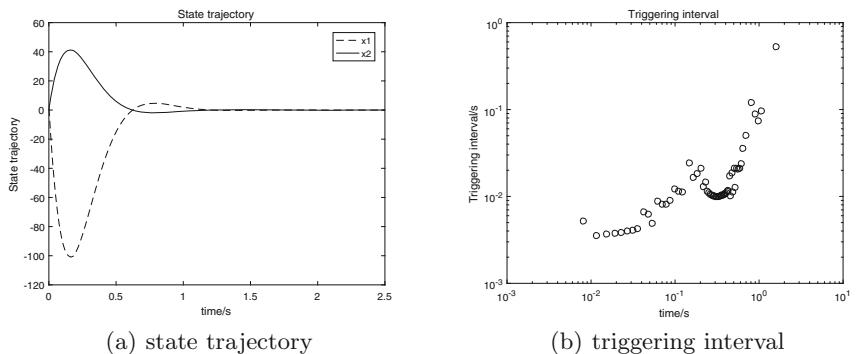


Fig. 2. Results for integral-based triggering condition (21) with time-varying threshold (22)

5 Conclusion

We have studied the stabilization problem of linear time invariant systems under event-triggered control with time-varying threshold in this paper. A novel triggering condition and the corresponding integral-based triggering condition with time-varying threshold were proposed. ETC with the aforementioned triggering conditions has been proven to assure the asymptotic stability of its origin and avoid Zeno behavior. Several numerical examples were shown to illuminate the feasibility and efficiency of the proposed methods. The simulation results showed that compared to the novel triggering condition, systems under the integral-based triggering condition have a larger triggering interval at the expense of a little control performance. Further extensions would include nonlinear systems based on event-triggered control with the abovementioned conditions.

References

1. Laila, D.S., Nesić, D., Astolfi, A., Panteley, E.: Sampled-Data Control of Nonlinear Systems. Springer, London (2006)
2. Dorf, R., Farren, M., Phillips, C.: Adaptive sampling frequency for sampled-data control systems. *IRE Trans. Autom. Control* **7**(1), 38–47 (1962)
3. Astrom, K.J., Bernhardsson, B.M.: Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In: Proceedings of the 41st IEEE Conference on Decision and Control, vol. 2, pp. 2011–2016 (2002)
4. Tabuada, P.: Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Trans. Autom. Control* **52**(9), 1680–1685 (2007)
5. Lunze, J., Lehmann, D.: A state-feedback approach to event-based control. *Automatica* **46**(1), 211–215 (2010)
6. Chen, X., Hao, F.: Stability of event-triggered output-feedback control systems. In: Proceedings of the 30th Chinese Control Conference, vol. 1416, no. 1, pp. 1184–1189 (2011)
7. Mousavi, S.H., Ghodrat, M., Marquez, H.J.: A novel integral-based event triggering control for linear time-invariant systems. In: 53rd IEEE Conference on Decision and Control, pp. 1239–1243 (2014)
8. Mousavi, S.H., Ghodrat, M., Marquez, H.J.: Integral-based event-triggered control scheme for a general class of non-linear systems. *IET Control Theory Appl.* **9**(13), 1982–1988 (2015)
9. Yu, H., Hao, F.: Input-to-state stability of integral-based event-triggered control for linear plants. *Automatica* **85**(2), 248–255 (2017)
10. Surhone, L.M., Timpledon, M.T., Marseken, S.F., Young, W.H., Inequality, H.: Young's Inequality. Betascript Publishing (2010)
11. Khalil, H.K.: Nonlinear Systems, pp. 71–89 (2020)
12. Guinaldo, M., Dimarogonas, D.V., Johansson, K.H., Snchez, J., Dormido, S.: Distributed event-based control strategies for interconnected linear systems. *IET Control Theory Appl.* **7**(6), 877–886 (2013)



Modelling of SFR for Wind-Thermal Power Systems via Improved RBF Neural Networks

Jianhua Zhang^{1(✉)}, Hongrui Li², Bo Hu³, Yong Min⁴, Qun Chen⁴, Guolian Hou², and Congzhi Huang²

¹ State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China

zjh@ncepu.edu.cn

² School of Control and Computer Engineering,
North China Electric Power University, Beijing 102206, China

³ State Grid Liaoning Electric Power Supply Co. Ltd., Shenyang, China

⁴ School of Control and Computer Engineering, Tsinghua University,
Beijing 100089, China

Abstract. In this paper, one kind of system identification algorithm based on an improved radial basis function (RBF) neural networks is presented for nonlinear and non-Gaussian systems. Survival information potential (SIP) of identification errors is employed to constructed the performance index to train the RBF neural networks. The data driven system identification algorithm implemented by the proposed RBF neural networks is applied to obtain the equivalent model of wind-thermal integrated power systems. Compared with the traditional RBF neural networks and back-propagation (BP) neural networks based on mean square error (MSE) criteria, simulation results demonstrate that the proposed system identification algorithm can obtain better SFR models for wind-thermal integrated power systems.

Keywords: Power system · Survival information potential · System identification · Radial basis function neural networks

1 Introduction

In order to reach the goals of energy security and environmental protection, it is essential to utilize renewable energy in terms of wind energy, solar energy, and so on. Hence, it's important to incorporate renewable energy into power grid. The imbalance between the electrical load and the power supplied by the connected generators leads to frequency deviation. Permanent abnormal frequency deviation may influence power system operation, security, reliability and efficiency. Therefore, it calls for investigating power system frequency response (SFR) under disturbances.

The existed methods to study SFR can be classified into full model time domain simulation method, linearization method and equivalent model method [1]. These methods all depend on physical models obtained by making some assumptions and simplifications. With the further increase of the proportion of centralized and distributed renewable energy connected to power grid, the power electronic interfaces of the power system increase [2] while the inertia of the power system decreases [3]. In addition, there are stochastic disturbances induced by renewable energy power generation. Consequently, it is difficult to obtain accurate frequency response model of power systems based on mechanism modeling.

On one hand, equivalent model method can build SFR model more quickly than the other two methods; On another hand, neural networks have been widely applied to industrial process modeling [4] because of their universal approximation ability. Accordingly, neural networks based system identification method is combined with hybrid equivalent model method to establish the SFR model for a wind-thermal integrated power system in this paper. An improved radial basis function neural networks (RBFNNs) is investigated to build hybrid equivalent model method using survival information potential (SIP) criterion.

The main contributions can be highlighted as follows: (i) an RBFNNs based system identification algorithm is developed for nonlinear non-Gaussian stochastic systems; (ii) a novel performance index is proposed for RBF neural networks based on the SIP of identification error; (iii) the proposed system identification algorithm is applied to build hybrid equivalent model for wind-thermal integrated power systems.

The remainder of the paper is organized as follows. Section 2 describes a wind-thermal integrated power system, the equivalent models of both wind power generation system and thermal power plant are explained in brief. Section 3 presents improved RBFNNs to identify the integrated power system by using SIP criterion. Section 4 discusses the application of the proposed RBFNNs and presents simulation examples to testify its effectiveness. Finally, Sect. 5 gives the concluding remarks.

2 Wind-Thermal Power System and Its Equivalent Models

The high penetration of wind power generation leads to the decline of the equivalent inertia of power systems and the weakening of the primary frequency regulation capacity, and even to some extent brings about the deterioration of the frequency characteristics [5]. A wind-thermal integrated power system is shown in Fig. 1, where two thermal power generation units and a wind farm are connected to power grid at the same point of common coupling.

In order to investigate its system frequency response, a simplified frequency response model is usually required to analyze frequency behavior of the integrated power system, in which the dynamic model of each power generation system must be included. The equivalent model takes the least cost among the above mentioned full model, linearized model and equivalent model.

The hybrid equivalent model shown in Fig. 2 includes two equivalent models. One is the equivalent model of the thermal power plant, in which there are some thermal power units; another one is the equivalent model of the wind farm consisting of some wind power generation units.

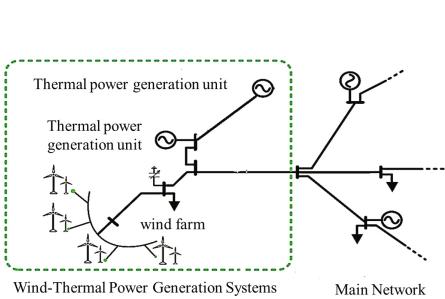


Fig. 1. Wind-thermal power system

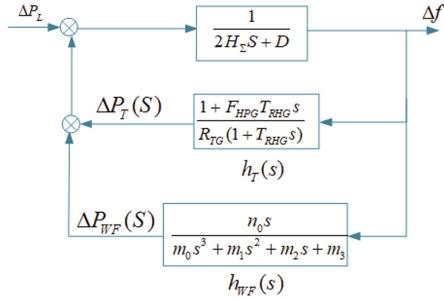


Fig. 2. Structure of hybrid equivalent model [6]

2.1 Equivalent Model of Thermal Power Plant

The physical model of the steam turbine-governor can be easily obtained for a single thermal power unit. Furthermore, the model can then be simplified to be a low-order model. If there are multiple thermal power units in the thermal power plant, the weighted equivalence method is employed to obtain the equivalent steam turbine-governor aggregation model [7], whose transfer function can be formulated as follows:

$$h_T(s) = \frac{\Delta P_T(s)}{\Delta f(s)} = \frac{1 + F_{HPG} T_{RHG} s}{R_{TG}(1 + T_{RHG} s)} \quad (1)$$

where R_{TG} , T_{RHG} and F_{HPG} are the aggregation parameters of the equivalent thermal power unit's adjustment coefficient, reheater time constant, and high-pressure turbine stage power ratio respectively. The mechanical power increment ΔP_T are regarded as the equivalent output and the system frequency increment Δf as the input.

2.2 Equivalent Model of Wind Farm

Variations of wind speed will influence the operating states of wind turbines. When considering the primary frequency modulation response of each wind turbine in a wind farm, we make an assumption that the speed difference of each wind turbine is small, therefore, the speed is almost same and approximately equal to the average speed. In addition, the system frequency deviation signals Δf detected by each wind turbine are identical. When we build the system response transfer function model for the wind farm, the wind farm can be equivalent to a group of wind turbines, and the equivalent model of the wind farm, $h_{WF}(s)$, can be obtained using the parameter aggregation method [8]:

$$h_{WF}(s) = \frac{\Delta P_{WF}(s)}{\Delta f(s)} = \frac{n_0 s}{m_0 s^3 + m_1 s^2 + m_2 s + m_3} \quad (2)$$

where n_0, m_0, m_1, m_2 and m_3 are the equivalent aggregation parameters of the transfer function $h_{WF}(s)$ respectively.

3 Modelling Based on Improved RBFNNs

Although the equivalent models of both thermal power plant and wind farm can be obtained using Eqs. (1) and (2), however, the above equivalent models has some limitations: 1) all governors or prime movers in the thermal power plant and wind farm have to be the same type, so the equivalent models can't deal with the power generation units with different types of governors or prime movers; 2) the equivalent models are obtained offline rather than online, they can't deal with the nonlinearity, time variation, coupling and uncertainties existed in wind-thermal power systems.

In order to obtain real time and accurate SFR model, RBFNNs based system identification method is combined with equivalent model method in this work. The conventional RBFNNs are trained under minimum mean-square error (MSE) criterion. However, it was observed that power grid frequency in Europe, United States and Japan is of random fluctuation, in addition, the frequency fluctuations are non-Gaussian [9]. Hence, the SIP of identification error is utilized to train RBFNNs when identifying the equivalent model of wind-thermal power systems.

RBFNNs shown in Fig. 3 is employed to identify the implicit relationship between the wind power/steam turbine power output and power grid frequency. RBFNNs is utilized to implement equivalent modelling by directly learning the input and output data of the power generation systems. The detailed structure of the RBFNNs is shown in Fig. 4, in which the input and output are

$x(k) = [x_1(k), x_2(k), \dots, x_n(k)]^T$ and $y_m(k)$ respectively. The activation function of the hidden layer takes following Gaussian basis function:

$$R_j(x) = \exp\left(-\frac{\|x - C_j\|^2}{2b_j^2}\right) \quad (3)$$

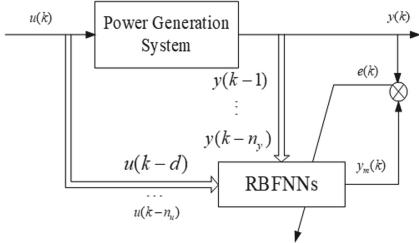


Fig. 3. Diagram of system identification based on RBFNNs

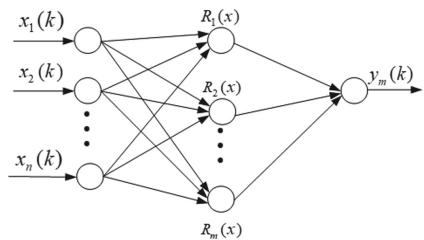


Fig. 4. The structure of RBFNNs

3.1 Forward Transmission of Signal

The input of the neuron in the input layer is $x(k) = [x_1(k), x_2(k), \dots, x_i(k), \dots, x_n(k)]^T$, $i = 1, 2, \dots, n$, the output of the neuron in the hidden layer can then be obtained from Eq. (3) as follows:

$$R_j(x(k)) = \exp\left(-\frac{\|x(k) - c_j(k-1)\|^2}{2b_j^2(k-1)}\right) \quad (4)$$

The output of the neuron in the output layer is:

$$y_m(k) = \sum_{j=1}^m \omega_j(k-1) R_j(x(k)) \quad (5)$$

The error used to train RBF network is:

$$e(k) = y(k) - y_m(k) \quad (6)$$

where $y(k)$ and $y_m(k)$ are actual power grid frequency and the identified power grid frequency using the proposed RBFNNs respectively.

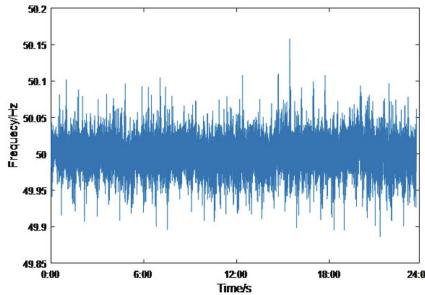


Fig. 5. Frequency fluctuation of power grid in Europe on a certain day

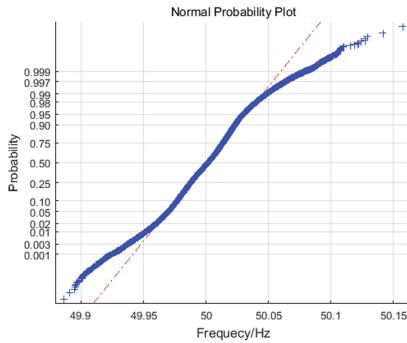


Fig. 6. Normal probability distribution diagram of power grid frequency

3.2 Performance Index

In actual power systems, the grid frequency has the characteristics of nonlinearity and randomness, and its fluctuation doesn't obey the Gaussian distribution [9]. Figure 5 shows the frequency fluctuation curve of a power grid in Europe on a certain day. It can be seen that the frequency fluctuates around 50 Hz. Figure 6 presents the probability of normal test of frequency data shown in Fig. 5. It is clear that the power grid frequency follows a non-Gaussian distribution.

Since the power grid frequency is non-Gaussian, the RBFNNs based system identification can't gain ideal results when using MSE criterion. SIP criterion can be used to investigate non-Gaussian systems because of its advantages [10]: 1) has good robustness; 2) does not involve the selection of core width; 3) does not have translation invariance. Therefore, the parameters of RBFNNs will be optimized by minimizing the SIP of the identification error in Eq. (6).

Let α be a natural number, the SIP of the estimation error $Z(k) \in Rm$ can be derived as follows [10]

$$\begin{aligned}
\hat{S}_\alpha(Z) &= \int_{R_+^m} \left(\frac{1}{N} \sum_{k=1}^N \iota(Z(k) > z) \right)^\alpha dz \\
&= \frac{1}{N^\alpha} \int_{R_+^m} \sum_{j_1, \dots, j_\alpha=1}^N (\iota(Z(j_1) > z) \times \dots \times \iota(Z(j_\alpha) > z)) dz \\
&= \frac{1}{N^\alpha} \sum_{j_1, \dots, j_\alpha=1}^N \int_{R_+^m} (\iota(Z(j_1) > z) \times \dots \times \iota(Z(j_\alpha) > z)) dz \\
&= \frac{1}{N^\alpha} \sum_{j_1, \dots, j_\alpha=1}^N \int_{R_+^m} \left(\prod_{i=1}^m \iota(\min(Z_i(j_1), \dots, Z_i(j_\alpha)) > z_i) \right) dz \\
&= \frac{1}{N^\alpha} \sum_{j_1, \dots, j_\alpha=1}^N \left(\prod_{i=1}^m \int_{R_+^m} \iota(\min(z_i(j_1), \dots, z_i(j_\alpha)) > z_i) dz_i \right) \\
&= \frac{1}{N^\alpha} \sum_{j_1, \dots, j_\alpha=1}^N \left(\prod_{i=1}^m \min(z_i(j_1), \dots, z_i(j_\alpha)) \right)
\end{aligned} \tag{7}$$

3.3 RBF Model Training

In this work, let $\alpha = 2$, the quadratic survival information potential of the identification error is used to construct the performance index for the RBFNN. The estimated SIP can be obtained at instant k using the error series $e(k-L+1 : k)$ as follows:

$$\hat{S}_\alpha(e) = \frac{1}{L^2} \sum_{j_1=1}^L \sum_{j_2=1}^L (\min(e(j_1), e(j_2))) \tag{8}$$

The learning algorithm of the weight $\omega_j(k)$ from the hidden layer to the output layer is:

$$\begin{aligned}
\Delta\omega_j(k) &= -\eta \frac{\partial E(k)}{\partial \omega_j(k-1)} \\
&= \eta \frac{1}{L^2} \left(\sum_{i=1}^L \sum_{j=1}^L \left(\frac{\partial \min(e(i), e(j))}{\partial e_1} \right) R_j(x(k)) \right)
\end{aligned} \tag{9}$$

$$\omega_j(k) = \omega_j(k-1) + \Delta\omega_j(k) + \bar{\alpha}(\omega_j(k-1) - \omega_j(k-2)) \tag{10}$$

The learning algorithm of $b_j(k)$ and $c_{ji}(k)$ is:

$$\Delta b_j(k) = -\eta \frac{\partial E(k)}{\partial b_j(k-1)} = \eta \frac{1}{L^2} \left(\sum_{i=1}^L \sum_{j=1}^L \left(\frac{\partial \min(e(i), e(j))}{\partial e} \right) \cdot \omega_j(k-1) R_j(x(k)) \frac{\|x(k) - c_j(k-1)\|^2}{b_j^3(k-1)} \right) \quad (11)$$

$$b_j(k) = b_j(k-1) + \Delta b_j(k) + \bar{\alpha}(b_j(k-1) - b_j(k-2)) \quad (12)$$

$$\Delta c_{ji}(k) = -\eta \frac{\partial E(k)}{\partial c_{ji}(k-1)} = \eta \frac{1}{L^2} \left(\sum_{i=1}^L \sum_{j=1}^L \left(\frac{\partial \min(e(i), e(j))}{\partial e} \right) \cdot \omega_j(k-1) R_j(x(k)) \frac{x_i(k) - c_{ji}(k-1)}{b_j^2(k-1)} \right) \quad (13)$$

$$c_{ji}(k) = c_{ji}(k-1) + \Delta c_{ji}(k) + \bar{\alpha}(c_{ji}(k-1) - c_{ji}(k-2)) \quad (14)$$

where η and $\bar{\alpha}$ are the learning rate and the momentum factor respectively.

4 Simulation Results

The simulated wind-thermal power system shown in Fig. 7 is a modified two-machine model. The simulation parameters of the thermal power unit model and the frequency modulation model are the same as those in reference [8], in which a 600 MW thermal power plant G1 was simulated. In this work, the wind farm whose installed capacity is 90 MW consists of 60 doubly fed Induction generator (DFIG) units. Each DFIG unit is identical. The first load L1 is 560 MW, and the second load L2 is 55 MW. The initial frequency of the simulation system is $f = 50$ Hz.

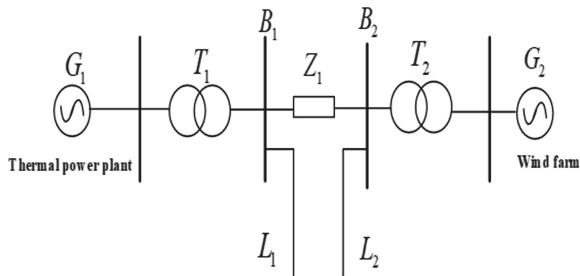


Fig. 7. Wind-thermal power system

A 7-10-1 RBFNNs is used to identify the equivalent model of the wind-thermal power system shown in Fig. 7. The excitation signals are needed during identifying the SFR models. In this simulation, Gaussian white noises are imposed on the first load L_1 , the output power of the thermal power plant and the wind farm, respectively. The initial weights are set to random number between $[-1, 1]$, the sampling time is 0.1 s. The simulation is performed over 1600 time instants. The learning rate and the momentum factor are set to $\eta = 0.01$ and $\bar{\alpha} = 0.05$ respectively.

Some comparative results are given to illustrate the superiority of the proposed RBFNNs based system identification method under SIP criterion. Compared with the RBFNNs and BPNNs under MSE criterion, it is clear from Fig. 8 that the proposed method has better performance whether at training stage or testing stage. In addition, it can also be observed from Table 1 that the proposed method can gain smallest identification error in term of root mean square error (RMSE), standard deviation (SD), mean absolute error (MAE) and R^2 .

In addition, Figs. 9 and 10 reveal the results using the proposed RBFNNs in detail. The probability density function of the identification error γ_e at typical instants is shown in Fig. 9. It can be seen that the identification error decreases towards a smaller randomness direction. Likewise, this can also be verified in the three-dimensional mesh plot of the system output error in Fig. 10. The shape of the probability density function of the identification error in Figs. 9 and 10 becomes narrower and sharper along with sampling time, it indicates that the proposed RBFNNs can deal with non-Gaussian frequency fluctuations.

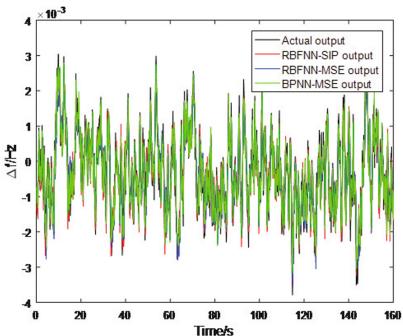


Fig. 8. Identification result

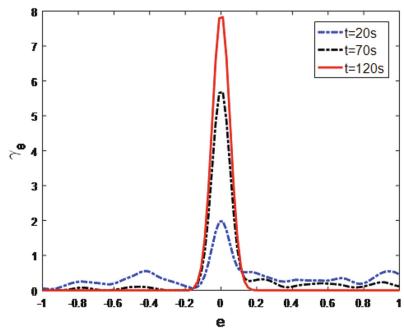


Fig. 9. Three-dimensional mesh map of system output error

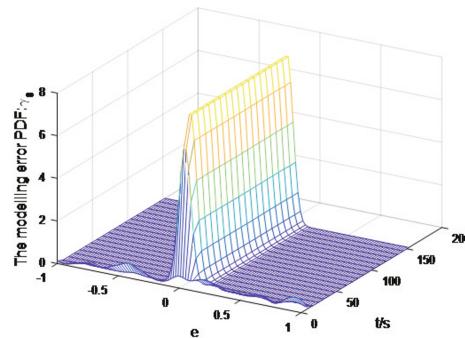


Fig. 10. Typical time PDF of system output error

Table 1. Comparison of training process

Method	RMSE	SD	MAE	R^2
BPNN-MSE method	4.58e-4	4.58e-4	3.56e-4	0.965
RBFNN- MSE method	4.26e-4	4.26e-4	3.39e-4	0.972
RBFNN- SIP method	3.87e-4	3.87e-4	3.12e-4	0.979

5 Conclusions

In this paper, an improved RBFNNs is presented to identify the model of non-Gaussian systems using SIP criterion due to reducing the uncertainty of the identification error. Accordingly, combined with the equivalent modelling method, the proposed system identification method is applied to build SFR model for wind-thermal power systems. Finally, the proposed data-driven modelling method is testified by an illustrative example. The presented system identification algorithm can effectively deal with non-Gaussian power grid frequency fluctuations. The SFR model represented by the RBFNNs can reveal the characteristics of the frequency responses under disturbances from load, thermal power generation and wind power generation.

Acknowledgments. This work was supported by National Key R&D Program of China No. 2019YFB1505402.

References

1. Liu, J., Wang, X., Lin, J., Teng, Y.: A hybrid equivalent model for prediction of power system frequency response. In: 2018 IEEE Power & Energy Society General Meeting, pp. 1–5 (2018)
2. Ye, Y., Qiao, Y., Lu, Z.: Revolution of frequency regulation in the converter-dominated power system. Renew. Sustain. Energy Rev. **111**, 145–156 (2019)

3. Dreidy, M., Mokhlis, H., Mekhilef, S.: Inertia response and frequency control techniques for renewable energy sources. *Renew. Sustain. Energy Rev.* **69**, 144–155 (2017)
4. Ueckerdt, F., Brecha, R., Luderer, G.: Analyzing major challenges of wind and solar variability in power systems. *Renew. Energy* **81**, 1–10 (2015)
5. Asmine, M., Langlois, C.: Field measurements for the assessment of inertial response for wind power plants based on HydroQuébec Transénergie requirements. *IET Renew. Power Gener.* **10**(1), 25–32 (2016)
6. Li, S., Cao, R., et al.: Frequency characteristics of power system involving virtual inertia and primary frequency regulation. *Renew. Energy Resour.* **37**(05), 694–700 (2019). (in Chinese)
7. Kundur. *Power System Stability And Control*. McGraw-Hill, New York (1994)
8. Li, S., Huang, Y., et al.: Modeling primary frequency regulation auxiliary control system of doubly fed induction generator based on rotor speed control. *Proc. CSEE* **37**(24), 9 (2017). (in Chinese)
9. Sch Fer, B., Beck, C., Aihara, K., et al.: Non-Gaussian power grid frequency fluctuations characterized by Lévy-stable laws and superstatistics. *Nat. Energy* **3**, 119–126 (2018)
10. Chen, B., Zhu, P., Jos, C.: Survival information potential: a new criterion for adaptive system training. *IEEE Trans. Signal Process.* **60**, 1184–1194 (2012)



An Efficient Action Recognition Framework Based on ELM and 3D CNN

Yiping Zou and Xuemei Ren^(✉)

Department of Automation, Beijing Institute of Technology, Beijing 100036, China
xmren@bit.edu.cn

Abstract. Deep neural network is shown to be the most efficient method for video representation and has achieved state-of-art results on different datasets of action recognition. In this paper, we proposed a hybrid architecture which integrates deep convolutional neural networks and extreme learning machine. The hybrid structure makes the most of their advantages: in the first stage the deep residual 3D network learns the features from both temporal and spatial sequences, then the ELM, instead of traditional classifiers, classifies the actions without tuning the parameters. The resulting network can not only extract the representation fully, but also obtain more accurate results faster. We show the effectiveness and outperformance of the proposed strategy on experiments.

Keywords: Action recognition · Extreme learning machine · Hybrid structure

1 Introduction

Action recognition in videos has many applications like human-machine interactions, intelligent household and etc. One of the most important task is how to extract the video features. Before the upsurge of deep learning, handcrafted features like interest point based local descriptors were popular. Among these descriptors, the combination of HOG (Histograms of Oriented Gradients) and HOF (Histograms of Optical Flow) were widely used. The works [1–4] adopted bag-of-features representations after computing feature descriptors extracted from videos. Inspired by the success of dense sampled feature points in image classification, Wang [5] proposed dense trajectories to describe the features in video sequences. By densely sampling on multiple scales, the trajectories were obtained through the dense optical flow field, which also produced dense trajectory descriptors. Wang [6] then proposed an improved method iDT (improved dense trajectories), which matched the feature points and eliminated the effect of camera motion. Meanwhile, iDT used Fisher Vector to code features. To date, iDT still achieved high performance among handcrafted features methods.

Since the great success of CNN (Convolutional Neural Network), many methods of action recognition adopt CNN to capture and fuse the spatial and temporal features in videos. Simonyan [7] proposed two-stream (spatial stream and

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 641–648, 2021.
https://doi.org/10.1007/978-981-15-8458-9_68

temporal stream) network, which merge two kinds of features extracted by spatial 2D CNN and temporal 2D CNN. Feichtenhofer [8] deepened the two-stream network and fuse the two streams with convolutional kernels instead of simple concatenation. Zhang [9] replaced optical flows with motion vector to decrease the time of calculation. Zhu [10] proposed Temporal Pyramid Pooling to obtain different temporal features. A main advantage of two-stream is that they can be initialized by the parameters obtained from image datasets such as ImageNet because the kernel dimensions are the same.

3D CNN is another useful method. Ji [11] applied 3D CNN in video action recognition. 3D convolution operate convolution on a three-dimensional tensor composed of multiple frames of images in videos, then integrating the spatial and temporal information. Tran [12] improved the 3D convolutional network, adding a 3D pooling layer, and deepen the network in order to train it on a larger data set. He also proposed C3D network which was then used as a general video feature extraction method. Using Inception [13] extended to 3D, Carreira [14] achieved good performance. One problem caused by 3D convolutional networks is the dramatically increased parameters. To solve this problem of 3D convolution kernels, Sun [15] decomposed 3D convolution kernels into 2D convolution kernels and 1D convolution kernels. Hara [16] explored 3D residual network(ResNet) structures and verified deep structures can achieve higher accuracy. These 3D CNNs above can directly extract spatio-temporal features. In this paper we mainly focus on 3D CNNs.

ELM [17, 18] is a single-hidden layer feed forward network(SLFN), which can calculate the output weights in an efficient and dramatically fast speed. Meanwhile, experiments have shown ELM has a better performance on avoiding overfitting compared with other learning method concerned with gradient-tuning, such as BP(back-propagation). Many works in different areas have presented that ELM can be applied to classify or approximate different kinds of datasets. However, when facing the situation of learning complex features such as big image datasets and videos datasets, it is hard to apply ELM directly on them, because the single hidden layer framework limits its ability of extracting features.

The main contribution of this work is the proposal of a hybrid CNN-ELM network to process action recognition in videos. It combines deep residual 3D CNN and ELM, which can learn more features of videos and classify labels in a faster and more efficient way. We also present the design of the layers and comparisons in different aspects in detail. Finally, the experiments verify the efficiency and the feasibility of our model.

This paper is organized as follows. The Sect. 2 introduces the details of the proposed network. And Sect. 3 presents experimental results and analysis its performance detailedly. We draw the conclusions in Sect. 4.

2 Method Description

2.1 Extreme Learning Machine

In the theory of ELM [17, 18], the network does not have to fine-tune the weights of hidden layers constantly until the loss and accuracy meets the requirement. Instead, the input weights can be initialized by random values, if the activation function is differentiable in any interval. Then, the output matrices can be obtained by generalized in-verse calculation. Overall, This straight way to get hidden wights explains why ELM could run extremely fast compared to other traditional SLFNs.

Assumed that the datasets has N examples and O classes, and the network has M hidden nodes, we write the input matrices as $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]$ and labels as $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_O]$ respectively. $l(\mathbf{x}_i)$ means the activation functions. We could use Sigmoid, Fourier, Multiquadratics as the activation functions in ELM. The hidden layer matrices \mathbf{L} can be written as

$$\mathbf{L} = \begin{bmatrix} l_1(\mathbf{x}_1) & \dots & l_M(\mathbf{x}_1) \\ \vdots & \vdots & \vdots \\ l_1(\mathbf{x}_N) & \dots & l_M(\mathbf{x}_N) \end{bmatrix} \quad (1)$$

If $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_O]$ is the output weights, we then have an equation

$$\sum_{i=1}^M \beta_i l(\mathbf{x}_k) = \mathbf{t}_k \quad k = 1, 2, \dots, M \quad (2)$$

It is also equal with

$$\mathbf{L}\boldsymbol{\beta} = \mathbf{T} \quad (3)$$

In fact, ELM aims to achieve the minimum training loss as well as the minimum norm of weights. In other words, ELM tries to solve:

$$C\|\mathbf{L}\boldsymbol{\beta} - \mathbf{T}\| + \|\boldsymbol{\beta}\|^\sigma \quad (4)$$

where C is a hyper-argument to balance loss and weights. For most applications, N is unequal with M . Thus, the optimal results of the Eq. (3) $\mathbf{L}\boldsymbol{\beta} = \mathbf{T}$ can be solved through \mathbf{L} 's pseudo-inverse matrix.

$$\hat{\boldsymbol{\beta}} = \mathbf{L}^\dagger \mathbf{T} \quad (5)$$

if $\sigma = 2$ [17, 19], then the result of $\boldsymbol{\beta}$ can be given as:

$$\hat{\boldsymbol{\beta}} = \begin{cases} (\frac{I}{C} + \mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T \mathbf{T} & N > M \\ \mathbf{L}^T (\frac{I}{C} + \mathbf{L}^T \mathbf{L})^{-1} \mathbf{T} & N \leq M \end{cases} \quad (6)$$

The ELM output function is given as

$$f(\mathbf{x}_j) = h(\mathbf{x}_j) \boldsymbol{\beta} = \begin{cases} h(\mathbf{x}_j) (\frac{I}{C} + \mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T \mathbf{T} & N > M \\ h(\mathbf{x}_j) \mathbf{L}^T (\frac{I}{C} + \mathbf{L}^T \mathbf{L})^{-1} \mathbf{T} & N \leq M \end{cases} \quad (7)$$

2.2 3D Residual Network

3D Residual Network stems from ResNet [20]. ResNets introduce a shortcut framework between the networks, which aims to solve the degradation in deeper network. Through shortcuts the network can learn an identity function, which makes training effect at least not become worse when the network goes deepen.

Table 1 shows the 3D Residual architecture. We take 18-layer and 34-layer ResNets as examples and mainly presents the detail in convolution layers here. 3D convolution and 3D pooling were implemented in 3D ResNets, where the temporal stride in convolution layer is 1 and in pooling layer is 2.

Table 1. 3D Residual Network architecture(18-layer and 34-layer) [16]

Layer name	Architecture	
	18-layer	34-layer
conv1	$7 \times 7 \times 7, 64$, temporal stride 1, spatial stride 2	
conv2_layer	$3 \times 3 \times 3, 64$, max pool, stride 2	
	$\begin{bmatrix} 3 \times 3 \times 3, 64 \\ 3 \times 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 \times 3, 64 \\ 3 \times 3 \times 3, 64 \end{bmatrix} \times 3$
conv3_layer	$\begin{bmatrix} 3 \times 3 \times 3, 128 \\ 3 \times 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 \times 3, 128 \\ 3 \times 3 \times 3, 128 \end{bmatrix} \times 4$
conv4_layer	$\begin{bmatrix} 3 \times 3 \times 3, 256 \\ 3 \times 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 \times 3, 256 \\ 3 \times 3 \times 3, 256 \end{bmatrix} \times 6$
conv5_layer	$\begin{bmatrix} 3 \times 3 \times 3, 512 \\ 3 \times 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 \times 3, 512 \\ 3 \times 3 \times 3, 512 \end{bmatrix} \times 3$

2.3 Design of Our CNN-ELM Model

The proposed CNN-ELM model is shown in Fig. 1. In our model, we still need to tune the parameters of CNN during the former part of training. The difference is that initial 3D Residual Network uses fully connected layer and softmax classifier to classify the labels during the whole process, while our model would invoke ELM when the loss would not decrease. At that time, the ELM could compute the hidden layer weights and β matrices.

3 Experiments

3.1 Dataset

In this experiments, we used UCF101 dataset to verify the feasibility of our method. Collected from Internet, UCF101 is an action recognition dataset of

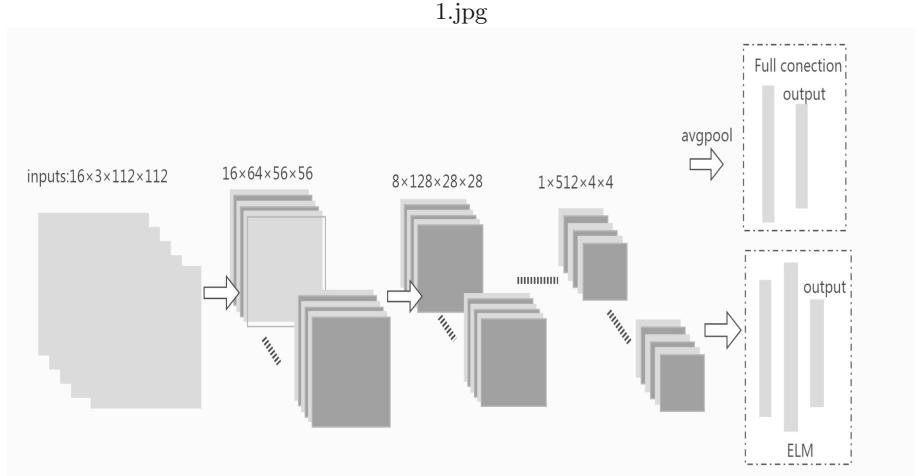


Fig. 1. The proposed framework of CNN-ELM for Action Recognition

realistic action videos with 101 action categories. From the 13320 videos in 101 action classes, UCF101 gives the greatest diversity, and there are large differences in camera movement, object representation and action, object size, observation point, clean background or not, light intensity, etc. This is still a challenging data up to now [21, 22].

3.2 Transfer Training

Transfer learning makes the network learn features fast and then save much time. In action recognition field, many work initial their model based on the parameters pretrained on ImageNet in order to arrive a better convergence point. However, ImageNet is a 2D convolutional network and is not suitable for pre-training our 3D CNN. In our work, before starting the training, our model is pre-trained on Kinetics dataset through which we could obtain a better start point.

3.3 Training

The code of our method is written in Pytorch and the available code of 3D ResNet [16]. All of the experiments are trained on GTX 1080TI. At first, we resize the input images as 112×112 pixels, and the size of input clips is $3 \times 16 \times 112 \times 112$. All the clips are multi-scale cropped in pre-processing stage. We use stochastic gradient descent(SGD) with 0.001 weight decay and momentum 0.9. Meanwhile, we set the learning rate as 0.1 at training stage, which would decrease when the validation loss does not descend for several epoches.

Table 2. Performance comparisons with other methods on UCF101. The data comes from [23] except for the citations and our methods.

Method	Accuracy
Two-stream method	
Two-stream ConvNet	73.0%
Two-stream + LSTM	82.6%
Two-stream fusion	82.6%
Long-term temporal ConvNet	82.4%
TSN	85.7%
Two-Stream I3D [14]	88.8%
Methods with handcrafted features	
IDT	85.9%
C3D+IDT	90.4%
3D CNN	
C3D	82.3%
P3D ResNet	88.6%
T3D [24]	90.3%
Our Method	92.7%

3.4 Result

We first compare with several popular methods in the context of video action recognition on UCF101 validation set. In the experiment we trained 3D ResNet-18. The performance comparisons are shown in Table 2. Like [23], we also briefly group the framework on UCF101 into 3 categories: Two-steam, 3D CNN, methods with handcrafted features. As shown in Table 2, the accuracy of our model can achieve 90.7%, making obvious improvement over original network ResNet by almost 5% and best competitor P3D by 2.1% respectively. This motivates us that if we use more advanced 3D architecture with ELM, the accuracy would get higher. Therefore, the potential of ELM in motion recognition is huge. Furthermore, the increase of accuracy could stem from ELM has better ability of separating different features than normal classifiers like softmax.

Table 3 shows the epochs that networks should take for a increase of 5%. We compare the original network and our proposed network. Clearly, although

Table 3. Epochs needed for same increase

MODEL	Time(s)/epoch
Our method	1
ResNet	80

the original network has tried its best to find the best solution, because of the gradient descent it still needs dozens of epochs to reach the same level as the proposed method. Every 1% increase the original network achieved consumes much time. While our method is efficient.

4 Conclusion

In this paper, we propose a CNN-ELM structure for action recognition in videos, using 3D residual networks and extreme learning machine. Making the best of the two networks, the proposed network can learn deep features and classify it in a faster and more efficient way. Meanwhile, we do not simply improve the performance but also give the settings in detail. Experiments has shown the clearly observed improvement. In the future, we will focus on how to make ELM more accessible to larger datasets, and the more details of this hybrid network such as the number of nodes or the model size. We may extend our method to other types of inputs like optical flow.

Acknowledgments. Our work is supported by Grant No.61973036, 61433003, from National Natural Science Foundation (NNSF) of China.

References

1. Laptev, I.: On space-time interest points. *Int. J. Comput. Vis.* **64**(2–3), 107–123 (2005)
2. Marszalek, M., Laptev, I., Schmid, C.: Actions in context. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2929–2936. IEEE (2009)
3. Qu, W., Zhang, Y., Feng, S., Wang, D., Yu, G.: Action-scene model for recognizing human actions from background in realistic videos. In: International Conference on Web-Age Information Management, pp. 566–577. Springer (2014)
4. Wang, H., Ullah, M.M., Klaser, A., Laptev, I., Schmid, C.: Evaluation of local spatio-temporal features for action recognition (2009)
5. Wang, H., Kläser, A., Schmid, C., Liu, C.L.: Action recognition by dense trajectories. In: CVPR 2011, pp. 3169–3176. IEEE (2011)
6. Wang, H., Kläser, A., Schmid, C., Liu, C.L.: Dense trajectories and motion boundary descriptors for action recognition. *Int. J. Comput. Vis.* **103**(1), 60–79 (2013)
7. Simonyan, K., Zisserman, A.: Two-stream convolutional networks for action recognition in videos. In: Advances in Neural Information Processing Systems, pp. 568–576 (2014)
8. Feichtenhofer, C., Pinz, A., Zisserman, A.: Convolutional two-stream network fusion for video action recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1933–1941 (2016)
9. Zhang, B., Wang, L., Wang, Z., Qiao, Y., Wang, H.: Real-time action recognition with enhanced motion vector CNNs. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2718–2726 (2016)
10. Zhu, W., Hu, J., Sun, G., Cao, X., Qiao, Y.: A key volume mining deep framework for action recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1991–1999 (2016)

11. Ji, S., Xu, W., Yang, M., Yu, K.: 3D convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(1), 221–231 (2012)
12. Tran, D., Bourdev, L., Fergus, R., Torresani, L., Paluri, M.: Learning spatiotemporal features with 3D convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4489–4497 (2015)
13. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
14. Carreira, J., Zisserman, A.: Quo vadis, action recognition. A new model and the kinetics dataset. *CoRR*, abs/1705.07750 2, 3 (2017)
15. Sun, L., Jia, K., Yeung, D.Y., Shi, B.E.: Human action recognition using factorized spatio-temporal convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4597–4605 (2015)
16. Hara, K., Kataoka, H., Satoh, Y.: Can spatiotemporal 3D CNNs retrace the history of 2D CNNs and imagenet? In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6546–6555 (2018)
17. Huang, G.B., Zhou, H., Ding, X., Zhang, R.: Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **42**(2), 513–529 (2011)
18. Huang, G.B., Zhu, Q.Y., Siew, C.K.: Extreme learning machine: theory and applications. *Neurocomputing* **70**(1–3), 489–501 (2006)
19. Huang, G.B., Ding, X., Zhou, H.: Optimization method based extreme learning machine for classification. *Neurocomputing* **74**(1–3), 155–163 (2010)
20. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. *Computing Research Repository*, pp. 770–778 (2015)
21. Jia, Y.: Robust control with decoupling performance for steering and traction of 4ws vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000)
22. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003)
23. Qiu, Z., Yao, T., Mei, T.: Learning spatio-temporal representation with pseudo-3D residual networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 5533–5541 (2017)
24. Diba, A., Fayyaz, M., Sharma, V., Karami, A.H., Arzani, M.M., Yousefzadeh, R., Van Gool, L.: Temporal 3D convnets: new architecture and transfer learning for video classification. arXiv preprint [arXiv:1711.08200](https://arxiv.org/abs/1711.08200) (2017)



Dynamic Modeling and Co-simulation for Active Suspension Systems

Huidong Hou, Shichang Han^(✉), Yingbo Huang, and Jing Na

Faculty of Mechanical and Electrical Engineering,
Kunming University of Science and Technology,
Kunming 650500, People's Republic of China
han_shichang@163.com

Abstract. A closed-loop dynamic simulator based on Adams and Matlab/Simulink software is modelled in this paper for active suspension systems to shorten the developing period, in which the active suspension system dynamic model and its control method are validated. A detailed 3D model of half-car active suspension system with hydraulic actuator is firstly built by using Solidworks. And then the virtual prototype of half-car active suspension system is established in Adams. Moreover, the proportion integration differentiation (PID) control method is designed to regulate the vehicle vertical displacement and pitch motion in Matlab/Simulink software. Finally, we combine the Adams and Matlab/Simulink software together to build a dynamic simulator, in which the correctness of the active suspension system dynamic model and effectiveness of the PID control method can be demonstrated.

Keywords: Active suspension systems · PID control · Co-simulation · Adams/Matlab

1 Introduction

With the rapid development of automotive industry, how to design superior vehicle suspension systems has become a hot topic. It is well-known that the suspension systems play an important role in ensuring driving safety and providing better maneuverability as one of the important components among vehicle chassis systems. Given the fact, different types of vehicle suspension structure have been designed. Normally, vehicle suspension systems can be grouped into three types, i.e., passive suspension system, semi-active suspension system and active suspension system. Compared with passive and semi-active suspension systems, active suspension systems can add or dissipate energy from the extra equipped actuator. Although active suspension systems have not been extensively applied in practical vehicle systems due to its complex mechanical structure and high energy consumption, its potential promising ability in improving driving comfort and safety has still attracted great attentions from both academical and industrial communities and considerable research work has been carried out [1–4].

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 649–656, 2021.

https://doi.org/10.1007/978-981-15-8458-9_69

In general, the dynamic model of active suspension system is built in Matlab/Simulink environment based on a set of specified dynamic equations among most available active suspension control research work [5–7]. Although this modeling procedure can contribute to designing the subsequent control method, the obtained results may create a gap to some extent between the theoretical study and practical application due to the ideal dynamic model. In this sense, how to design an elaborated active suspension system dynamic model to reflect the overall dynamic characteristics of active suspension system deserves to further investigate. It is well-known that Adams can build a virtual prototype before manufacturing the prototype, which can greatly reduce the cost of production and test time, and shorten the design cycle. This fact can be also found in vehicle active suspension control research work, especially for the initial stages of mechanical structure developments of active suspension system [8–10].

Apart from the dynamic model, control method is also important for active suspension system to achieve satisfactory performance and numerous advanced control methods are reported, such as adaptive control [11], sliding mode control [12] and H_∞ control method [13], etc. From the practical point of view, however, advanced control method may request that the practitioners should have solid background in control theory to implement the proposed control strategy in practical active suspension system successfully. This request may be quite stringent in practice. In this respect, the traditional proportion integration differentiation (PID) control method is still prevalent and does not fade in practical engineering control field.

Inspired by the above discussions, this paper will present a dynamic modeling and co-simulation approach for active suspension system. Firstly, a half-car active suspension dynamic model is built based on the theory of rigid body dynamics via the Solidworks and then imported into Adams to obtain an elaborate model. Then, the PID control method is designed in Matlab/Simulink to reduce the vertical and pitch motions. Finally, a combined dynamic simulator based on Adams and Matlab/Simulink are conducted to illustrate the correctness of the active suspension system modeling and the efficacy of the proposed PID control method.

2 Dynamic Modeling

The dynamic model of half-car active suspension system is generally described as follows as mentioned in [4,6]:

$$\begin{cases} M\ddot{y}_c + F_{df} + F_{sf} + F_{dr} + F_{sr} = u_y \\ I\ddot{\varphi} + a(F_{df} + F_{sf}) - b(F_{dr} + F_{sr}) = u_\varphi \\ m_f\ddot{y}_1 - F_{sf} - F_{dr} + F_{tf} + F_{bf} = -u_1 \\ m_r\ddot{y} - F_{sr} - F_{dr} + F_{tr} + F_{br} = -u_2 \end{cases} \quad (1)$$

where M and I represent the mass of vehicle body and the mass moment of inertia for the pitch motion; m_f and m_r represent the unsprung masses of front

and rear wheels; F_{df} and F_{sf} denote the forces produced by the front damper and spring; F_{dr} and F_{sr} denote the forces produced by the rear damper and spring; F_{tf} , F_{bf} , F_{tr} and F_{br} denote the elasticity forces and damping forces of the tires; y_c is the vertical displacement and φ is the pitch angle; y_1 and y_2 are the displacements of the unsprung mass; a and b refer to the distances between the suspension system to the vehicle body center; u_1 and u_2 are the actual control inputs. $u_y = u_1 + u_2$ and $u_\varphi = au_1 + bu_2$ are the control actions, which are used to calculate the actual control inputs as follows:

$$u_1 = \frac{bu_y + u_\varphi}{a + b}, u_2 = \frac{au_y - u_\varphi}{a + b} \quad (2)$$

In most available active suspension control research work, the dynamic model is generally directly built in Matlab/Simulink. In practice, however, each component of active suspension systems is regarded as rigid body, which means that there will exist inevitable influence between each rigid component. In this sense, considering the internal influence of each component during modeling procedure is important to achieve a complete model, which is contribute to perfectly and accurately reflecting the dynamic characteristic of active suspension systems.

Hence, we build a 3D model of half-car active suspension system in terms of Solidworks software, whose mechanical structure schematic is given in Fig. 1. This model is composed of three major components: the rigid body with dampers, springs and hydraulic actuators; two excitation platforms driven by the servo hydraulic system are placed at the bottom of the vehicle body to simulate the road excitation; two servo hydraulic actuators, which are connected to vehicle body at one end and to the front and rear wheels at the other end, respectively.

3 Co-simulation of Adams and Matlab

In this section, a combined dynamic simulator will be conducted based on Adams and Matlab/Simulink to verify the correctness and efficacy of the established half-car active suspension model and the associated active suspension control method, accordingly. Firstly, the obtained 3D model of half-car active suspension system via Solidworks will be imported into the Adams, in which an elaborated and completed half-car active suspension model will be established. Furthermore, the PID control strategy is designed in Matlab/Simulink to regulate the vehicle motions. The flow chart of the dynamic modeling and co-simulation process can be found in the Fig. 2.

3.1 Adams Setting

In order to obtain an elaborated model of half-car active suspension, the 3D model as shown in Fig. 1 is imported in Adams. The half-car active suspension model has four degrees of freedom, i.e., vertical displacement and pitch angle of the vehicle body, vertical displacement of the left and right wheels. In Adams, the components of the half-car active suspension model are set as rigid and all

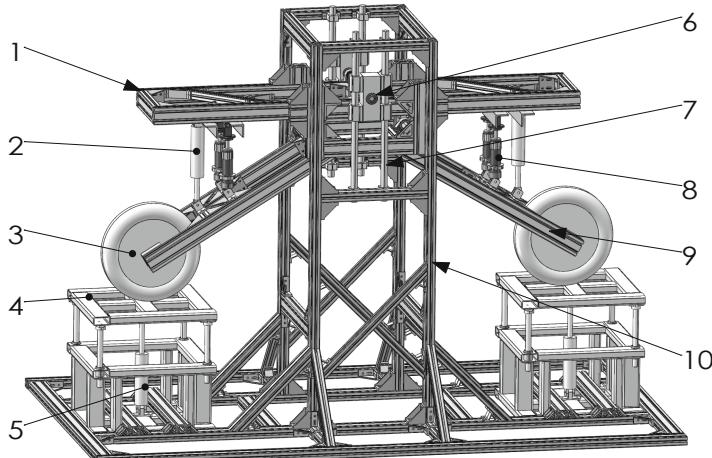


Fig. 1. 3D model of half-car active suspension system (1-Vehicle body. 2- Hydraulic actuator. 3-Wheel. 4-Vibration platform. 5-Hydraulic actuator. 6-Rotating bearing. 7-Linear guide. 8-Spring damper. 9-Swing arm beam. 10-Bracket)

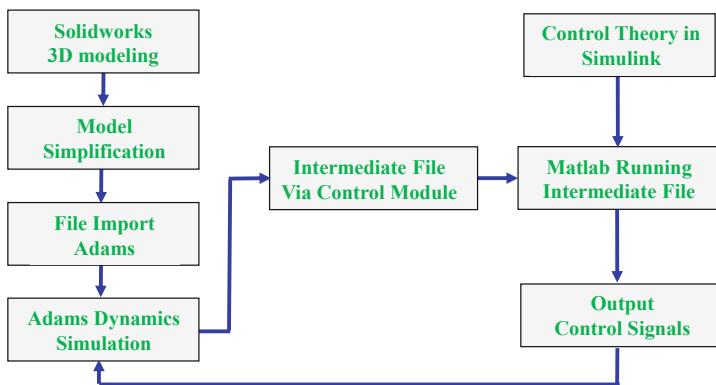


Fig. 2. Block diagram of the proposed co-simulation approach

motion pairs are rigid connections by default. The fiction and clearance between motion pairs are neglected. In virtual prototyping, the relevant parameter and main constraints joints setting in Adams are listed in Table 1.

3.2 Control Design

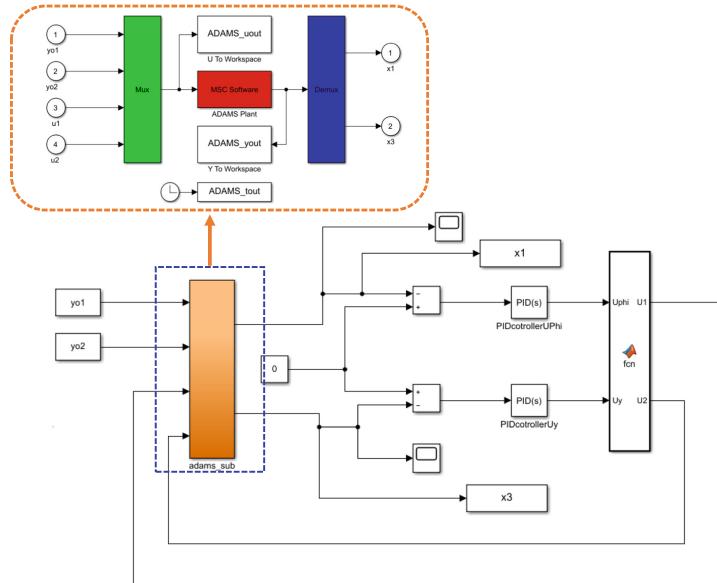
In this subsection, we will introduce a traditional PID control method to regulate the half-car active suspension system vertical displacement y_c and pitch motion φ . For this purpose, we define $e_1 = y_c$ and $e_2 = \varphi$ as the control error of vertical and pitch motions, respectively. Hence, the two PID controllers u_y and u_φ can be designed as:

Table 1. Parameter and main motion pairs setting

Joint	Type	Component 1	Component 2	Number
Joint1	Fixed Joint	Ground	Linear Guide	1
Joint2	Revolute Joint	Swing Arm Beam	Vehicle Body	2
Joint3	Revolute Joint	Hydraulic Actuator	Vehicle Body	2
Joint4	Revolute Joint	Hydraulic Pole	Swing Arm Beam	2
Joint5	Revolute Joint	Wheel	Swing Arm Beam	2
Joint6	Revolute Joint	Slider	Vehicle Body	1
Joint7	Cylinder Joint	Hydraulic Actuator	Hydraulic Pole	2
Joint8	Slider	Linear Guide	Vehicle Body	1

$$\begin{aligned} u_y(t) &= K_{P1}e_1(t) + K_{I1} \int_0^T e_1(t)dt + K_{D1} \frac{de_1(t)}{dt} \\ u_\varphi(t) &= K_{P2}e_2(t) + K_{I2} \int_0^T e_2(t)dt + K_{D2} \frac{de_2(t)}{dt} \end{aligned} \quad (3)$$

where K_{Pi} , K_{Ii} and K_{Di} , $i = 1, 2$ represent proportional, integral and differential coefficients.

**Fig. 3.** The structure of the combined dynamic simulator

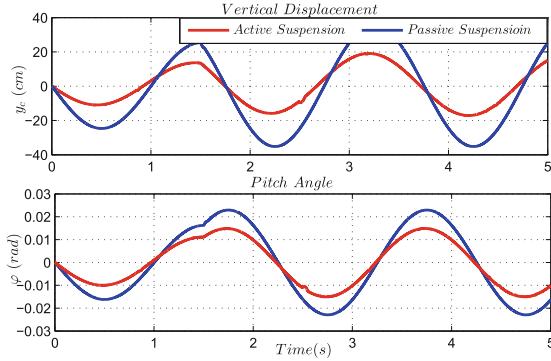


Fig. 4. Comparative results of vehicle motions

Therefore, the realistic control inputs u_1 and u_2 can be calculated based on (2) and then injected into Adams to regulate the vehicle motions.

4 Simulations

In this section, co-design simulation results based on Adams and Matlab/Simulink are provided to illustrate the effectiveness of the proposed method. The combined dynamic simulator diagram can be found in Fig. 3.

The parameters of half-car active suspension model in Adams are setted as $M = 120\text{ kg}$, $I = 60\text{ kgm}$, $m_f = 10\text{ kg}$, $m_r = 10\text{ kg}$, $k_{f1} = k_{r1} = 28000\text{ N/m}$, $k_{f2} = k_{r2} = 1000\text{ Ns/m}$, $b_c = b_e = 200000\text{ N/m}$, $b_{f2} = b_{r2} = 1500\text{ Ns/m}$, $a = b = 0.7\text{ m}$. The road disturbance is set as $y_{o1} = y_{o2} = 0.05\sin(\pi t)$. Moreover, we set the u_y and u_φ as 0, by which the active suspension can be regarded as passive suspension for comparison. Simulation results are given in Figs. 4 and 5. As shown in Fig. 4, the responses of vertical displacement and pitch angle with

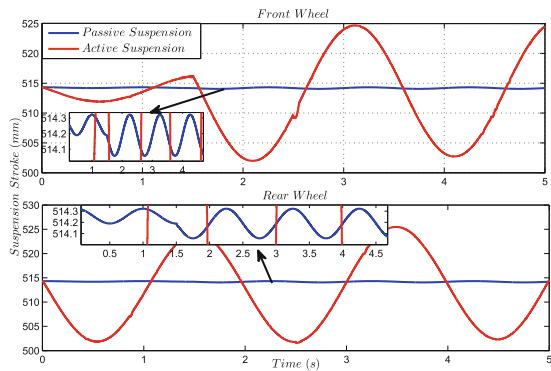


Fig. 5. Comparative results of suspension stroke

the proposed method have lower peaks compared with that of passive suspension, which means that the improved ride comfort is obtained with the proposed active suspensions and associated control method. Furthermore, the suspension stroke profiles are given in Fig. 5. One can find from Fig. 5 that the suspension stroke responses of active suspensions are larger than that of passive suspensions. This is natural because the active suspensions can provide better ride comfort as shown in Fig. 4, which means that larger suspension stroke is required in comparison to passive suspensions. Based on the yielded simulation results, we can conclude that the established dynamic model of half-car active suspension system can completely reflect the characteristics of active suspension system and the proposed active suspension control method can improve the ride comfort.

5 Conclusion

In this paper, a co-simulation approach for half-car active suspension system was proposed. A more elaborated dynamic model of half-car active suspension was constructed in terms of Adams, which is different with most available work that built dynamic model with Matlab/Simulink. Then, the traditional PID control method was employed in Matlab/Simulink to generate the control signal and then injected into Adams to regulate the vehicle motions. Simulation results were provided and demonstrated that the constructed dynamic model can reflect the characteristics of half-car active suspensions comprehensively and the vibrations of vehicle body can be significantly reduced.

Acknowledgment. This work was supported by the National Natural Science Foundation of China under Grant 61873115, and the Scientific Research Fund of Yunnan Education Department under Grant 2020J0067 and Grant 2019J0046.

References

1. Hrovat, D.: Survey of advanced suspension developments and related optimal control applications. *Automatica* **33**(10), 1781–1817 (1997)
2. Cao, D., Song, X., Ahmadian, M.: Editors' perspectives: road vehicle suspension design, dynamics, and control. *Veh. Syst. Dyn.* **49**(1), 3–28 (2011)
3. Cao, J., Liu, H., Li, P., Brown, D.J.: State of the art in vehicle active suspension adaptive control systems based on intelligent methodologies. *IEEE Trans. Intell. Transp. Syst.* **9**, 392–405 (2008)
4. Huang, Y., Na, J., Wu, X., Gao, G.: Approximation-free control for vehicle active suspensions with hydraulic actuator. *IEEE Trans. Ind. Electron.* **65**(9), 7258–7267 (2018)
5. Lian, R.-J.: Enhanced adaptive self-organizing fuzzy sliding-mode controller for active suspension systems. *IEEE Trans. Ind. Electron.* **60**(3), 958–968 (2012)
6. Sun, W., Gao, H., Kaynak, O.: Adaptive backstepping control for active suspension systems with hard constraints. *IEEE/ASME Trans. Mechatron.* **18**(3), 1072–1079 (2013)

7. Rath, J.J., Defoort, M., Karimi, H.R., Veluvolu, K.C.: Output feedback active suspension control with higher order terminal sliding mode. *IEEE Trans. Ind. Electron.* **64**(2), 1392–1403 (2016)
8. Azadeh, F., Reza, K.: Fuzzy control for active suspension in ADAMS/car full vehicle. In: Proceedings of 25th Chinese Control Conference, pp. 1188–1193 (2006)
9. Michele, I., Patrizio, T., Mauro, M.: Development of a heavy truck semi-active suspension control. *Control Eng. Pract.* **14**(3), 305–312 (2006)
10. Atte, R., Enso, I., Toni, L.: Validation of a nonlinear two-dimensional MacPherson suspension system model with multibody simulations. In: Proceedings of 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), pp. 1–5 (2018)
11. Huang, Y., Na, J., Wu, X., Liu, X., Guo, Y.: Adaptive control of nonlinear uncertain active suspension systems with prescribed performance. *ISA Trans.* **54**, 145–155 (2015)
12. Li, H., Yu, J., Hilton, C., Liu, H.: Adaptive sliding-mode control for nonlinear active suspension vehicle systems using T-S fuzzy approach. *IEEE Trans. Ind. Electron.* **60**(8), 3328–3338 (2013)
13. Chen, H., Guo, K.-H.: Constrained H_∞ control of active suspensions: an LMI approach. *IEEE Trans. Control Syst. Technol.* **13**(3), 412–421 (2005)



Design of Terminal Sliding Mode Adaptive Controller for Multiple Mobile Manipulators with Flexible Joints

Baigeng Wang and Shurong Li^(✉)

Automation College, Beijing University of Posts and Telecommunications,
Beijing 100876, China
upc_wbg@163.com, lishurong@bupt.edu.cn

Abstract. In this paper, a terminal sliding mode adaptive control algorithm is proposed for the control of multiple mobile manipulators. Firstly, the model of each mobile manipulator is established, and the object model is built at the same time. In addition, the influence of the flexible joint is considered in each mobile manipulator model. Secondly, considering the uncertainty of the model, the adaptive algorithm is used to estimate the model error. Then, the terminal sliding mode control algorithm is used to design the controller, which makes the position tracking error asymptotically converge to 0 in finite time and the force error asymptotically converge to 0 in finite time. Finally, simulation example verifies the feasibility of the strategy.

Keywords: Multiple mobile manipulators · Terminal sliding mode control · Adaptive control · Flexible joint

1 Introduction

In this paper, our research object is multiple mobile manipulators system. In the working range, compared with the fixed manipulator, the mobile manipulator has a wider operating space. In terms of functionality, due to the existence of the manipulator, the multiple mobile manipulators are more functional than the simple mobile robot cars.

Compared with the traditional fixed manipulator, the mobile manipulator is more difficult to build a model and design a controller. In order to solve these problems, many researchers have made a lot of efforts and achieved a lot of results. In reference [1], Yingmin Jia designs a nonlinear input-output decoupling controller to solve the control problem of velocity subsystem described by acceleration/braking force. In reference [2], Naijian Chen proposes an adaptive sliding mode control algorithm based on backstepping in order to solve the problems of dynamic coupling, nonholonomic constraint and nonlinear disturbance in the mobile manipulator system. In reference [3], a improved LMI control strategy is designed to deal with the control problems of continuous time systems.

In reference [4], an adaptive motion/force control is proposed to estimate the uncertain parameters.

Different from the single manipulator modeling process, the multiple manipulators system modeling not only needs to build the manipulator model, but also to build the object model. And then the complete multiple manipulators system model could be obtained. In reference [5], Dongya Zhao and Quanmin Zhu propose a new position synchronised control algorithm. In reference [6], in order to solve uncertain kinematics and dynamics problem of multiple manipulators, a new adaptive synchronised tracking control approach is developed. In reference [7], a class of extended inverse dynamics control algorithms combining with distributed sliding-mode estimators are developed to study the time-varying formation tracking for multiple manipulator systems. Many achievements have been made in the research of multiple manipulators system, but there are still some problems not considered to be perfect, such as flexible joint.

In this paper, the control problem of multiple mobile manipulators with flexible joints is studied. Firstly, the model of each mobile manipulator is established, and the object model is built at the same time. And then a complete system model is obtained by changing the coordinate system of the mobile manipulators and the object. Secondly, the adaptive algorithm is used to estimate the error of the model, and a controller based on terminal sliding mode is designed to guarantee that each mobile manipulator can accurately track the desired trajectory and the force error converges to 0 asymptotically. Finally, Lyapunov equation is used to verify the stability of the controller, and the simulation is carried out in MATLAB.

2 Dynamic Model

In order to facilitate the establishment of the dynamic model of the multiple mobile manipulators system, several assumptions are given as follow.

Assumption 1: The mobile platforms of all mobile manipulators are identical. The number of joints of the manipulator fixed on the platform is n_a and the degree of freedom is n_a .

Assumption 2: The object is a rigid regular object, the deformation is ignored, and there is no relative movement between the object and the end actuator of every mobile manipulator.

Assumption 3: The degree of freedom of the rigid object is the same as each mobile manipulator.

Considering the multiple mobile manipulators system composed of χn_a -link mobile manipulators, the dynamic model of the i th mobile manipulator in coordinate system $O_iX_iY_iZ_i$ can be written as follows:

$$M_i(q_i)\ddot{q}_i + C_i(q_i, \dot{q}_i)\dot{q}_i + G_i(q_i) + J_i^T(q_i)F_i = Z_i(q_i)\tau_i \quad (1)$$

where

$$M_i(q_i) = \begin{bmatrix} M_{ci} & M_{cai} \\ M_{aci} & M_{ai} \end{bmatrix}, C_i(q_i, \dot{q}_i) = \begin{bmatrix} C_{ci} & C_{cai} \\ C_{aci} & C_{ai} \end{bmatrix}, G_i(q_i) = \begin{bmatrix} G_{ci} \\ G_{ai} \end{bmatrix},$$

$$J_i^T = \begin{bmatrix} D_i^T & 0 \\ J_{ci}^T & J_{ai}^T \end{bmatrix}, F_i = \begin{bmatrix} F_{ci} \\ F_{ai} \end{bmatrix}, Z_i(q_i) = \begin{bmatrix} Z_{ci} & 0 \\ 0 & Z_{ai} \end{bmatrix}, \tau_i = \begin{bmatrix} \tau_{ci} \\ \tau_{ai} \end{bmatrix}.$$

where $q_i = [q_{ci}^T, q_{ai}^T]^T \in R^n$ denotes the state vector of mobile manipulator, $q_{ci} \in R^{n_c}$ is the state vector of the mobile platform in each mobile manipulator, $q_{ai} \in R^{n_a}$ is the state vector of each manipulator fixed on the mobile platform, $n = n_c + n_a$. $M_i(q_i) \in R^{n \times n}$ is the inertia matrix of each mobile manipulator. $C_i(q_i, \dot{q}_i) \in R^{n \times n}$ is the matrix of Coriolis force and centripetal force. $G_i(q_i) \in R^n$ is the matrix of gravity. $Z_i \in R^{n \times (n-h)}$ is the input conversion matrix. $\tau_i \in R^{n-h}$ denotes the vector of joint torque. $J_i(q_i) \in R^{(n-h) \times n}$ is coordinate transformation matrix from actuator to joint space. $F_i \in R^{n-h}$ is the force acting on the object by the actuator.

The flexible joints of each manipulator are expressed as follows:

$$J_{mi}\ddot{q}_{mi} + \tau_i = u_i, \quad \tau_i = K_{mi}(q_i - q_{mi}) \quad (2)$$

where J_{mi} is the rotational inertia of motor. q_{mi} denotes the motor rotation angle. u_i is the torque of motor. K_{mi} is the transfer coefficient matrix. In an ideal state, the joint is assumed to be completely rigid, $q_{mi} = q_i$, $\tau_i = u_i$.

Because the mobile platform of the mobile manipulator is constrained by h dimension nonholonomic constraints. Through kinematic analysis, the constraints on the mobile platform can be expressed as $D(q_{ci})\dot{q}_{ci} = 0$. Define the matrix $L(q_{ci}) \in R^{n_c \times (n_c-h)}$, $L^T(q_{ci})D^T(q_{ci}) = 0$, and then there is $\mu_i \in R^{n_c-h}$ which satisfies the following equations:

$$\dot{q}_{ci} = L(q_{ci})\dot{\mu}_i, \quad \ddot{q}_{ci} = \dot{L}(q_{ci})\dot{\mu}_i + L(q_{ci})\ddot{\mu}_i \quad (3)$$

Define $\gamma_i = [\mu_i^T, q_{ai}^T]^T \in R^{n-h}$, from Eq. (1) (2), the dimension reduction model of the manipulator can be expressed as follows:

$$M'_i(\gamma_i)\ddot{\gamma}_i + C'_i(\gamma_i, \dot{\gamma}_i)\dot{\gamma}_i + G'_i(\gamma_i) + {J'_{ai}}^T F_{ai} + Z'_i J_{mi} \ddot{q}_{mi} = Z'_i u_i \quad (4)$$

In Assumption 2 and Assumption 3, the object is known to be regular rigid and the degree of freedom is known, then Newton Euler method is used to model the object as follow:

$$M_w(x)\ddot{x} + C_w(x, \dot{x})\dot{x} + G_w(x) = J_w^T F_a = F_w \quad (5)$$

where $x, \dot{x}, \ddot{x} \in R^{n-h}$ denotes the position, velocity and acceleration vector of the object's centre of mass.

Define the generalized inverse matrix of matrix J_w^T as $(J_w^T)^+$. From Eq. (5), the force on the object at the end actuator of the manipulator is:

$$F_a = (J_w^T)^+ [M_w(x)\ddot{x} + C_w(x, \dot{x})\dot{x} + G_w(x)] + F_N \quad (6)$$

where F_N is the internal force of the object and does not affect the movement of the object, $J_w^T F_N = 0$.

Equation (6) is decomposed into each mobile manipulator model as follow:

$$F_{ai} = (J_{wi}^T)^+ [M_w(x)\ddot{x} + C_w(x, \dot{x})\dot{x} + G_w(x)] + F_{Ni} \quad (7)$$

The velocity vector of the end actuator is transformed into joint and task space through the transformation matrix:

$$\dot{x}_{ai} = J'_{ai}\dot{\gamma}_i, \quad \dot{x}_{ai} = J_{wi}\dot{x} \quad (8)$$

Combining Eq. (8), the following relation can be obtained:

$$\dot{\gamma}_i = J'_{ai}^{-1} J_{wi} \dot{x}, \quad \ddot{\gamma}_i = J'_{ai}^{-1} J_{wi} \ddot{x} + (J'_{ai}^{-1} \dot{J}_{wi} - J'_{ai}^{-1} J'_{ai} J'_{ai}^{-1} J_{wi}) \dot{x} \quad (9)$$

Take Eq. (7) and (9) into Eq. (4), and then the dynamic model of task space can be obtained as follows:

$$\bar{M}_i(x)\ddot{x} + \bar{C}_i(x, \dot{x})\dot{x} + \bar{G}_i(x) + \bar{J}_{ai}^T F_{Ni} + \bar{Z}_i J_{mi} \ddot{q}_{mi} = \bar{Z}_i u_i \quad (10)$$

In order to facilitate the controller design in the following chapters, it is necessary to decouple the position and force of the multiple mobile manipulators system. Multiply left by $J_{wi}^T J'_{ai}^{-T}$ on both sides of formula (10):

$$\tilde{M}_i(x)\ddot{x} + \tilde{C}_i(x, \dot{x})\dot{x} + \tilde{G}_i(x) + J_{wi}^T F_{Ni} + \tilde{Z}_i J_{mi} \ddot{q}_{mi} = \tilde{Z}_i u_i \quad (11)$$

From formula (4), the model of multiple mobile manipulators system composed of mobile manipulator is expressed as follows:

$$M'(\gamma)\ddot{\gamma} + C'(\gamma, \dot{\gamma})\dot{\gamma} + G'(\gamma) + J_a'^T(\gamma)F_a + Z'J_m \ddot{q}_m = Z'u \quad (12)$$

where

$$\gamma = [\gamma_1^T, \dots, \gamma_i^T, \dots, \gamma_\chi^T], \quad M'(\gamma) = \text{blockdiag}[M'_1(\gamma), \dots, M'_\chi(\gamma)].$$

From formula (8), The coordinate transformation relationship of the multiple mobile manipulators system can be obtained:

$$J_a'(\gamma)\dot{\gamma} = J_w(x)\dot{x} \quad (13)$$

By transforming formula (13) and deriving the two sides of the formula:

$$\begin{aligned} \dot{\gamma} &= J_a'^{-1} J_w \dot{x} \\ \ddot{\gamma} &= J_a'^{-1} J_w \ddot{x} + (J_a'^{-1} \dot{J}_w - J_a'^{-1} J_a' J_a'^{-1} J_w) \dot{x} \end{aligned} \quad (14)$$

Taken formula (6) and (14) into formula (12), dynamic model of multiple mobile manipulators system in task space can be obtained as follow:

$$\bar{M}(x)\ddot{x} + \bar{C}(x, \dot{x})\dot{x} + \bar{G}(x) + \bar{J}_a^T F_N + \bar{Z} J_m \ddot{q}_m = \bar{Z} u \quad (15)$$

Multiply left by $J_w^T J_a'^{-T}$ on both sides of formula (15):

$$\tilde{M}(x)\ddot{x} + \tilde{C}(x, \dot{x})\dot{x} + \tilde{G}(x) + J_w^T F_N + \tilde{Z} J_m \ddot{q}_m = \tilde{Z} u \quad (16)$$

According to the derivation of the dynamic model, the following properties can be obtained:

Property 1: Matrix $\tilde{M}(x)$ is positive definite symmetric and bounded.

Property 2: Matrices $\dot{\tilde{M}} - 2\dot{C}$ and $\tilde{M} - 2\tilde{C}$ are skew symmetric.

Property 3: Given a set of intermediate variables $\alpha \in R^n$, the dynamic equation of the manipulator is linearized as follows:

$$\tilde{M}\dot{\alpha} + \tilde{C}\alpha + \tilde{G} = Y(x, \dot{x}, \alpha, \dot{\alpha})P$$

where $Y(x, \dot{x}, \alpha, \dot{\alpha}) \in R^{n \times r}$ is regression matrix and $P \in R^r$ is parameter vector. In the same way, each single mobile manipulator system has the above properties.

In this paper, considering that the model has an upper bound but the upper bound is unknown, for the model (11), the inertia matrix, Coriolis force and centripetal force matrix and gravity matrix of the manipulator are rewritten as follows:

$$\begin{aligned} \tilde{M}_i(x) &= \tilde{M}_{ib}(x) + \Delta\tilde{M}_i(x) \\ \tilde{C}_i(x, \dot{x}) &= \tilde{C}_{ib}(x, \dot{x}) + \Delta\tilde{C}_i(x, \dot{x}) \\ \tilde{G}_i(x) &= \tilde{G}_{ib}(x) + \Delta\tilde{G}_i(x) \end{aligned} \quad (17)$$

where $\tilde{M}_{ib}(x)$, $\tilde{C}_{ib}(x, \dot{x})$ and $\tilde{G}_{ib}(x)$ represent the nominal model of the system; $\Delta\tilde{M}_i(x)$, $\Delta\tilde{C}_i(x, \dot{x})$ and $\Delta\tilde{G}_i(x)$ are the errors of manipulator model.

Take formula (17) into (11), and then

$$\tilde{M}_{ib}(x)\ddot{x} + \tilde{C}_{ib}(x, \dot{x})\dot{x} + \tilde{G}_{ib}(x) + J_{wi}^T F_{Ni} + \tilde{Z}_i J_{mi} \ddot{q}_{mi} = \tilde{Z}_i u_i + \rho_i \quad (18)$$

where $\rho_i = \Delta\tilde{M}_i(x) + \Delta\tilde{C}_i(x, \dot{x}) + \Delta\tilde{G}_i(x)$, $\|\rho_i\| \leq c_{i0} + c_{i1}\|x\| + c_{i2}\|\dot{x}\|^2$, c_{i0}, c_{i1}, c_{i2} are unknown constants.

According to Eq. (18), the multiple mobile manipulators system model can be written as follows:

$$\tilde{M}_b(x)\ddot{x} + \tilde{C}_b(x, \dot{x})\dot{x} + \tilde{G}_b(x) + \tilde{Z} J_m \ddot{q}_m = \tilde{Z} u + \rho \quad (19)$$

3 Controller Design

According to the multiple mobile manipulators system model (19) and each manipulator model (18), the following error equation is defined:

$$e = x - x_d, e_{fi} = F_{Ni} - F_{Nid}, e_{Fi} = J_{wi}^T (F_{Ni} - F_{Nid}) \quad (20)$$

where e is the position error of object, x_d is the desired position of the object. e_{fi} is the internal force error of the mobile manipulator actuator acting on the object, F_{Nid} is the desired internal force.

The terminal sliding mode surface with position error and force error is designed as follows:

$$s_i = \dot{e} + \Lambda_1 e^{a/b} + \Lambda_2 e_{F_i} \quad (21)$$

where Λ_1, Λ_2 are positive definite diagonal constant matrices. a, b are positive odd numbers, and $a < b < 2a$.

According to the nature of the object, the internal forces can offset each other. From Eq. (21), we can obtain:

$$s = \sum_{i=1}^{\chi} s_i = \chi(\dot{e} + \Lambda_1 e^{a/b}) \quad (22)$$

Define an intermediate variable as follows:

$$\dot{x}_{ri} = \dot{x}_d - \Lambda_1 e^{a/b} - \Lambda_2 e_{F_i} \quad (23)$$

Take Eq. (23) into Eq. (21) to get:

$$s_i = \dot{x} - \dot{x}_{ri} \quad (24)$$

The controller is designed as follows:

$$\begin{aligned} u_i = & \tilde{Z}^+ \left\{ \tilde{M}_{ib} \left(\ddot{x}_d + \left(-\frac{\Lambda_1 a}{b} \text{diag}(e^{\frac{a}{b}-1}) \dot{e} \right) - K_i s^{a/b} - \Lambda_2 \dot{e}_{F_i} \right) \right. \\ & + \tilde{C}_{ib} \dot{x} + \tilde{G}_{ib} - \left(\hat{c}_{i0} + \hat{c}_{i1} \|x\| + \hat{c}_{i2} \|\dot{x}\|^2 \right) \text{sgn} \left(\tilde{M}_{ib}^{-T} s_i \right) \\ & \left. + J_{wi}^T F_{Ni} + J_{mi} \ddot{q}_m \right\} \end{aligned} \quad (25)$$

where $\hat{c}_{i0}, \hat{c}_{i1}, \hat{c}_{i2}$ are the estimate of c_{i0}, c_{i1}, c_{i2} , and K_i is a positive definite constant diagonal matrix.

Design the adaptive rate of the estimate as follows:

$$\dot{\hat{c}}_{i0} = \kappa_{i0} \left\| \tilde{M}_{ib}^{-1} \right\| \|s_i\|, \quad \dot{\hat{c}}_{i1} = \kappa_{i1} \left\| \tilde{M}_{ib}^{-1} \right\| \|x\| \|s_i\|, \quad \dot{\hat{c}}_{i2} = \kappa_{i2} \left\| \tilde{M}_{ib}^{-1} \right\| \|\dot{x}\|^2 \|s_i\| \quad (26)$$

where $\kappa_{i0}, \kappa_{i1}, \kappa_{i2}$ are positive constants.

Define the estimation error as follows:

$$\tilde{c}_{i0} = c_{i0} - \hat{c}_{i0}, \quad \tilde{c}_{i1} = c_{i1} - \hat{c}_{i1}, \quad \tilde{c}_{i2} = c_{i2} - \hat{c}_{i2} \quad (27)$$

Derivative both sides of formula (27):

$$\dot{\tilde{c}}_{i0} = -\dot{\hat{c}}_{i0}, \quad \dot{\tilde{c}}_{i1} = -\dot{\hat{c}}_{i1}, \quad \dot{\tilde{c}}_{i2} = -\dot{\hat{c}}_{i2} \quad (28)$$

Select the following Lyapunov equation:

$$V_i = V_{i1} + V_{i2} = \frac{1}{2} s_i^T s_i + \frac{1}{2} \sum_{j=0}^2 \frac{\tilde{c}_{ij}^2}{\kappa_{ij}} \quad (29)$$

Derivative both sides of formula (29):

$$\dot{V}_i = \dot{V}_{i1} + \dot{V}_{i2} = s_i^T \dot{s}_i + \sum_{j=0}^2 \frac{\tilde{c}_{ij} \dot{\tilde{c}}_{ij}}{\kappa_{ij}} \quad (30)$$

From Eq. (21), we can obtain the expression of \dot{s} as follows:

$$\dot{s}_i = \ddot{x} - \ddot{x}_d + \frac{A_1 a}{b} \text{diag} \left(e^{\frac{a-b}{b}} \right) \dot{e} + A_2 \dot{e}_{F_i} \quad (31)$$

Take formula (18) and control law (25) into formula (31):

$$\dot{s}_i = -K_i s_i^{a/b} + \tilde{M}_{ib}^{-1} \rho - \tilde{M}_{ib}^{-1} \psi \text{sgn} \left(\tilde{M}_{ib}^{-1} s_i \right) \quad (32)$$

where $\psi = \hat{c}_{i0} + \hat{c}_{i1} \|x\| + \hat{c}_{i2} \|\dot{x}\|^2$.

Bring the adaptive law (26), Eq. (18) and Eq. (32) into Lyapunov equation (30):

$$\begin{aligned} \dot{V}_i &= s_i^T \left(-K_i s_i^{a/b} + M_{ib}^{-1} \rho - M_{ib}^{-1} \psi \text{sgn} \left(M_{ib}^{-1} s_i \right) \right) - \\ &\quad \tilde{c}_{i0} \left\| \tilde{M}_{ib}^{-1} \right\| \|s_i\| - \tilde{c}_{i1} \|x\| \left\| \tilde{M}_{ib}^{-1} \right\| \|s_i\| - \tilde{c}_{i2} \|\dot{x}\|^2 \left\| \tilde{M}_{ib}^{-1} \right\| \|s_i\| \\ &\leq -s_i^T K_i s_i^{a/b} \leq 0 \end{aligned} \quad (33)$$

According to Lyapunov stability criterion, s_i , c_{i0} , c_{i1} , c_{i2} are all bounded. Because the terminal sliding mode is adopted in this control scheme, the finite time convergence can be further proved.

According to formula (26), the adaptive laws are not negative. The integral from 0 to t for the adaptive law can be obtained as follows:

$$\begin{aligned} \tilde{c}_{i0} &= c_{i0} - \hat{c}_{i0} = c_{i0} - \hat{c}_{i0}(0) - \int_0^t \kappa_{i1} \left\| M_{ib}^{-1} \right\| \|s_i\| dt \\ \tilde{c}_{i1} &= c_{i1} - \hat{c}_{i1} = c_{i1} - \hat{c}_{i1}(0) - \int_0^t \kappa_{i1} \left\| \tilde{M}_{ib}^{-1} \right\| \|x\| \|s_i\| dt \\ \tilde{c}_{i2} &= c_{i2} - \hat{c}_{i2} = c_{i2} - \hat{c}_{i2}(0) - \int_0^t \kappa_{i2} \left\| \tilde{M}_{ib}^{-1} \right\| \|\dot{x}\|^2 \|s_i\| dt \end{aligned} \quad (34)$$

By selecting the appropriate value of $\hat{c}_{i0}(0)$, $\hat{c}_{i1}(0)$ and $\hat{c}_{i2}(0)$, ensure that $\hat{c}_{i0}(0) > c_{i0}$, $\hat{c}_{i1}(0) > c_{i1}$ and $\hat{c}_{i2}(0) > c_{i2}$. And then, it can be seen that $\tilde{c}_{i0} < 0$, $\tilde{c}_{i1} < 0$ and $\tilde{c}_{i2} < 0$.

It can be seen from the above derivation process that $\sum_{j=0}^2 \frac{\tilde{c}_{ij} \dot{c}_{ij}}{\kappa_{ij}} > 0$, and then formula (30) can be rewritten as follows:

$$\dot{V}_{i1} \leq \dot{V}_i \leq -s_i^T K_i s_i^{a/b} \leq -\alpha_i \left(\sum_{j=1}^n \frac{1}{2} s_{ij}^2 \right)^\nu = -\alpha_i V_{i1}^\nu \quad (35)$$

where $\alpha_i = k_{i\min} 2^\gamma > 0$, $k_{i\min}$ is the minimum elements of positive definite diagonal matrix K_i , $0 < \gamma = \frac{a/b+1}{2} < 1$.

Therefore, V_{i1} converges to 0 in finite time. And then, s_i also converges to 0 in finite time. Further, s converges to 0 in finite time.

In order to prove the finite time convergence of object position error, select another Lyapunov equation:

$$V_e = \frac{1}{2} e^T e \quad (36)$$

It is known that s converges to 0 in finite time. Therefore, it can be obtained from formula (22):

$$\dot{e} = -A_1 e^{a/b} \quad (37)$$

Derivative both sides of Eq. (36):

$$\dot{V}_e = -A_1 e^T e^{a/b} \leq -\lambda_{\min} 2^\gamma \left(\sum_{i=1}^n \frac{1}{2} e_i^2 \right)^\nu = -\alpha \bar{V}^\nu \quad (38)$$

where $\alpha = 2^\gamma \lambda_{\min} > 0$, λ_{\min} is the minimum elements of positive definite diagonal matrix A_1 , $0 < \nu = \frac{a/b+1}{2} < 1$.

Therefore, \bar{V} converges to 0 in finite time. And then, the position error e also converges to 0 in finite time. Further, the vector error \dot{e} converges to 0 in finite time.

Bring e and \dot{e} into the sliding mode surface (21), it can be seen that the internal force error e_{Fi} converges to 0 in finite time. Furthermore, from Eq. (20), e_{fi} also converges to 0 in finite time.

4 Simulation

The simulation model adopts a multiple mobile manipulators system composed of double two-link mobile manipulators. The controller simulation parameters are as follows:

$$\begin{aligned} A &= \text{diag}(1), \Gamma = \text{diag}(1), K = \text{diag}(10), K_f = \text{diag}(0.5), \\ \kappa_{i0} &= 1, \kappa_{i1} = 1, \kappa_{i2} = 1, \hat{c}_{i0} = 0.5, \hat{c}_{i1} = 0.5, \hat{c}_{i2} = 0.5, a = 3, b = 5. \end{aligned}$$

The initial position of the object's center of mass and the expected trajectory of the object are respectively:

$$x(0) = \begin{bmatrix} 1 \\ 5 \\ 2 \\ 1 \end{bmatrix}, x_d = \begin{bmatrix} 2.5 \sin t \\ 2.5 \cos t \\ 2d_1 + 2d_2 \sin(\theta_2) \\ \sin t + \cos t \end{bmatrix}$$

The simulation results of the double mobile manipulators system are shown in the figure. Among them, Figs. 1, 2 and 3 show the track tracking and its error of the object. It can be seen that the actual object track can be traced to the desired track through control. The internal force error simulation selects the mobile manipulator 1 as the simulation object, and its internal force error to the object is shown in Fig. 4. In this paper, terminal sliding mode adaptive control strategy has been presented systematically to control the multiple mobile manipulators system which carry a rigid object in the presence of flexible joint and model uncertain. Aiming at the uncertainty of the model, the model is rewritten into the form of nominal model plus error model, and the error is estimated by the adaptive algorithm. Further, the terminal sliding mode controller is designed to make the system motion converge to the desired trajectory and ensure the internal force converge to 0. The simulation results show the effectiveness of the control algorithm.

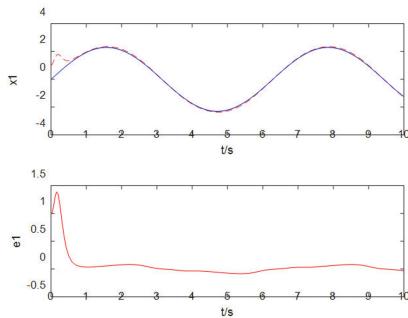


Fig. 1. Position tracking and error of object in x-axis

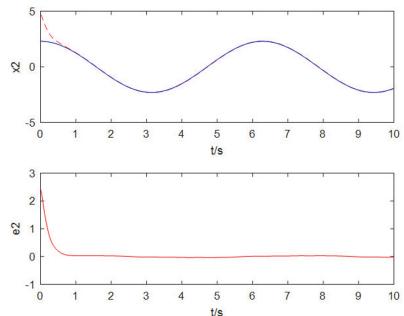


Fig. 2. Position tracking and error of object in y-axis

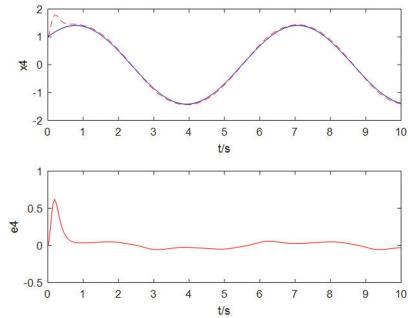


Fig. 3. Tracking and error of object rotation angle

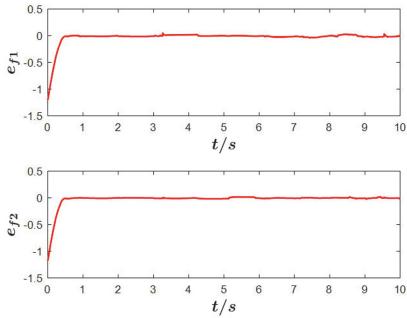


Fig. 4. The internal force error of manipulator

References

1. Jia, Y.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000)
2. Chen, N., Song, F., Li, G., et al.: An adaptive sliding mode backstepping control for the mobile manipulator with nonholonomic constraints. *Commun. Nonlinear Sci. Numer. Simul.* **18**(10), 2885–2899 (2013)
3. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003)
4. Boukattaya, M., Mezghani, N., Damak, T.: Adaptive motion/force control of uncertain nonholonomic mobile manipulator with estimation of unknown external force. *Multibody Syst. Dyn.* **44**, 223–250 (2018)
5. Zhao, D., Zhu, Q.: Position synchronised control of multiple robotic manipulators based on integral sliding mode. *Int. J. Syst. Sci.* **45**(1–3), 556–570 (2014)

6. Zhao, D., Li, S., Zhu, Q.: Adaptive synchronised tracking control for multiple robotic manipulators with uncertain kinematics and dynamics. *Int. J. Syst. Sci.* **47**(4), 791–804 (2015)
7. Ge, M., Guan, Z., Yang, C., et al.: Time-varying formation tracking of multiple manipulators via distributed finite-time control. *Neurocomputing* **202**, 20–26 (2016)



CT Tracing Algorithm Based on Optical Flow Field

Lei Xiao^(✉), Minghai Xu, and Zhongyi Hu^(✉)

College of Computer and Artificial Intelligence, Wenzhou University, Wenzhou, China
 {xiaolei,huzhongyi}@wzu.edu.cn, xmhemail@126.com

Abstract. Currently, there have already been plenty of visual tracing algorithms that were put forward in succession, but in the field of ship tracking, it has still been faced with challenge. The article has mainly solved problem brought by variation disturbance of ship tracking light and contrast in CCTV system. On the basis of compressing tracking frame, the article has put forward a kind of CT ship tracking algorithm based on optical flow field improvement. The article has reduced the alternating sample by above 50% to obtain faster tracking performance by utilizing the CT tracing solving scheme improved by optical flow field. At the same time, the article has improved the robustness and accuracy of tracking algorithm under illumination variation disturbance. The article has conducted verification and analysis on plenty of shipping tracking collection, and the algorithm of this article can realize more accurate tracking to the ship target.

Keywords: Optical flow · Compressing tracking · Random projection

1 Introduction

The target tracking based on machine vision has been the study direction that has possessed extreme challenge in the computer vision field study, which has attracted sights of many research scholars within the industry, the article has put into numerous manpower and material resources to conduct analytical research to the related subjects. With the constant enhancement of our country's comprehensive national strength, effect of harbour region shipping to economic development of our country has been far-reaching day by day, which has been the important component of transport system, whose main advantages have included big freight volume, low energy consumption, low environmental protection as well as cost, etc. [1]. In order to establish comprehensive and high efficient modernized shipping safety supervision system, the objectivity must ensure the accuracy, robustness of shipping tracking as well as timeliness of shipping tracking analysis. Although plenty of visual tracing algorithms have been put forward in succession currently, they have still faced with challenges [1–8]. As the target shipping may suffer all kinds of appearance variance, such as the treatment to

shielding disturbance, scale variation, etc., which often have caused the accuracy of tracing is not that high. For this purpose, the article has put forward a kind of CT shipping tracking algorithm based on optical flow field improvement.

2 Introduction of Optical Flow Field Method

The optical flow field [9] has been the method inferring direction of object movement speedometer by detecting change rule with time of brightness for picture pixel point. Its basic principle is: set velocity vector for every pixel to form the sports ground corresponding with picture, get the corresponding relation between the pixel point on the picture and the actual object by utilizing its projection relationship at some point, then find out the sports object according to variance character of the vector, which therefore is to extract and divide sports objects [9–11]. The optical flow filed has reflected the variance tendency of every point brightness in the picture, which can be regarded as the instantaneous flow field generated by picture's movement on the plane, which has also been a kind of approximate evaluation to real sports ground.

The optical flow field can be regarded as the velocity vector field of object, which has contained two components (u, v) . Assume the time interval between adjacent two frames pictures have been very little, at the same time when the brightness variance of picture has been very small, its optical flow constraint equation can be got as:

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (1)$$

Make first order taylor series expansion, it can get:

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt \quad (2)$$

Ie.:

$$I_x dx + I_y dy + I_t dt = 0 \quad (3)$$

Command $u = \frac{dx}{dt}$, $v = \frac{dy}{dt}$, then:

$$I_x u + I_y v = -I_t \quad (4)$$

$I(x, y, t)$ has been the brightness value of pixel point (x, y) in t moment, u, v have respectively been the optical flow component of this point along x, y directions, I_x, I_y, I_t have respectively been the partial derivative of brightness value I to x, y and t . As within very short time dt range, the picture brightness is the fixed value, the range of movement is narrow, at the same time of keeping spatial consistency, therefore formula (1) and formula (4) are tenable. According to the assumed condition, the article has calculated the movement of pixel point position for adjacent frames among t to $t + dt$. Main target of optical flow method is to solve optical flow constraint equation by utilizing least square method to obtain a solution u, v approximate to the ideal assumed condition.

Advantage of optical flow field method is that it does not need to consider the prior knowledge of background, the segmentation is accurate, the segmentation effect to condition of mutational background has been very good; its deficiency lies in that the algorithm complexity is rather high, the anti-noise performance is not that good, in addition, if the shelter or external light variation condition appear, the optical flow field is not equal to sports ground, then the optical flow field method will lose efficacy.

3 Compressive Tracking Frame

Based on the compressive sensing theory, the literature [2] has designed a kind of very sparse casual observation matrix, which has extracted the compressive domain character of the object to establish the appearance model of the object:

$$\mathbf{y} = \mathbf{Rx} \quad (5)$$

In formula (5), $\mathbf{x} \in \Re^m$ has expressed vector quantity of high-dimensional space, $\mathbf{y} \in \Re^m$ has expressed vector quantity of low-dimensional space. $\mathbf{R} \in \Re^{m \times m}$ has expressed casual observation matrix, whose element is defined as:

$$r_{i,j} = \sqrt{s} \times \begin{cases} 1 & \text{with probility } \frac{1}{2s} \\ 0 & \text{with probility } 1 - \frac{1}{s} \\ -1 & \text{with probility } \frac{1}{2s} \end{cases} \quad (6)$$

By analyzing literature [12, 13] it can get that the extracted compressive domain has been the linear combination of Haar-like character in essence, which is shown as formula (7):

$$y_i = \sum_{k=1}^{NR} r_{i,k} rect_{i,k} \quad (7)$$

In formula (7), $rect(\cdot)$ has expressed the rectangle character frame generated casually, NR has expressed sum total of rectangle frame, value of $r_{i,k}$ has been 1 or -1 , assume elements in are mutually independent, the confidence coefficient $H(y)$ can be calculated from Naive Bayes Classifier:

$$H(\mathbf{y}) = \sum_{i=1}^n \log \left(\frac{p(y_i|Y=1)}{p(y_i|Y=0)} \right) \quad (8)$$

$\mathbf{Y} \in \{0, 1\}$ has expressed label of sample. Literature [2] has modeled numerator and denominator of fraction in formula (8) as Gaussian distribution.

In order to adapt to the appearance variance during object movement process, classifier parameter dynamic updating mode is as below:

$$p(y_i|Y=1) \sim N(\mu_i^1, \sigma_i^1) \quad (9)$$

$$p(y_j|Y=0) \sim N(\mu_j^0, \sigma_j^0) \quad (10)$$

In order to adapt to the appearance variance during object movement process, classifier parameter dynamic updating mode is as below:

$$\mu_i^1 \leftarrow \gamma \mu_i^1 + (1 - \gamma) \mu^1 \quad (11)$$

$$\sigma_i^1 \leftarrow \sqrt{[\gamma (\sigma_i^1)] + (1 - \gamma) (\sigma_i^1)^2 + \gamma (1 - \gamma) (\mu_i^1 - \mu^1)^2} \quad (12)$$

and $\mu^1 = \frac{1}{n} \sum_{k=0|y=1}^{n-1} v_i(k)$ have respectively expressed variance and mean value.

Finally, make the alternating sample whose H is biggest in formula (8) to be the position of object in the current frame.

Via analysis it is not difficult to find that the classifier construction and parameter updating method described in formula (9)–(12) have played the key role in the CT tracing frame. The article has extracted compressive domain character from the whole object region to establish object appearance model, as is shown by formula (9) and (10), then the article has updated classifier parameter by directly using formula (11) and (12) to adapt to the constant appearance variance of object. Under general simple scene, formula (9)–(12) can effectively describe the appearance as well as variance of object. However, when the object has experienced partial/overall shielding or cluttering background scene, CT tracking frame is easy to extract and learn the appearance character as well as variance of sheltering object/background to make the property of classifier degrade and cause the tracking drifting. Aiming at this, the article has relieved the sheltering tracking problem by using idea of sub region [14–19] as reference, which has solved cluttering background tracking problem via method of multiple characters fusion.

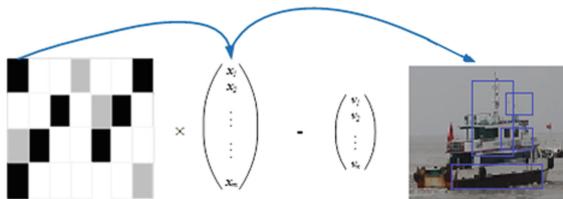


Fig. 1. Random projection of high-dimensional character

The original signal property has contained the complete information of signal, especially the original picture signal, which has contained rich color character, shape character, texture character as well as spatial relationship character. The expressing form of picture, that is to say to extract key character information of picture to conduct representation to picture, therefore, it has decided character of shipping object original image signal. The especially key point is that the range of random projection for shipping object can be confirmed from this, which therefore will restricted the range of picture signal character in low-dimensional

character space. In order for thorough comprehension and convenient calculation, the article has expanded visualized description in form of Fig. 1. In order to improve the robustness and accuracy based on random projection tracking algorithm, it needs to solve at the same time that the light variance disturbing algorithm is stable, it still needs to lower problem that contrast ratio variance disturbance affects tracking algorithm. Literature [20] has given out a kind of new tracking improved algorithm based on random projection, which has adopted phase equalization idea to lower the effect brought by light and contrast ratio disturbance. The method has combined phase equalization with response value, whose concrete steps are: in the $t - 1$ th frame, after obtaining the position of target ship, it has generated target ship alternation region of t th picture and it has made Fast Fourier Transform (FFT, Fast Fourier Transform), based on this, it has quested the position whose phase is most consistent in target ship alternating component, which therefore has obtained picture of phase equalization:

$$PC(x) = \frac{\sum_o \sum_n W_o(x) [A_{no}(x) \Delta\phi_{no}(x) - T_o]}{\sum_o \sum_n A_{no}(x) + \varepsilon} \quad (13)$$

Where the phase component of the n th Fourier series is recorded as $\phi_{no}(x)$, the index of its direction is marked as O, and he corresponding amplitude component is recorded as $A_n(x)$.

4 Experiment

As show in Fig. 2, A standard assumption in the classical control theory is that the data transmission required by the control or state estimation algorithm can be performed with infinite precision. However, due to the growth in communication technology, it is becoming more common to employ digital limited capacity communication networks for exchange of information between system components.

In order to verify the validity of this algorithm, the article has simultaneously adopted standard data set in visual tracking field and the harbor region shipping tracking video database to conduct experiment. Experiment parameter setting: sub domain quantity is 4, low-dimensional character dimension n is 50, searching radius of positive sample is 4, searching radius of negative sample is 8–20, learning rate of classifier γ is 0.85, threshold value of parameter updating is 0. The visual tracking algorithm detecting and evaluating method formulated by literature [21] has obtained wide recognition by researchers, therefore the article has adopted success and precision parameter quantitative evaluation tracking effect defined by [21] and has adopted spatial robustness evaluation(Spatial Robustness Evaluation, SRE), temporal robustness evaluation (Temporal Robustness Evaluation, TRE) and one-pass evaluation (One-pass Evaluation, OPE) three kinds of strategy evaluating algorithm defined by [21] to the initial sensitivity.

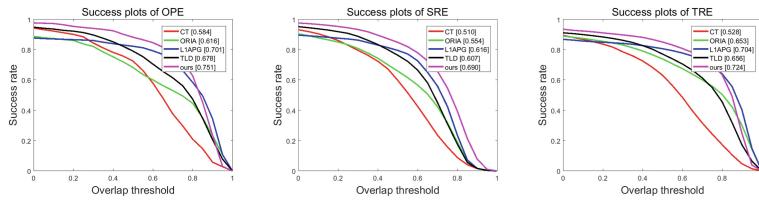


Fig. 2. Interface of computer terminal for remote control over host computer

5 Conclusion

The article has mainly put forward the CT tracking solving scheme based on optical flow field improvement aiming at shipping tracking to solve sample field problem. On the basis of compressive sensing, the article has conducted effective improvement to reduce the alternating sample by more than 50% and obtain faster tracking property. At the same time, method of this article has combined phase equalization with response value and it has improved the robustness and accuracy of tracking algorithm under light variance disturbance condition.

Acknowledgments. This study was financially supported by the Natural Science Foundation of Zhejiang Province Major Project (LZ20F020004), the Science and Technology Plan Major Science and Technology Projects of Wenzhou (ZY2019020), the Natural Science Foundation of Zhejiang Province (LY16F020022), and Wenzhou Science and Technology Planning Project (S20180017) of China.

References

- Peng, W.: Opportunity and challenge of inland navigation development in China. Port Waterw. Eng. **2**, 11–15 (2010)
- Zhang, K., Zhang, L., Yang, M.H.: Real-time compressive tracking., Florence, Italy, pp. 864–877 (2012)
- Ling, H.: Online robust picture alignment via iterative convex optimization, Providence, RI, United States, pp. 1808–1814 (2012)
- Bao, C., Wu, Y., Ling, H., et al.: Real time robust L1 tracker using accelerated proximal gradient approach, Providence, RI, United States, pp. 1830–1837 (2012)
- Kalal, Z., Mikolajczyk, K., Matas, J.: Tracking-learning-detection. IEEE Trans. Pattern Anal. Mach. Intell. **34**(7), 1409–1422 (2012)
- Lei, X., Huigang, W., Zhongyi, H.: Visual tracking via adaptive random projection based on sub-regions. IEEE Access **6**, 41955–41965 (2018)
- Lei, X., Minghai, X., Zhongyi, H.: Real-time inland CCTV ship tracking. Math. Prob. Eng. **2018**, 1–10 (2018)
- Xiao, L., Wang, H.G., Zou, M.L., et al.: A ship tracking algorithm of harbor channel based on orthogonal particles filter. In: Proceedings of 2018 Chinese Intelligent Systems Conference, vol. 1 (2019)
- Heeger, D.J.: Optical flow using spatiotemporal filters. Int. J. Comput. Vis. **1**(4), 279–302 (1988)

10. Young, S.I., Girod, B., Taubman, D.: Fast optical flow extraction from compressed video. *IEEE Trans. Pict. Process.* **PP**(99), 1 (2020)
11. Fei, X., Mengmeng, S., Ximming, G., et al.: Comprehensive evaluation of facial paralysis classification by combining the facial texture features and optical flow. *J. Northwest Univ. (Nat. Sci. Ed.)* (2019)
12. Fei, T., Liu, Q.: Multi-scale ship tracking via random projections. *Signal Pict. Video Process.* **8**(6), 1069–1076 (2014)
13. Fei, T., Liu, Q.: Robust multi-scale ship tracking via multiple compressed features fusion. *Signal Process. Pict. Commun.* **31**, 76–85 (2015)
14. Godec, M., Roth, P.M., Bischof, H.: Hough-based tracking of non-rigid objects. *Comput. Vis. Pict. Underst.* **117**(10), 1245–1256 (2013)
15. Lu, X., Zhang, J., Li, S., et al.: Person tracking with partial occlusion handling, Hangzhou, China, pp. 1–6 (2015)
16. Rui, Y., Shi, Q., Shen, C., et al.: Part-based robust tracking using online latent structured learning. *IEEE Trans. Circuits Syst. Video Technol.* **27**(6), 1235–1248 (2017)
17. Shu, G.: Part-based multiple-person tracking with partial occlusion handling, Providence, RI, United States, pp. 1815–1821 (2012)
18. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003)
19. Jia, Y.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000)
20. Yan, J., Wang, Z., Ai, S., et al.: Real-time tracking of targets with complex state based on ICT algorithm. *J. Huazhong Univ. of Sci. Tech.* (3), 107–112 (2015)
21. Yi, W., Jongwoo, L., Ming-Hsuan, Y.: Object tracking benchmark. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(9), 1834–1848 (2015)



Photovoltaic Power Forecasting Based on Randomized Multi-scale Kernels

Yang Deng¹, Yeqiong Liu², and Xuemei Dong^{3(✉)}

¹ School of Finance, Zhejiang Gongshang University, Hangzhou 310018, China

² School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018, China

³ School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China
dongxuemei@zjgsu.edu.cn

Abstract. Due to the randomness and instability of photovoltaic power generation, accurate power generation forecasting is of great significance to ensure grid stability and economic dispatch. This paper proposes a photovoltaic (PV) power output prediction model based on randomized multi-scale kernels, where the centers of the kernels are extracted by K-means method and the scales are sampled from a pre-defined uniform distribution. Through the experiment analysis, we believe that for generation data, establishing a model according to the season can improve the prediction accuracy. Compared with the existing models, the results show that the proposed method in this paper has a better effect on power generation prediction.

Keywords: Photovoltaic power generation · Prediction model · K-mean clustering · Multi-scale kernels

1 Introduction

With the urgency of energy and environmental issues becoming increasingly prominent around the world, renewable energy sources have attracted widespread attention. China is rich in solar energy resources, and PV power generation technology is gradually mature. However, due to the randomness and instability of PV power generation, large-scale grid-connected circuits will be affected to some extent. Therefore, an accurate prediction of PV power generation is significant to the development of PV power generation industry and energy supply.

Up to now, many PV power forecasting models have been proposed. Chen et al. [1] proposed a neural network prediction model with weather forecasting information, but this method depended heavily on environmental data and its computational complexity was high. Yu et al. [2] presented a support vector machine prediction model which used the fuzzy similarity matrix to classify the sample set and leave-one-out algorithm to optimize the parameters, but the

prediction accuracy still need to be improved. Zhao et al. [3] established a deep belief network model by using the reconstruction difference method to determine the hidden layers of the belief network, but the calculation was easy to fall into local convergence and the generalization ability was poor.

Considering the low prediction accuracy, poor generalization ability and long convergence time in the above models, this paper proposes a PV power generation prediction model based on randomized multi-scale Gaussian kernels. This model is a linear combination of multiple Gaussian kernels with different scales in which the scale parameters are randomly selected from a uniform distribution $U[0, \Omega]$, the centers are determined by the K-means method and the linear combination coefficients are solved by the l_2 regularization method. The results of real data analysis show that the model makes full use of the fitting ability of different scale kernels to different fluctuation characteristics of data, and it can achieve high prediction accuracy in the case of low computational complexity.

2 Randomized Multi-scale Kernels Regression Algorithm

We first introduce some related concepts and symbols.

Given a training sample set

$$S = \{z_i = (\mathbf{x}_i, y_i) : i = 1, 2, \dots, n\},$$

where each z_i is independently and identically distributed extracted from space $Z = X \times Y$ and $X \subset R^d$ denotes the input space, $Y \subset R$ is the output space, the goal of a regression problem is to find a appropriate function f , which describes the relationship between the input variable x and the output variable y , from a pre-given function space based on S .

A commonly used function space is the reproducing kernel Hilbert space (RKHS), which is defined as,

$$\mathcal{H} = \{f : f(x) = \sum_{j=1}^m \beta_j K(\mathbf{x}, \mathbf{x}_j)\}, \quad (1)$$

where K is a kernel function, which can be a Sigmoid Kernel, a Polynomial Kernel or a Gaussian Kernel. And \mathbf{x}_j is generally selected from the input data.

2.1 Algorithm Introduction

Considering the limitation of the single-kernel learning ability [4,5] in (1), we introduce a method based on multiple kernels, that is, choose a Gaussian kernel set with m scales

$$\left\{ \exp\left\{-\frac{\|\mathbf{x} - \mathbf{x}_1\|^2}{2\sigma_1^2}\right\}, \exp\left\{-\frac{\|\mathbf{x} - \mathbf{x}_2\|^2}{2\sigma_2^2}\right\}, \dots, \exp\left\{-\frac{\|\mathbf{x} - \mathbf{x}_m\|^2}{2\sigma_m^2}\right\} \right\}.$$

Here, $\{\mathbf{x}_j : j = 1, 2, \dots, m\}$ are the centers of the kernels, which will be determined by K-means method. Generally, the kernels with small scales can fit the sharp

part of data while the kernels with large ones can fit the gentle part [6]. These scales will be randomly drawn from a uniform distribution $U[0, \Omega]$, in which, Ω can be obtained through cross-validation technique.

Based on this set of Gaussian kernels and the training data, we want to find a linear combination function,

$$f(x) = \sum_{j=1}^m \beta_j K_{\sigma_j}(\mathbf{x}, \mathbf{x}_j),$$

by minimizing the following l_2 regularization function,

$$R_{reg} = \frac{1}{n} \sum_{i=1}^n (y_i - f(\mathbf{x}_i))^2 + \lambda \|\boldsymbol{\beta}\|_2^2, \quad (2)$$

with $\|\boldsymbol{\beta}\|_2^2 = \sum_{j=1}^m \beta_j^2$ and λ being a regularization parameter, which can be solved by cross-validation technique.

Through matrix calculation, the optimal coefficient vector $\boldsymbol{\beta}$ is obtained as,

$$\boldsymbol{\beta} = (\mathbf{K}^T \mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{K}^T \mathbf{Y}, \quad (3)$$

where \mathbf{I} is the identity matrix, $\mathbf{Y} = (y_1, y_2, \dots, y_n)^T$ and \mathbf{K} is the kernel matrix,

$$\mathbf{K} = \begin{bmatrix} K_{\sigma_1}(\mathbf{x}_1, \mathbf{x}_1) & K_{\sigma_2}(\mathbf{x}_1, \mathbf{x}_2) & \dots & K_{\sigma_m}(\mathbf{x}_1, \mathbf{x}_m) \\ K_{\sigma_2}(\mathbf{x}_2, \mathbf{x}_1) & K_{\sigma_2}(\mathbf{x}_2, \mathbf{x}_2) & \dots & K_{\sigma_m}(\mathbf{x}_2, \mathbf{x}_m) \\ \vdots & \vdots & \ddots & \vdots \\ K_{\sigma_n}(\mathbf{x}_n, \mathbf{x}_1) & K_{\sigma_2}(\mathbf{x}_n, \mathbf{x}_2) & \dots & K_{\sigma_m}(\mathbf{x}_n, \mathbf{x}_m) \end{bmatrix}.$$

2.2 The Proposed Algorithm Description

The aforementioned process can be described as the following randomized multi-scale kernels learning (RMSKL) algorithm.

3 Experiment Results and Analysis

This section provides several experiment verification results on the data which is collected from a certain area of Hangzhou from 2016 to 2018. Considering the low efficiency of PV power generation at night, we select the data from 7:00–18:00 every day, finally we get a data set with size 23170. The features are recorded in Table 1. To eliminate the influence of the dimension, we scale all data into [?1,1]. Generally, we randomly select 75% data as the training set and the remaining 25% data as the test set.

In order to measure the performance of the model, we choose the Mean Square Error (MSE) as the metric which is given by,

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2,$$

Algorithm 1. The RMSKL algorithm

Input: Training set S , positive parameter $\Omega \in \{\Omega_t, t = 1, 2, \dots, T\}$, regularization parameter $\lambda \in \{\lambda_l, l = 1, 2, \dots, L\}$, and number of kernels m .

- 1: Select m kernel centers by K-means method.
- 2: **for** $t=1,2,\dots,T$ **do**
- 3: **for** $l=1,2,\dots,L$ **do**
- 4: Randomly draw $\sigma_j \in U[0, \Omega_t], j = 1, 2, \dots, m$,
- 5: Set $K_{\sigma_j} = \exp\{\|\mathbf{x} - \mathbf{x}_j\|^2 / 2\sigma_j^2\}$
- 6: $\mathcal{H} = \{f : f(x) = \sum_{j=1}^m \beta_j K_{\sigma_j}(\mathbf{x}, \mathbf{x}_j)\}$
- 7: Calculate $f = \arg \min_{f \in \mathcal{H}} \left\{ \frac{1}{n} \sum_{i=1}^n (y_i - f(\mathbf{x}_i))^2 + \lambda_l \|\beta\|_2^2 \right\}$
- 8: **end for**
- 9: **end for**
- 10: **return** f and corresponding Ω, λ .

where, \hat{y}_i represents the predicted value of the i th group of samples, y_i is the true value corresponding to the i th group of samples.

3.1 Experiment 1. Selection of the Kernel Center Coordinates and the Number

We determine the centers of kernels by K-means method and the number m of clusters by the test MSE, which is calculated by the algorithm of this article in the following experiment. Specifically, we select m from 50 to 2000, train at 50 intervals, and calculate the training MSE and the test MSE to draw the line chart in Fig. 1.

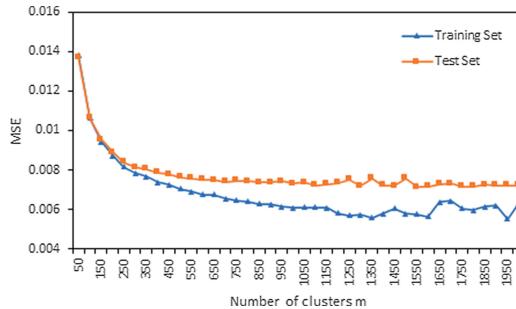


Fig. 1. Results of clustering

As we can see in Fig. 1, as m increases, that is, the number of multi-scale kernels increases, the training MSE and the test MSE are gradually reduced within a certain range. When m exceeds 1000, the training MSE still has some fluctuations, while the test MSE tends to be stable. In summary, we select the number of clusters is 1550 in the subsequent experiment using overall data modeling.

Table 1. Data description.

Variables/Unit	Data range	Mean	Std
Positive active/kWh	[497269, 9911.99]	7430.998	1413.127
Instantaneous active/kW	[0, 368.565]	116.4661	90.52297
Instantaneous reactive/Kvar	[0, 27.915]	5.885674	5.448669
AI/A	[0, 533.25]	170.8673	132.1055
BI/A	[0, 532.8]	169.8945	131.9285
CI/A	[0, 533.1]	170.4825	132.0312
AU/V	[216.9, 244.2]	228.1459	4.701758
BU/V	[219.3, 245.6]	229.788	4.519587
CU/V	[218.9, 244.6]	229.092	4.507124
Net total radiant illumination/Wm ⁻²	[1,866]	172.7555	165.9234
Total radiation exposure/0.01 MJm ⁻²	[0.01, 4.02]	1.326498	0.993778
Net total radiation exposure/0.01 MJm ⁻²	[0.01, 2.55]	0.618764	0.568434
Average of temperature/	[-1.6, 35.62]	18.81441	8.563422
Maximum daily temperature/	[2.8, 41.3]	23.52436	9.199896
Minimum daily temperature/	[-4.4, 29.7]	15.25909	8.401505

3.2 Experiment 2. Comparison of Fitting Results of Different Models to the Overall Data

We choose three models to compare with the RMSKL model, and the parameters of four models are selected as follows.

- SVM model with a RBF kernel, scale parameter $\gamma \in \{2^{-5}, 2^{-4}, \dots, 2^{10}\}$, and the penalty term $C \in \{10^{-20}, 10^{-19}, \dots, 10^5\}$.
- BP neural network model, which uses relu as its activation function, and the number of hidden layers $l \in \{1, 2, 3, 4\}$, the number of nodes $n \in \{10, 20, \dots, 300\}$.
- Linear ridge regression model, regularization parameter $\lambda \in \{10^{-20}, 10^{-19}, \dots, 10^5\}$.
- RMSKL model, the positive parameter $\Omega \in \{2^{-5}, 2^{-4}, \dots, 2^{10}\}$, and the regularized parameter $\lambda \in \{10^{-20}, 10^{-19}, \dots, 10^5\}$.

According to experiment 1, we choose m to be 1550. When Ω is 6.06287 and λ is $1e-20$, the results in Table 2 show that the training MSE and the test MSE of the RMSKL are lower than those of other three models. Meanwhile, the low variance of 20 times experiments demonstrates the stability of the RMSKL.

3.3 Experiment 3. Modeling by Season

Taking into account the seasonality of PV power generation, we establish four models for the four seasons respectively. Because of the randomness, we repeat the experiment 20 times and take the mean and the standard deviation in Table 3.

Table 2. Comparative results of four models.

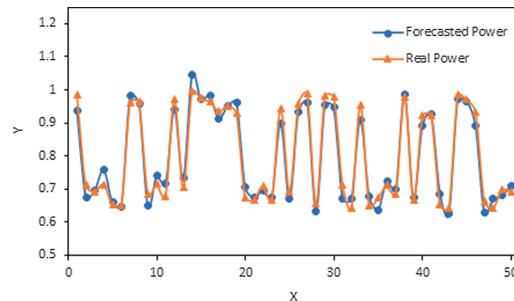
Models	Training MSE	Test MSE
RMSKL	0.005742 ± 0.000033	0.007140 ± 0.000086
SVM	0.006819	0.008422
BP	0.009008	0.009444
Ridge Regression	0.013558	0.013602

Table 3. Forecasting results of RMSKL for different seasons.

Seasons	m	Training MSE	Test MSE
Spring	900	0.003549 ± 0.000029	0.004434 ± 0.000059
Summer	640	0.005253 ± 0.000040	0.006771 ± 0.000089
Fall	750	0.005214 ± 0.000063	0.006628 ± 0.000222
Winter	550	0.004798 ± 0.000053	0.005842 ± 0.000184

As shown in Table 3, the test MSE in Spring is significantly lower than those in other seasons, and the forecasting results are more stable, which indicates that the Spring model has a better forecasting effect. In addition, the results of models based on seasons are better than those based on the overall data as in Table 2.

In order to represent the effect of the RMSKL model in predicting PV power generation, we randomly select 50 series of data in the test set to depict the change curve between the predicted value and the real value of Spring, and the data is well fitted in Fig. 2.

**Fig. 2.** Predicted results of Spring

To further study the influence of different sizes of training samples on the performance of four models, we choose 7280 data of Spring to carry out the following experiment. Then, we fix the test samples, continuously adjust the number of

training samples, and record the results of different models on different training samples with 20 trials (Table 4).

Table 4. Experiment results of four models.

Size	RMSKL-SCR		SVM	
	Training MSE	Test MSE	Training MSE	Test MSE
N = 1820	0.004762 ± 0.000067	0.005758 ± 0.000099	0.005935	0.009002
N = 2730	0.004233 ± 0.000037	0.005248 ± 0.000118	0.006253	0.007257
N = 3640	0.004090 ± 0.000034	0.004967 ± 0.000088	0.005458	0.005990
N = 4550	0.004007 ± 0.000029	0.004847 ± 0.000075	0.005012	0.005765
N = 5460	0.003549 ± 0.000029	0.004434 ± 0.000059	0.004630	0.005519
Size	BP		Ridge regression	
	Training MSE	Test MSE	Training MSE	Test MSE
N = 1820	0.010106	0.010454	0.010196	0.010014
N = 2730	0.008254	0.009023	0.010259	0.010019
N = 3640	0.006453	0.007717	0.010183	0.010030
N = 4550	0.007404	0.007912	0.010185	0.010051
N = 5460	0.006840	0.007436	0.010135	0.010015

It can be seen that the MSE of the RMSKL model is lower than other models, and the forecasting effect is the best. As the training samples increase in size, the MSEs of the RMSKL, BP and SVM models are gradually decreasing, indicating that the bigger size is beneficial to model performance, and under small training samples, the RMSKL model has better advantages.

4 Conclusions

This paper contributes to propose a PV power generation forecasting model based on randomized multi-scale kernels. Based on the overall data experiment, we find that the increase in the number of clusters is conducive to the forecasting effect. And using the model to compare with the other three models, the results show that the RMSKL model has better forecasting effect. By establishing the RMSKL model in different seasons, it is found that the model based on spring has the best forecasting effect, and the model with the bigger size has higher accuracy.

Acknowledgements. This work is supported partly by First Class Discipline of Zhejiang-A (Zhejiang Gongshang University-Statistics), Zhejiang college students science and technology innovation activity plan (Xinmiao talent plan).

References

1. Chen, C.S., Duan, S.X., Yin, J.J.: Design of photovoltaic array power generation prediction model based on neural network. *J. Electr. Eng.* **24**(09), 153–158 (2009)
2. Yu, Q.L., Xu, C.Q., Li, S., Liu, H., Song, Y., Liu, X.O.: Short-term photovoltaic power prediction based on fuzzy clustering and support vector machine. *J. Power Syst. Autom.* **28**(12), 115–118 (2016)
3. Zhao, L., Liu, Y.B., Yu, L.N., Liu, J.Y.: Short-term power generation forecast of photovoltaic power plants based on deep belief network. *Power Syst. Protect. Control* **47**(18), 11–19 (2019)
4. Wu, Q., Ying, Y., Zhou, D.X.: Multi-kernel regularized classifiers. *J. Complex.* **23**(1), 108–134 (2007)
5. Bao, Z.J., Pi, D.Y., Sun, Y.X.: Nonlinear model predictive control based on multi-core support vector machine (English). *Chin. J. Chem. Eng.* **05**, 691–697 (2007)
6. Dong, X.M., Weng, H., Shi, J., Gu, Y.H.: Randomized multi-scale Kernels learning with sparsity constraint regularization for regression. *Int. J. Wavelets Multiresolut. Inf. Process.* **17**(06), 13 (2019)



Environmental Sound Recognition Based on Residual Network and Stacking Algorithm

Haoyuan Wang, Xuemei Ren^(✉), and Zhen Zhao

School of Automation, Beijing Institute of Technology, Beijing 100081, China
xmren@bit.edu.cn

Abstract. Environmental sound recognition is one of the important tasks in the field of audio research. Because the environment is complex and there is a lot of useless sound information, the traditional methods have low recognition accuracy, which is gradually replaced by related methods of deep learning. In this paper, combined with the latest research in this field, the recognition algorithm based on residual network and stacking method is proposed. The whole is divided into two parts: a feature extractor and a classifier. The residual network is responsible for extracting features with high recognition rate and the stacking algorithm is responsible for accurate recognition. The method is applied to the representative datasets ESC-50 and UrbanSound8k. We obtain a higher accuracy and the model is more clear and simple.

Keywords: Environment sound recognition · MFCC · Residual network · Stacking algorithm

1 Introduction

At present, speech recognition is more mature in the field of audio recognition [1], which has been widely used in our life, such as the mobile phone voice assistant. In recent years, a large number of applications have also emerged in environmental sound recognition, such as the Internet of Things [2] that advocates the Internet of Everything and the field of autonomous driving [3,4]. These fields often face the test of complex outdoor environments, processing information from the environment. Relying only on image information, the judgment is usually incomplete [5,6]. The acoustic-based scene analysis and event detection can make up for the corresponding deficiencies [7].

The main difficulty affecting the accuracy of environmental sound recognition is that a large number of useless sounds are inevitably received in the process of sound collection, which seriously influence our accurate detection [8]. Therefore, two directions to solve this problem are to extract key features and use appropriate classifiers. The traditional environmental sound classifiers revolve around

feature extraction, such as log-mel [9], matrix decomposition [10], GTCC [11], MFCC [12] and mfcc-based improved feature FMF [13]. The classifier mainly uses machine learning methods such as GMM [14]. However, with the rapid development of deep learning, it has brought tremendous changes to the field of the sound recognition. The feature extraction introduces network structures such as CNN [15] and RNN [16]. The deep network structure can extract more distinguishing features or deep features, greatly improving recognition accuracy. In terms of classifiers, CNN [17] and other methods have also achieved good results.

This paper improves on the currently applied deep learning methods. The feature extractor combines the MFCC features and the residual network structure to avoid overfitting of complex network structures such as CNN, but maintains the ability to extract deep features. The classifier adopts the ensemble learning method, which can gather the advantages of multiple models to complete more accurate recognition.

2 Method

How to extract acoustic features more accurately is the key to complete the recognition work. Our algorithm structure is roughly divided into the feature extractor and classifier. The feature extractor will firstly extract the MFCC features, then carry out the residual training to get the features with high recognition.

2.1 MFCC Features Extraction

The MFCC features are widely used in audio feature extraction due to the Mel frequency close to the human ear hearing characteristics. The calculation of the MFCC features goes through the following steps:

Step 1: Pre-emphasis. As the first step of audio signal preprocessing, the audio signal is pre-emphasized that the main purpose is to highlight the high frequency part of the audio signal.

Step 2: Subframe. Divide the pre-emphasized audio signal into multiple short audio frames because the audio signal has a large fluctuation range. The size of each frame is generally 20 ms to 40 ms.

Step 3: Windowing. Windowing is performed on each audio frame after framing. Windowing can make the audio signal more continuous. Windowing will cause signal energy loss at both ends of the audio frame. To avoid this situation, the adjacent frames will overlap. The time-domain signal is $x(n)$ and the window function is $w(n)$. The time-domain signal is truncated by the piercing function to obtain the N-point sequence $x(n)$, so the expression can be obtained:

$$x_n = w(n)x(n) \quad (1)$$

Step 4: FFT transform. The audio signal is unstable in the time domain and the analysis is difficult. It is necessary to convert the signal from the time domain to the frequency domain. The main implementation method is discrete Fourier transform, which is faster than Fourier transform. The main representation of FFT transform is shown:

$$G(k) = \sum_{n=0}^{N-1} g(n)e^{-j\frac{2\pi}{N}nk}, 0 \leq k \leq N \quad (2)$$

Where N represents the number of FFT points.

step 5: Modulus square. Get energy spectrum, As shown in formula :

$$\begin{aligned} G(k) &= ae^{-j\theta_k t} + ja \sin \theta_k k = a_k + jb_k \\ E(k) &= (a^2 + b^2)^2 \end{aligned} \quad (3)$$

Step 6: Passing the Mel filter. The filter is composed of multiple triangular band-pass filters, which can perform a certain smooth operation on the spectrum signal, reduce the influence of harmonics and highlight the formant of the original speech. In the actual operation process, in order to further reduce the amount of calculation, a Mel filter bank consisting of 24 triangular band-pass filters is usually defined that the range covered by each triangular window filter is a critical bandwidth of the human ear similar. These similarity improvements can simulate the masking effect of the human ear. The frequency response of the triangular filter $H_m(k)$ is:

$$H_m(k) = \begin{cases} 0, k < f(m-1) \\ \frac{2(k-f(m-1))}{(f(m+1)-f(m-1))(f(m+1)-f(m))}, f(m-1) \leq k \leq f(m) \\ 0, k > f(m+1) \end{cases} \quad (4)$$

Step 7: Take log energy. In the audio processing process, the logarithmic energy of the audio signal reflects the audio energy information, which is a very important feature for the processing of audio data. After the above steps, the MFCC features can be obtained and converted into the Mel spectrum. As shown below:

$$S(m) = \ln\left(\sum_{k=0}^{N-1} E(k)H_m(k)\right) \quad (5)$$

2.2 Residual Network

In recent years, the deep networks have become more and more widely used in computer vision. The network structures have emerged such as Alexnet [18], which have achieved higher accuracy in the classification of visual tasks. However, with the complexity of the network structure, a series of problems such as disappearing gradients and complicated training also have appeared. In response to these problems, the paper [19] proposes a residual network. The residual network introduces a shortcut connection, which can skip one or more layers of the

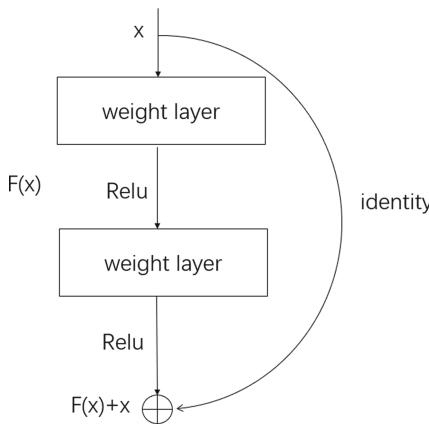


Fig. 1. Residual block

network to solve the problem that the model is difficult to train. The structure of the residual module is shown in Fig. 1.

Assuming that the network input is x and the network output is $H(x)$, the residual is defined as:

$$F(X) = H(X) - X \quad (6)$$

Specific Structure of Residual Network. In this paper, the ResNet50 network is used to modify the fully connected layer to extract the features of the corresponding dimensions. As shown in Fig. 2, the ResNet50 is mainly composed of six stages. The difference from the second to fifth stage is the number of IDBlocks. The sixth stage outputs the flattecenterenging features we need in the next stage.

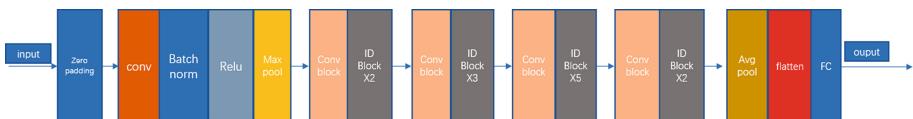


Fig. 2. Resnet-50 structure

2.3 Stacking Algorithm

Stacking algorithm is a common integrated learning method nowadays [20]. By integrating different models, the different models can be combined strongly. Different from other integrated algorithms, the classifiers of the stacking method can be based on different algorithms. The general process of stacking algorithm

is to train the learner of the next layer to combine the output of the learner of the previous layer [21]. Suppose that the model to be stacked is denoted by L and there are N in total, which are $L_1 \dots L_N$ and process the same data set S, where $S_i(x_i, y_i)$, x_i is explanatory variables and y_i is the explained variable. Firstly, the predicted value of base-level is $c_1 \dots c_n$ through upper layer learning, where $s_i = L_i(S)$. Then take the output result of the first layer as the input result of the second layer and put it in the secondary learner for learning.

The paper uses a two-layer stacking model as shown in Fig. 3. The first layer learner uses the LR and RF model, the second layer uses the SVM model and the output of the first layer is input into the SVM model.

Logistic Regression Model. It can predict the probability of a certain situation under the different independent variables. In the training and recognition time of the model, the LR model has great advantages over support vector machine and RF [22].

Random Forest Model. Taking samples by autonomous sampling to construct a decision tree, the final output depends on the output of the constructed decision tree. The RF model is very suitable for solving high-dimensional samples and the problem of overfitting [23].

Support Vector Machine Model. Achieve nonlinear regression by introducing spatial transformation and kernel function. SVM can solve machine learning problems with small samples [24].

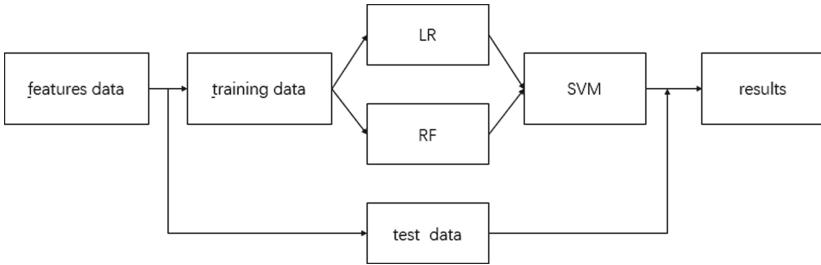


Fig. 3. Stacking model

Because the above three methods have certain shortcomings and have their own areas of expertise, the prediction results generated when they process the same dataset will have relatively independent characteristics, which fully satisfies the principle that the errors are irrelevant and better than random prediction.

3 Experiments

This article will be tested on the ESC-50 and UrbanSound8k dataset. The ESC-50 dataset includes 2000 environmental sound files in the format of wav, divided

into 50 categories. The UrbanSound8k includes 8372 labeled sounds in 10 categories. Relatively speaking, ESC-50 has fewer samples in each category with the high quality. The UrbanSound8k has more samples, but fewer categories.

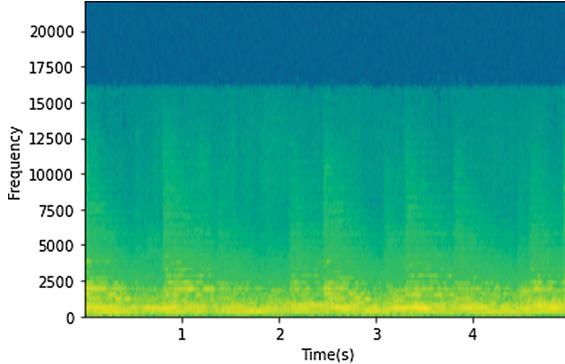


Fig. 4. Mel spectrum

Table 1. Comparison of accuracy of different methods (%)

Model	ESC-50	UrbanSound8k
Piczak-CNN [25]	76.9	76.5
D-CNN [26]	68.1	85.9
Envnet-v2 [27]	84.9	78.3
MelFB-LGTFB-CNN [28]	88.1	85.8
Raw Data-LM [29]	–	92.2
Our Method	91.7	93.6

The specific experimental process is as follows:

Step 1: Extract the MFCC features to get the Mel spectrum. As shown in Fig. 4, the Mel spectrum obtained has a typical periodicity. Different types of sounds in different datasets will show different characteristics. Taking the ESC-50 dataset as an example, one of the major types is indoor sound, which receives less interference and shows regularity. The sound collected in the city is mixed with some noise, which makes it more difficult for our identification work.

Step 2: Put the Mel spectrum into the residual network for training and after reaching the accuracy requirements, extract deep features.

Step 3: Put deep features into the stacking model to get the classification results. The Table 1 compares the accuracy of different methods on the dataset ESC-50 and UrbanSound. It can be seen from the Table 1 that our method exceeds 90% on both types of datasets, reaching 91.7% and 93.6% respectively. The accuracy rate is significantly improving. The recognition rate is also more balanced. The structure we designed is also relatively clear and modular. The future improvements can be made on this base, targeting classifiers and feature extractors.

4 Conclusion

This paper proposes an environmental sound recognition method based on residual network and stacking method. This method uses the residual network structure to extract deep-level features, and uses stacking to fuse multiple models to classify deep-level features better. The combination of these two levels of improvement has made the entire model more excellent. However, the environmental sounds in the actual scene are more complicated, mixing the sounds of many different objects and a lot of meaningless noise. Therefore, the accurate recognition of the complete target may require the support of sound source separation and other technologies to achieve better results. In the future, we will apply this method to other actual scenes, process more different types of sound signals and design a complete system in combination with image processing.

Acknowledgments. The research is supported by National Natural Foundation(NNSF) of China under Grant No. 61973036,61433003.

References

1. Wang, W.: Machine Audition: Principles, Algorithms and Systems (2010)
2. Alsouda, Y., Pllana, S., Kurti, A.: A machine learning driven IoT solution for noise classification in smart cities. arXiv preprint [arXiv:1809.00238](https://arxiv.org/abs/1809.00238) (2018)
3. Niles, J.E.: U.S. Patent Application No. 14/521,866 (2016)
4. Jia, Y.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. IEEE Trans. Control Syst. Technol. **8**(3), 554–569 (2000)
5. Amalia, L., Jesús, G.-B., Alejandro, C., et al.: Optimal representation of Anuran call spectrum in environmental monitoring systems using wireless sensor networks. Sensors **18**(6), 1803 (2018)
6. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. IEEE Trans. Autom. Control **48**(8), 1413–1416 (2003)
7. Stowell, D., Stylianou, Y., Wood, M., et al.: Automatic acoustic detection of birds through deep learning: the first bird audio detection challenge. Methods Ecol. Evol. **10**, 368–380 (2018)
8. Ma, L., Milner, B., Smith, D.: Acoustic environment classification. ACM Trans. Speech Lang. Process. **3**(2), 1–22 (2006)

9. Deng, L., Droppo, J., Acero, A.: Enhancement of log mel power spectra of speech using a phase-sensitive model of the acoustic environment and sequential estimation of the corrupting noise. *IEEE Trans. Speech Audio Process.* **12**(2), 133–143 (2004)
10. Ghoraani, B., Krishnan, S.: Time-frequency matrix feature extraction and classification of environmental audio signals. *IEEE Trans. Audio Speech Lang. Process.* **19**(7), 2197–2209 (2011)
11. Vacher, M., Serignat, J.-F., Chaillol, S.: Sound classification in a smart room environment: an approach using GMM and HMM methods. In: The 4th IEEE Conference on Speech Technology and Human-Computer Dialogue (SpeD), vol. 1, pp. 135–146. Publishing House of the Romanian Academy (Bucharest) (2007)
12. Salamon, J., Jacoby, C., Bello, J.P.: A dataset and taxonomy for urban sound research. In: Proceedings of the 22nd International Conference on Multimedia, Orlando, Florida, pp. 1041–1044 (2014)
13. Zhu, W., O'Shaughnessy, D.: Incorporating frequency masking filtering in a standard MFCC feature extraction algorithm. In: 2004 7th International Conference on Signal Processing, Proceedings, ICSP 2004. IEEE (2004)
14. Scheirer, E., Slaney, M.: Construction and evaluation of a robust multi feature speech/music discriminator. In: IEEE International Conference on Acoustics, pp. 1331–1334. IEEE (1997)
15. Kumar, A., Khadkevich, M., Fügen C.: Knowledge transfer from weakly labeled audio using convolutional neural network for sound events and scenes. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 326–330 (2018)
16. Vu, T.H., Wang, J.-C.: Acoustic scene and event recognition using recurrent neural networks. In: Detection and Classification of Acoustic Scenes and Events (2016)
17. Chi, Z., Ying, L., Cheng, C.: Deep convolutional neural network combined with concatenated spectrogram for environmental sound classification. In: 2019 IEEE 7th International Conference on Computer Science and Network Technology (ICC-SNT). IEEE (2019)
18. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems (2012)
19. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
20. Bohanec, M., Cestnik, B.: A schema for using multiple knowledge. In: The Workshop on Computational Learning Theory and Natural Learning Systems, pp. 157–170. MIT Press (1994)
21. Wolpert, D.H.: Stacked Generalization. Springer, New York (2011)
22. Yao, J., Liu, D.: Logistic regression analysis of risk factors for intracranial infection after multiple traumatic craniotomy and preventive measures. *J. Craniofac. Surg.* **30**(7), 1946–1948 (2019)
23. Baumann, T.: Decision tree usage for incremental parametric speech synthesis. In: IEEE International Conference in Acoustics, Speech and Signal Processing, pp. 3819–3823. IEEE, Italy (2014)
24. Adankon, M.M., Cheriet, M.: Support vector machine. *Comput. Sci.* **1**(4), 1–28 (2002)
25. Piczak, K.J.: Environmental sound classification with convolutional neural networks. In: Proceedings IEEE 25th International Workshop Machine Learning Signal Processing, pp. 1–6 (2015)

26. Kopparapu, S.K., Laxminarayana, M.: Choice of Mel filter bank in computing MFCC of a resampled speech. In: 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010). IEEE (2010)
27. Tokozume, Y., Ushiku, Y., Harada, T.: Learning from between-class examples for deep sound recognition. In: Proceedings of International Conference on Learning Representations, pp. 1–13 (2018). <https://openreview.net/forum?id=B1Gi6LeRZ>
28. Park, H., Yoo, C.D.: CNN-based learnable Gammatone Filterbank and equal-loudness normalization for environmental sound classification. IEEE Signal Process. Lett. **27**, 411–415 (2020)
29. Li, S., Yao, Y., Hu, J., Liu, G., Yao, X., Hu, J.: An ensemble stacked convolutional neural network model for environmental event sound recognition. Appl. Sci. **8**, 1152 (2018)



Robust Online Filter Based on a Second-Order Adaptive Model

Shenglun Yi and Xuemei Ren^(✉)

Beijing Institute of Technology, Beijing 100081, China
3120185460@bit.edu.cn

Abstract. This paper considers robust online filtering problems based on a second-order adaptive model. On the one hand, an adaptive second-order model is applied to characterize the nominal model of the data collected in real-time. On the other hand, we introduce a risk sensitivity parameter to express the modeling error between the nominal model and the actual model. More precisely, the proposed algorithm consists of two parts, one is the state estimation based on the robust Kalman filter, the other is the parameter identification using the Yule-Walker algorithm. Finally, results show the superiority of the proposed robust online filter.

Keywords: Online filtering · Robust Kalman filter · Modeling error · Yule-Walker algorithm

1 Introduction

Any discipline that is concerned with the sensor technologies more likely needs to the a huge number of data. However, data acquisition is always accompanied by noise, due to such reasons as measuring error, sensor fault, or loss of data. Hence, the key point is that how to reduce the influence of the noise and extract the useful information from the data set.

Currently, there are many algorithms focus on the off-line filtering problem. This means a full set of data is necessary before we deal with the data. About the off-line filtering, common solutions can be divided into two ways, one is translating the data set from time domain to frequency domain, the other is dealing with data set in time domain. More precisely, with respect to the off-line filtering in frequency domain, Bhati. D. et al. [1] design time-frequency localized three-band biorthogonal linear phase wavelet filter bank for epileptic seizure electroencephalograph (EEG) signal classification. Then, Kumar. M. K. P. et al. [2] propose a method for single channel speech separation (SCSS) by combining empirical mode decomposition (EMD) and speech specific information. In addition, there are also some off-line filtering which imposes the data to the time domain. For instance, Moon. T. et al. [3] present a new framework of applying deep neural networks (DNN) to devise a universal discrete denoise.

Yang. X. et al. [4] presents a new model based on linear diffusion equation and noise estimation. However, these methods need all the data for the denoising process, and is not suitable for online purposes.

In addition, Kalman filter, as a well-known online filtering, has been widely used in a lot of applications. However, the performance of Kalman filter is highly dependent on the model precision. On the one hand, in practical application of the online filtering, it is difficult to find an accuracy actual model which is equal to the nominal model. Hence, people propose many difficult statistic models, such as CV model and CA model [5]. These two models are suitable to some slightly fluctuating data. Then, in order to dispose the practical application, the Singer model [6] is proposed to deal with the colored process noise. Next, current model (CM) [6] is designed based on the Singer model and this model assumes that the gradient of data change at the next time tick is always within the neighborhood of the current value. Moreover, In order to better describe the dynamic characteristics of the system, Yi. S. L. et al. [7] propose a second-order adaptive statistic model to achieve the processing of the practical data set with colored process noise, and the characteristics of the colored noise were considered in the dynamic model via adaptive parameters.

On the other hand, even if we try to design a nominal model that is very close to the actual model, there is always a modelling error between the nominal model and the actual model. To address this, robust versions based on the standard Kalman filter was designed to guarantee some degree of immunity. Risk-sensitive filter [8] is the most critical one of these robust versions. In this approach, standard quadratic loss function is replaced by an exponential quadratic loss function. This exponential nature can penalize severely large errors. Furthermore, in [9], a ball is constructed by placing the bound on the Kullback-Leibler (KL) divergence between the nominal probability density and the actual probability density and it represents an allowed mismodeling budget. Thus, a time-invariant risk sensitivity parameter over an entire simulation horizon is also introduced and larger parameter are able to penalize larger errors. Then, the Bayes estimator is robust under modeling errors.

In this paper, we apply an adaptive second-order model to characterise the nominal model of the data collected in real time. Moreover, in order to address the modelling error between the nominal model and the actual model, we introduce a risk sensitivity parameter. Then, the effectiveness of method was demonstrated via some experiments, which could have the best performance when the risk sensitivity parameter is set as 0.005.

Finally, the structure of this paper is as follows. Section 2 presents the robust online filtering. The overview of the experiment is provided in Sect. 3. Some conclusions are given in Sect. 4.

2 Robust Online Filtering

We consider a nominal state space model of the form

$$\begin{aligned} x_{t+1} &= A_t x_t + B_t u_t + w_t \\ y_{t+1} &= C_{t+1} x_{t+1} + v_{t+1} \end{aligned} \quad (1)$$

where x_t , u_t and y_{t+1} are the state vector, the input vector and the measurement vector, respectively. A_t , B_t and C_{t+1} are the state process matrix, the input matrix and measurement matrix, respectively. w_t is the process noise with its variance Q_t and v_{t+1} is the measurement noise with the known variance R .

2.1 State Estimation Based on the Robust Kalman Filter

Here, we introduce the second-order adaptive model [7] to characterise the nominal model of the data collected in real time. Let \mathbf{x} and $\dot{\mathbf{x}}$ denote the data itself and the gradient of the data, respectively. Then, the state vector should be $x_t := [\mathbf{x}_t \ \dot{\mathbf{x}}_t]^T$. Next, the input vector is denoted by the mean of the estimated gradient of the data in the time interval of $[0, t]$, i.e.

$$u_t = \frac{1}{t} \sum_{k=0}^t \hat{\mathbf{x}}_k. \quad (2)$$

In addition, let $\hat{x}_{0|0}$, $u_0 = \hat{x}_{0|0}$ and $P_{0|0}$ be the initial value of the estimated state vector, input vector and covariance matrix, respectively. Moreover,

$$A_t = \begin{bmatrix} 1 & \frac{(1-e^{-\alpha_t T})}{e^{-\alpha_t T}} \\ 0 & e^{-\alpha_t T} \end{bmatrix}, \quad B_t = \begin{bmatrix} T - \frac{1-e^{-\alpha_t T}}{e^{-\alpha_t T}} \\ 1 - e^{-\alpha_t T} \end{bmatrix}. \quad (3)$$

Accordingly,

$$Q_t = 2\alpha_t \sigma_t^2 \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix} \quad (4)$$

where

$$\begin{aligned} q_{11} &= \frac{1}{2\alpha_t^3} (4e^{-\alpha_t T} - 3 - e^{-2\alpha_t T} + 2\alpha_t T) \\ q_{12} &= \frac{1}{2\alpha_t^2} (e^{-2\alpha_t T} + 1 - 2\alpha_t T) \\ q_{22} &= \frac{1}{2\alpha_t} (1 - e^{-2\alpha_t T}). \end{aligned} \quad (5)$$

where α_t and σ_t are self-adjusting parameters [11], T is the sample interval. In addition, we assume that C_{t+1} is a time-invariant matrix with $[0 \ 1]$ and R is a known constant value.

Algorithm 1. robust Kalman filter

Require: $y_t, \hat{x}_{t|t}, P_{t|t}$

- 1: $\hat{x}_{t+1|t} = A_t \hat{x}_{t|t} + B_t u_t$
 - 2: $P_{t+1|t} = A_t P_{t|t} A_t^T + Q_t$
 - 3: $\tilde{P}_{t+1|t} = (P_{t+1|t}^{-1} - \lambda_t I)^{-1}$
 - 4: $K_{t+1} = \tilde{P}_{t+1|t} C_{t+1}^T (C_{t+1} \tilde{P}_{t+1|t} C_{t+1}^T + R)^{-1}$
 - 5: $\hat{x}_{t+1|t+1} = \hat{x}_{t+1|t} + K_{t+1} (y_{t+1} - C_{t+1} \hat{x}_{t+1|t})$
 - 6: $P_{t+1|t+1} = (I - K_{t+1} C_{t+1}) P_{t+1|t}$
-

Since we obtain the state space model (1), it is not difficult to see that Kalman filter is commonly used to solve the state estimation problems of this model. However, in the practical application, the actual model is always different from this nominal model. Therefore, a risk sensitivity parameter λ_t [10] is introduced to express the modelling error between the nominal model and the actual model. Note that if the risk sensitivity parameter is equal to 0, the robust Kalman filter is equivalent with the Kalman filter. Then, the implementation is given by Algorithm 1. Here, it is worth noting that this risk sensitivity parameter $\lambda_t \in (0, 1)$ is given by the expert experience. In other words, in different applications we have different empirical value which is calculated based on a large number of experiments. In addition, in our robust Kalman filter, there are two unknown parameters α_t and σ_t in A_t , B_t and Q_t . Accordingly, it is not difficult to see that if we want to know the estimated state $\hat{x}_{t+1|t+1}$, we must get α_t and σ_t at first.

2.2 Parameter Identification Based on the Yule-Walker Algorithm

Then, we tend to find the self-adjusting parameters α_t and σ_t based on the Yule-Walker algorithm. Let α_0 and σ_0 denote the initial value of these two parameters. First, we define that \hat{g}_t is the gradient of the data at time t and it is the second element of $\hat{x}_{t|t}$, i.e. $\hat{x}_{t|t}^2$. Then,

$$\begin{aligned} r_t(1) &= r_{t-1}(1) + \frac{1}{t} [\hat{g}_t \hat{g}_{t-1} - r_{t-1}(1)] \\ r_t(0) &= r_{t-1}(0) + \frac{1}{t} [\hat{g}_t \hat{g}_t - r_{t-1}(0)] \end{aligned} \quad (6)$$

where $r_t(0)$ is the correlation function of the gradient between time t and time $t-1$, and $r_t(1)$ is the autocorrelation function of the gradient at time t . Therefore, the parameters updating is given by [7]

$$\begin{aligned} \alpha_{t+1} &= -\frac{\ln r_t(1) - \ln r_t(0)}{T} \\ \sigma_{t+1}^2 &= \frac{r_t(0) - \alpha_t r_t(1)}{1 - \left(\frac{r_t(1)}{r_t(0)}\right)^2}. \end{aligned} \quad (7)$$

In what follows, the Eq. (7) is applied to A_t , B_t and Q_t . Finally, we get a closed loop of the robust online filter as shown in Fig. 1.

3 Experiment

The configuration of the experiment is shown in Fig. 2. It was a quasi-static test for the column made by Chinese Grade 345 steel and C30 Grade concrete [12]. Through this experiment, deformation displacement at different time samples were obtained, which correspond to the measurements in the proposed algorithm. Then, via the off-line filtering based on the whole data, the reference data is not difficult to collect.

In what follows, we introduce the proposed robust online filter to solve this denoising problem. As shown in the first figure in Fig. 3, it is not difficult to see that the reference is submerged by the serious noise, so it is very necessary to

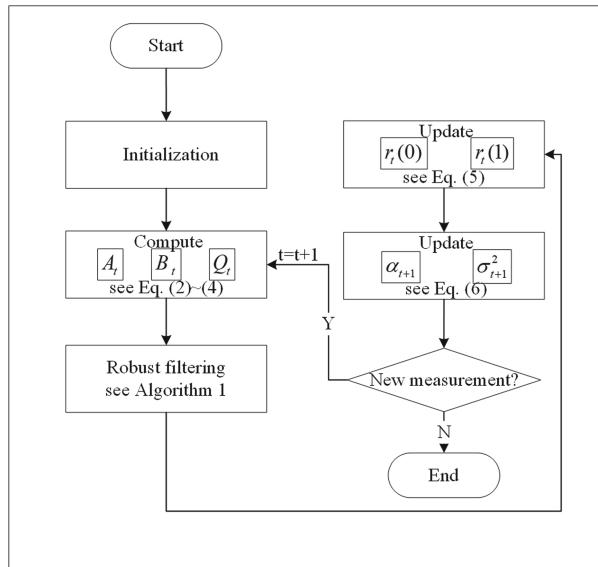


Fig. 1. The flow chart of the proposed robust online filtering algorithm.

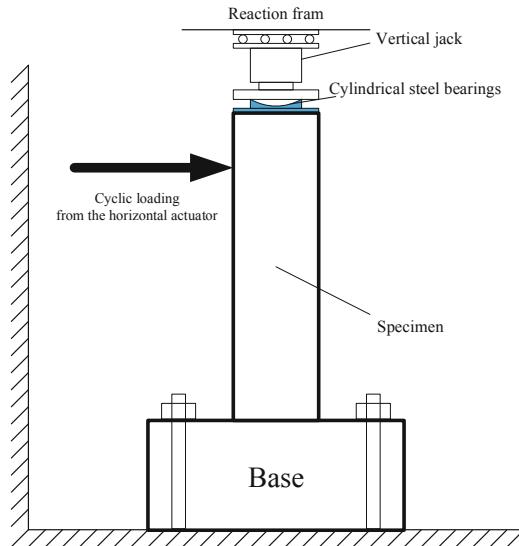


Fig. 2. The configuration of the experiment.

deal with this data set. On the basis of the expert experience, we set the risk sensitively parameter as 0.005. Then, the second figure in Fig. 3 shows a very satisfactory denoising effect. In order to better prove this conclusion, we select another five value of the risk sensitively parameter.

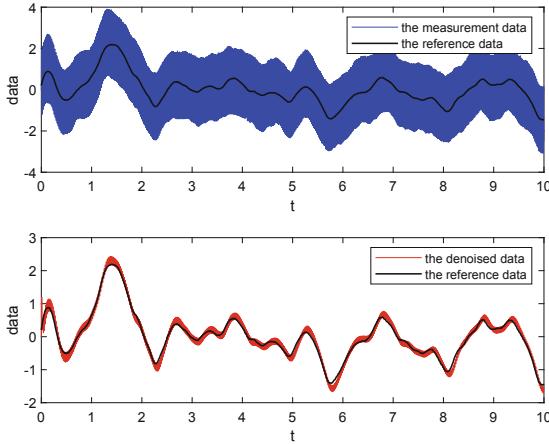


Fig. 3. First figure shows the measurement data and the reference data, and the second figure contains the reference data and the denoised data with the risk sensitively parameter $\lambda_t = 0.005$.

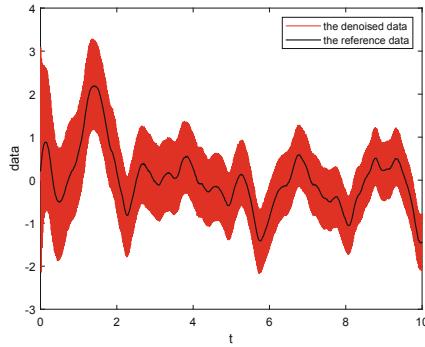


Fig. 4. The reference data and the denoised data with $\lambda_t = 0$.

First, we set the risk sensitively parameter as 0. This means the proposed robust Kalman filter is equivalent with the Kalman filter. Comparing with the robust one with the standard one, it is clear to show in Fig. 4 the performance of the robust online filtering is much better than the latter.

Then, Fig. 5 shows the performance of the reference data and the denoised data with the other different four risk sensitively parameters. In order to evaluate these results, the root-mean-square error (RMSE) is used. RMSE is very commonly used and makes for an excellent general purpose error metric for numerical predictions. More precisely, RMSE can be expressed as the following:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (d_i - r_i)^2}{n}} \quad (8)$$

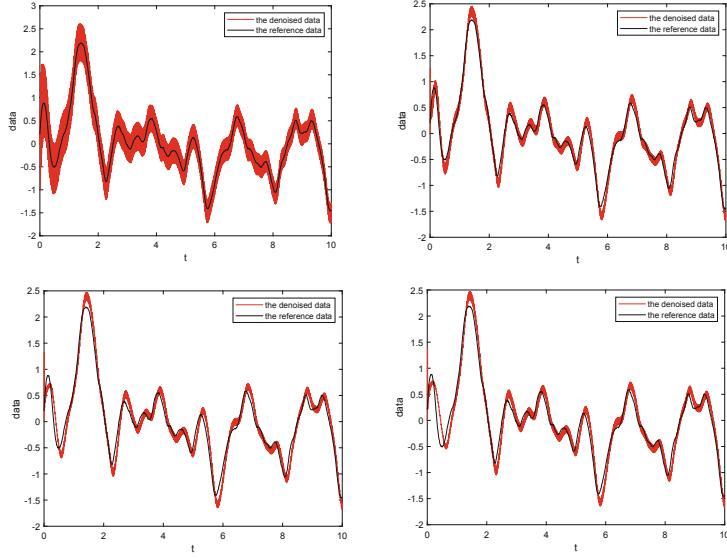


Fig. 5. The reference data and the denoised data with $\lambda_t = 0.001$, $\lambda_t = 0.01$, $\lambda_t = 0.05$, $\lambda_t = 0.1$, respectively.

where n is the number of the measurements, d_i is the i^{th} denoised data and r_i is the corresponding reference data. Hence, Table 1 shows the performance comparison between different risk sensitively parameters. As shown in Table 1, it is not difficult to see that $\lambda_t = 0.005$ is the best choice in this application, and as the increase or the decrease of λ_t , the performance of the proposed algorithm would be worse. However, they all better than the situation of $\lambda_t = 0$. i.e. the standard Kalman filter.

Table 1. Performance comparison between different risk sensitively parameters.

λ_t	0	0.001	0.005	0.01	0.05	0.1
RMSE	0.6049	0.2310	0.1142	0.1275	0.1597	0.1684

4 Conclusion

In this paper, we propose a robust online filter based on the second-order adaptive model. Here, we focus on the practical case that the actual model is different from our nominal model, i.e. second-order adaptive model. Therefore, we introduce the risk sensitively parameter to express this modelling error. However, we only consider the expert experience of this risk sensitively parameter. Finally,

the effectiveness of method was demonstrated via some experiments, which could have the best performance when the risk sensitively parameter is set as 0.005. In the future, we would consider to find the theoretical value of this risk sensitively parameter based on the minimax game [13].

Acknowledgments. This work is partially supported by NSFC under Grant Nos. 61973036 and 61433003.

References

1. Bhati, D., Sharma, M., Pachori, R.B., et al.: Time-frequency localized three-band biorthogonal wavelet filter bank using semidefinite relaxation and nonlinear least squares with epileptic seizure EEG signal classification. *Digit. Signal Process.* **62**, 259–273 (2017)
2. Kumar, M.K.P., Kumaraswamy, R.: Single-channel speech separation using combined EMD and speech-specific information. *Int. J. Speech Technol.* **20**(4), 1037–1047 (2017)
3. Moon, T., Min, S., Lee, B., et al.: Neural universal discrete denoiser. In: *Advances in Neural Information Processing Systems*, pp. 4772–4780 (2016)
4. Yang, X., Zhou, B., Feng, J.: A partial differential equation model based on noise estimation for image denoising. In: *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pp. 1–5. IEEE (2017)
5. Li, X.R., Jilkov, V.P.: Survey of maneuvering target tracking: dynamic models. In: *Signal and Data Processing of Small Targets 2000*, vol. 4048, pp. 212–235. International Society for Optics and Photonics (2000)
6. Xiaolong, C., Yongjie, P., Ye, L., et al.: AUV sensor fault diagnosis based on STF-singer model. *Chin. J. Sci. Instrum.* **31**, 1502–1508 (2010)
7. Yi, S.L., Jin, X.B., Su, T.L., et al.: Online denoising based on the second-order adaptive statistics model. *Sensors* **17**(7), 1668 (2017)
8. Whittle, P., Whittle, P.R.: *Risk-Sensitive Optimal Control*. Wiley, New York (1990)
9. Levy, B.C., Nikoukhah, R.: Robust least-squares estimation with a relative entropy constraint. *IEEE Trans. Inf. Theory* **50**(1), 89–104 (2004)
10. Zorzi, M.: Robust Kalman filtering under model perturbations. *IEEE Trans. Autom. Control* **62**(6), 2902–2907 (2016)
11. Jin, X., Du, J., Jia, B.A.O.: Maneuvering target tracking by adaptive statistics model. *J. China Univ. Posts Telecommun.* **20**(1), 108–114 (2013)
12. Tang, Z., Ma, H., Guo, J., et al.: Experimental research on the propagation of plastic hinge length for multi-scale reinforced concrete columns under cyclic loading. *Earthq. Struct.* **11**(5), 823–840 (2016)
13. Levy, B.C., Nikoukhah, R.: Robust state space filtering under incremental model perturbations subject to a relative entropy tolerance. *IEEE Trans. Autom. Control* **58**(3), 682–695 (2012)



Extended-State-Observer-Based Funnel Control for Motor Servo System with Quantized Input

Yun Cheng and Xuemei Ren^(✉)

School of Automation, Beijing Institute of Technology, Beijing 100081, China
xmren@bit.edu.cn

Abstract. For improving the tracking performance of the general motor servo systems with quantized control input and disturbances, a new quantized funnel control strategy is proposed. A new simple virtual output is introduced to address the high relative degree problem in the applications of funnel control, and the tracking error will be also prescribed in a funnel region. This strategy removes the complex design of funnel controller in systems of relative degree two. In addition to, the unknown system states and disturbances are estimated by an extended state observer, and the quantization error is compensated. Then, the stability of the closed-loop control system is guaranteed by the Lyapunov theory. Some simulations illustrate the effectiveness of the proposed control strategy.

Keywords: Motor servo systems · Tracking control · Quantized control input · Funnel control · Extended state observer

1 Introduction

Motor servo systems (MSS) have been wildly used in various fields, ranging from industry to military. However, nonlinear frictions, disturbances and modeling error existing in the MSS make the control performance of traditional methods worse. In the applications of the MSS, tracking control is an important part for the practical requirements, so many nonlinear control methods have been proposed to improve the tracking performance, for example, sliding-mode control [1], dynamic surface control [2], adaptive neural network control [3], robust control [4], and LMI [5]. But, the above methods cannot deal with the transient performance of the tracking error efficiently.

Moreover, it is difficult to obtain the actual velocity of the motor through the differential of position signal in applications, thus the velocity is an unknown state for controller design. To solve this problem, the extended state observer (ESO) is introduced in this paper. ESO was developed by Han [6], and the key idea of ESO is that the external disturbances and the unknown system dynamics are considered as a new state variable, and then, the new extended state and

system states are estimated by an ESO. The advantage of ESO used in tracking control is not only making the controller robust to modeling error, parameters change and measurement noise but also simplifying the controller design.

The prescribed performance control (PPC) approach based on error transformation can prescribe the convergence of the tracking error, in other words, it can guarantee the transient behavior without losing the steady-state performance. However, considering the potential singularity problem [7] and the complexity of error transformation in PPC, funnel control [8] is a better way to complete the similar tasks. For ensuring the evolution of the systems tracking error within a prescribed performance funnel, reference [7] propose and develop funnel control. Compared with PPC, the simple structure of funnel control makes it more effective in engineering applications. In addition to, the funnel control methods are expanded from relative degree one to relative degree two [9]. But, they cannot guarantee small overshoot and steady-state error, especially in reference [9]. Recently, this method has been used for tracking control in robotic systems [10] and turntable servomechanisms [11].

Considering more and more applications of information technology in modern control engineering systems [12], it is meaningful to design a high-performance funnel control strategy for the MSS with constraints of input quantization. In case of the quantization of the control signal will make the nonlinearity of the MSS stronger and difficult to deal with, a new control strategy needs to be designed to compensate it. Based on the backstepping method, an adaptive quantized controller for the tracking problem of uncertain nonlinear systems is proposed in [13], but the transient behavior of the tracking error is not considered.

In this paper, a new quantized funnel control strategy is proposed to address the tracking performance of the MSS with the nonlinearity, uncertainty and input quantization. This strategy can not only prescribe the evolution of the position tracking error but also compensate the quantization error and the uncertain nonlinearity, and the unknown state of the MSS is also estimated by an ESO.

2 System Description and Preliminaries

The general MSS with input quantization and disturbances is considered as

$$J\ddot{\theta} + b(t)\dot{\theta} + d(\dot{\theta}, t) = Q(u(t)) \quad (1)$$

where J is the inertia, $b(t)$ is the coefficient of viscous friction, θ and $\dot{\theta}$ is position and velocity, $d(\dot{\theta}, t)$ is the unknown disturbances, and $Q(u(t))$ is the output of a quantizer. The new state variables are selected as $[x_1, x_2]^T = [\theta, \dot{\theta}]^T$, then the dynamics of system (1) can be rewritten as

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = \frac{1}{J}(-b_0 x_2 - \tilde{b}(t)x_2 - d(x_2, t) + Q(u(t))) \\ y = x_1 \end{cases} \quad (2)$$

where $b(t) = (b_0 + \tilde{b}(t))$, b_0 is a positive constant, and $\tilde{b}(t)$ is the modeling error.

Assumption 1. The reference signal y_r and its derivative are continuous and bounded, the unknown disturbance $d(\theta, t)$ is also bounded, and the modeling error $\tilde{b}(t)$ is continuous and bounded with a small positive constant.

2.1 Extended State Observer

Considering the unknown environment noise, the angle velocity of the motor servo system cannot be obtained accurately through differentiating the information collected by position sensor, an extended state observer is introduced here to estimate the unknown states of system (2). We define the total disturbances as $x_3(t) = -(\tilde{b}(t)x_2 + d(x_2, t))/J$, so the state variable of system (2) can be noted as $x(t) = [x_1, x_2, x_3]$. Note $h(t)$ as the derivative of $x_3(t)$, then system (2) can be described as

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = \frac{1}{J}(-b_0 x_2 + Q(u(t))) + x_3 \\ \dot{x}_3 = h(t) \end{cases} \quad (3)$$

For the observability matrix of system (2) is full rank, a linear ESO [14] is introduced as

$$\begin{cases} \dot{\hat{x}}_1 = \hat{x}_2 - 3\omega_0(\hat{x}_1 - x_1) \\ \dot{\hat{x}}_2 = \frac{1}{J}(-b_0 \hat{x}_2 + Q(\hat{u}(t))) + \hat{x}_3 - 3\omega_0^2(\hat{x}_1 - x_1) \\ \dot{\hat{x}}_3 = -\omega_0^3(\hat{x}_1 - x_1) \end{cases} \quad (4)$$

where ω_0 is a small constant and represents the bandwidth of the observer error system [14], $\hat{u}(t)$ is continuous control input designed by estimated information and control method proposed in Sect. 3.

2.2 Funnel Control

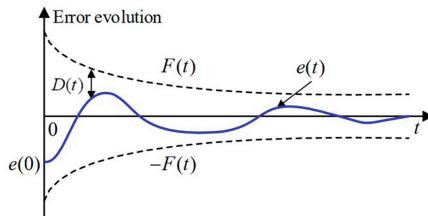


Fig. 1. Performance funnel of the tracking error

Funnel control can prescribe the evolution of the tracking error with the high-gain property of the system. Figure 1 presents the error evolution within a bounded funnel. It shows that when the tracking error approaches the boundary

of the given funnel, the controllers gain will be adjusted to a high value, so the error will not contact the funnel boundary.

Define the performance funnel as,

$$F(t) := \{(t, e) \in R \times R^n | \varphi(t) \cdot \|e(t)\| < 1\} \quad (5)$$

where $e(t) = y(t) - y_r$ is the tracking error. $F(t) = 1/\varphi(t)$ is funnel boundary, where $\varphi(t) > 0, (\forall t > 0)$. One simple form of the funnel controllers is

$$\begin{cases} u(t) = -k_F(t) \cdot e(t) \\ k_F(t) = \frac{1}{D(t)} \end{cases} \quad (6)$$

where $D(t) = F(t) - \|e(t)\|$ is the distance from tracking error to funnel boundary $F(t) > 0$.

Here, we introduce a simple virtual output [15] to solve the high relative degree problem.

$$s(t) = e_1(t) + \mu e_2(t) \quad (7)$$

3 Main Results

In order to guarantee asymptotic performance and transient behavior of the tracking error under input quantization, a funnel control based on virtual output is proposed in this section.

3.1 Controller Design

In this part, we will show a new quantized input feedback controller based on virtual output (7). The structure of the new controller is shown in Fig. 2.

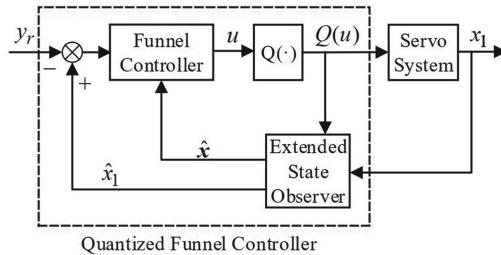


Fig. 2. The structure of quantized funnel controller

Differentiating (7)

$$\begin{aligned} \dot{s}(t) &= \dot{e}_1(t) + \mu \dot{e}_2(t) \\ &= \dot{e}_1(t) + \mu \left(\frac{1}{J} (-b_0 x_2 + Q(u(t))) + x_3 - \ddot{y}_r \right) \end{aligned} \quad (8)$$

Select the candidate Lyapunov function as,

$$V_c = \frac{1}{2}s^2 \quad (9)$$

Combine (8) and the time derivative of (9)

$$\begin{aligned} \dot{V}_c &= s\dot{s} \\ &= s(e_2 + \mu(\frac{1}{J}(-b_0x_2 + Q(u(t))) + x_3 - \ddot{y}_r)) \end{aligned} \quad (10)$$

For ensuring the convergence of the closed-loop system, the funnel controller is designed as

$$u_f = \frac{J}{\mu}(-e_2 - \frac{s}{F(t)-|s|}) + b_0x_2 - J \cdot x_3 + J \cdot \ddot{y}_r \quad (11)$$

In order to compensate the quantization error [13], the controller without ESO is design as follows

$$u_1 = u_f - u_m \tanh(u_m \cdot s/\lambda) \quad (12)$$

where $u_m = \max\{u_0, l\}$ represents the upper bound of quantization error, l is the quantization length, and $\lambda > 0$ is a designed parameter.

According to [13], the uniform quantizer ($\delta = 0$) in this paper has the following property.

$$\begin{aligned} sQ(u_1) &\leq su_1 + u_m|s| \\ &= su_f - (su_m) \tanh(s \cdot u_m/\lambda) + u_m|s| \\ &\leq su_f + M_1 \end{aligned} \quad (13)$$

where $M_1 = 0.2785\lambda$ with the property $|\phi| - \phi \tanh(\phi/\lambda) \leq 0.2785\lambda$ in [16].

Substituting (13) into (10), one can obtain

$$\dot{V}_c \leq s(e_2 + \mu(\frac{1}{J}(-b_0x_2 + u_f) + x_3 - \ddot{y}_r)) + \frac{\mu}{J}M_1 \quad (14)$$

Integrating (14) and (11)

$$\begin{aligned} \dot{V}_c &\leq s(e_2 + (-e_2 - \frac{s}{F(t)-|s|}) + \frac{\mu}{J}M_1) \\ &= -\frac{s^2}{F(t)-|s|} + \frac{\mu}{J}M_1 \end{aligned} \quad (15)$$

Based on the above design and considering the only measurement of position signal, an ESO is introduced to construct the virtual out. The actual funnel controller based on ESO is designed as

$$\hat{u} = \hat{u}_f - u_m \tanh(u_m \cdot \hat{s}/\lambda) \quad (16)$$

where $\hat{s} = \hat{e}_1 + \mu\hat{e}_2(t)$ is constructed by ESO.

3.2 Stability Analysis

Theorem 1. For the servo system described by (3), the upper bound of quantization error $u_m = \max\{u_0, l\}$, the proposed controller (11), (16) based on ESO (4) can guarantee that the evolution of tracking error is uniformly bounded in the prescribed funnel region bounded by $F(t)$.

Proof. Select a Lyapunov function as

$$V = V_c + \frac{1}{2} \varepsilon^T P \varepsilon \quad (17)$$

where $\varepsilon_i = \tilde{x}_i / \omega_0^{i-1}$, $\tilde{x}_i = x_i - \hat{x}_i$, $i = 1, 2, 3$, According to ESO (4), we can get

$$\dot{\varepsilon} = \omega_0 A \varepsilon + M \frac{h(t)}{\omega_0^2} \quad (18)$$

where A is a Hurwitz matrix [14].

As shown in [11], Lemma 1, for bounded $h(t)$, there exists a positive constant σ_i , and a finite constant T so that

$$|\tilde{x}_i| \leq \sigma_i, \sigma_i = o\left(\frac{1}{\omega_0^2}\right), i = 1, 2, 3, \forall t \geq T \quad (19)$$

According to (10), (13), (16) and (18), the time derivative of V can be written as

$$\begin{aligned} \dot{V} &= \dot{V}_c + \frac{1}{2} \varepsilon^T (A^T P + PA) \varepsilon + \varepsilon^T PM \frac{h(t)}{\omega_0^2} \\ &\leq s(e_2 + \mu(\frac{1}{J}(-b_0 x_2 + \hat{u}_f)) + \hat{x}_3 + \tilde{x}_3 - \ddot{y}_r) \\ &\quad + \frac{\mu}{J} M_1 + \frac{1}{2} \varepsilon^T (A^T P + PA) \varepsilon + \varepsilon^T PM \frac{h(t)}{\omega_0^2} \end{aligned} \quad (20)$$

Using Youngs inequality, one can get

$$\begin{aligned} s\tilde{x}_3 &\leq \frac{1}{2}s^2 + \frac{1}{2}\tilde{x}_3^2 \\ \varepsilon^T PM \frac{h(t)}{\omega_0^2} &\leq \frac{1}{2}\varepsilon^T \varepsilon + \frac{1}{2}\frac{\lambda_{\max}(PM)(h(t))^2}{\omega_0^4} \end{aligned} \quad (21)$$

Substituting (11) and (21) into (20)

$$\begin{aligned} \dot{V} &\leq -\frac{s^2}{F(t)-|s|} + \frac{1}{2}\mu s^2 - \omega_0 \varepsilon^T \varepsilon + \frac{1}{2}\varepsilon^T \varepsilon + \frac{\lambda_{\max}(PM)(h(t))^2}{\omega_0^4} \\ &\leq -\left(\frac{1}{F(t)-|s|} + \frac{1}{2}\mu\right)s^2 - (\omega_0 - \frac{1}{2})\|\varepsilon\|^2 + \frac{1}{2}\mu\sigma_3^2 + \frac{\lambda_{\max}(PM)(h(t))^2}{\omega_0^4} \end{aligned} \quad (22)$$

where $P > 0$, and $A^T P + PA = -2I$.

And then, (22) can be described as follows

$$\dot{V} \leq -\eta V + \zeta \quad (23)$$

where $\eta = \min\{2/(F(t)-|s|) - \mu, 2\omega_0 - 1\}$, $\zeta = 0.5\mu\sigma_3^2 + (\lambda_{\max}(PM)(h(t))^2)/\omega_0^4$.

(23) implies that the signals s and ε are bounded, and the tracking error $e_1 = x_1 - y_r = s - \mu e_2$ can be guaranteed within the funnel region bounded by $F(t) + \kappa_s^*$, $\kappa_s^* > 0$, which is proofed in [15], Therorem 1. \square

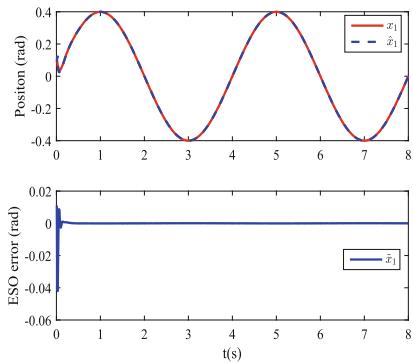


Fig. 3. ESO state \hat{x}_1 and estimation error

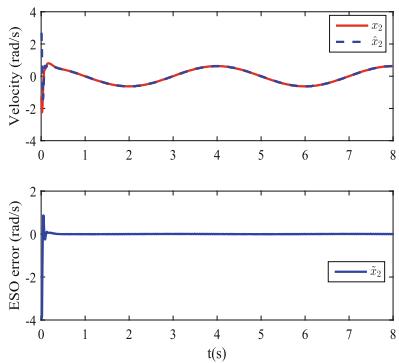


Fig. 4. ESO state \hat{x}_2 and estimation error

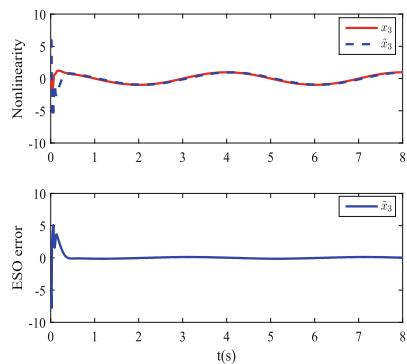


Fig. 5. ESO state \hat{x}_3 and estimation error

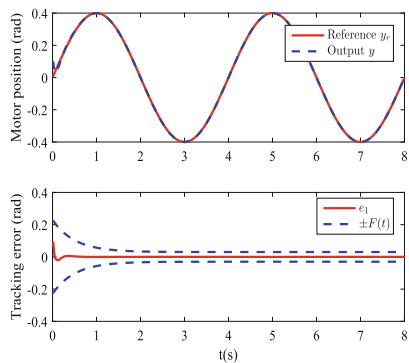


Fig. 6. Performance of the tracking error

4 Simulations

In this section, some simulation results and comparisons were given to show the effectiveness of the proposed control strategy based on high-gain observer and funnel error.

1) Results of funnel control based on virtual output (FC-VO).

The initial state of system (2) is $[x_1, x_2]^T = [0.1, 0.1]^T$. The parameters of system (2) are given as $J = 0.1 \text{ kg/m}^2$, $b = 0.5 \text{ Nm} \cdot \text{s/rad}$, and $x_3 = 0.8(\sin(x_2) + x_2)$. The initial bound of the uniform quantizer is $u_0 = 0.04$, and the quantization interval is $l = 0.1$. The designed parameters of quantized control law are $\lambda = 0.2$ and $\mu = 0.03$. The parameter in ESO (4) is $\omega_0 = 30$, the reference signal is $y_r = 0.4 \sin(2\pi t/4)$, and the virtual output tracking error performance function is $F(t) = 0.2e^{-2t} + 0.03$. Figure 3 and Fig. 4 show the states estimation of system (2). Figure 5 shows the estimation of the extended state x_3 . The evolution of the position tracking error is shown in Fig. 6.

Figure 3 and Fig. 4 show that the states of the system (2) can be observed accurately by ESO (4), and the observer errors are small. In Fig. 5, the extended state x_3 is estimated accurately. Moreover, according to Fig. 6, the position tracking error under the proposed quantized control strategy is not only in the prescribed funnel region but also with small steady-state error.

2) Comparisons of FC-VO and funnel control for arbitrary relative degree (FC-ARD) in [9].

With the same ESO (4), quantization error compensation design (15), and all parameters in 1), the comparison of the two controllers is shown in Fig. 7. Comparing with FC-ARD method, the FC-VO has better control performances in both transient behavior and steady-state performance of the position tracking error.

5 Conclusions

Considering quantization of control input and some nonlinearity in the general motor servo system, a new quantized control strategy based on virtual output and ESO is proposed in this paper. The evolution of the virtual output constructed by an ESO can be prescribed through a funnel function, and then the position tracking error can be guaranteed in a bigger funnel function than virtual output such that the transient behavior and asymptotic convergence of it can be prescribed. In addition to, nonlinear disturbances and modeling error of the system are extended as a new state and estimated by an ESO with saturation. The convergence of the tracking error bounded by funnel function and the stability of the closed-loop system are proved by Lyapunov theory. Some simulations show the effectiveness of the quantized control strategy.

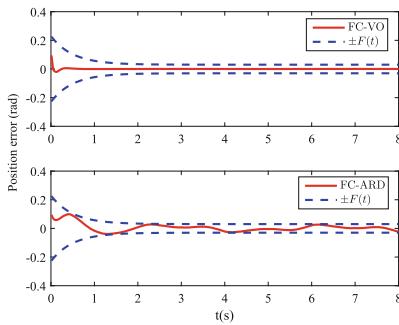


Fig. 7. Performance of the tracking error under FC-VO and FC-ARD

References

1. Baik, I.C., Kim, K.H., Youn, M.J.: Robust nonlinear speed control of PM synchronous motor using boundary layer integral sliding mode control technique. *J. IEEE Trans. Control Syst. Technol.* **8**(1), 47–54 (2000)
2. Wang, M., Ren, X., Chen, Q., et al.: Modified dynamic surface approach with bias torque for multi-motor servomechanism. *J. Control Eng. Pract.* **50**, 57–68 (2016)
3. Ren, X., Lewis, F.L., Zhang, J.: Neural network compensation control for mechanical systems with disturbances. *J. Automatica* **45**(5), 1221–1226 (2009)
4. Jia, Y.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *J. IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2008)
5. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic uncertainty: a predictive approach. *J. IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003)
6. Han, J.: From PID to active disturbance rejection control. *J. IEEE Trans. Ind. Electron.* **56**(3), 900–906 (2009)
7. Han, S.I., Lee, J.M.: Recurrent fuzzy neural network backstepping control for the prescribed output tracking performance of nonlinear dynamic systems. *J. ISA Trans.* **53**(1), 33–43 (2014)
8. Ilchmann, A., Ryan, E.P., Sangwin, C.J.: Tracking with prescribed transient behaviour. *J. ESAIM: Control Optim. Calculus Var.* **7**, 471–493 (2002)
9. Berger, T., Le, H.H., Reis, T.: Funnel control for nonlinear systems with known strict relative degree. *J. Automatica* **87**, 345–357 (2018)
10. Hackl, C.M., Endisch, C., Schroder, D.: Contributions to non-identifier based adaptive control in mechatronics. *J. Robot. Autonom. Syst.* **57**(10), 996–1005 (2009)
11. Wang, S., Ren, X., Na, J., et al.: Extended-state-observer-based funnel control for nonlinear servomechanisms with prescribed tracking performance. *J. IEEE Trans. Autom. Sci. Eng.* **14**(1), 98–108 (2017)
12. Jiang, Z.P., Liu, T.F.: Quantized nonlinear control a survey. *J. Acta Autom. Sinica* **39**(11), 1820–1830 (2013)
13. Zhou, J., Wen, C., Wang, W.: Adaptive control of uncertain nonlinear systems with quantized input signal. *J. Autom.* **95**, 152–162 (2018)

14. Zheng, Q., Linda, Q., Gao, Z.: On stability analysis of active disturbance rejection control for nonlinear time-varying plants with unknown dynamics. In: Proceeding of 46th IEEE Conference on Decision Control, New Orleans, LA, USA, pp. 3501–3506 (2007)
15. Dhrubajit, C., Khalil, H.K.: Funnel control for nonlinear systems with arbitrary relative degree using high-gain observers. *Automatica* **105**, 107–116 (2019)
16. Polycarpou, M.M.: Stable adaptive neural control scheme for nonlinear systems. *J. IEEE Trans. Autom. Control* **41**(3), 447–451 (1996)



Control and Application of Tail-Like Mechanism in Self-balance Robot

Lufeng Zhang, Xuemei Ren^(✉), and Yun Cheng

Beijing Institute of Technology, Beijing 100081, China
zlf2018@163.com, xmren@bit.edu.cn, 18236881947@163.com

Abstract. This paper proposes a self-balance robot with tail-like mechanism and corresponding dynamic model. Unlike the self-balance vehicle which adjust the tilt angle using the wheel motor to control the balance, the balance of the self-balance robot in this paper is realized by controlling the tail-like mechanism. Firstly, an underactuated system model is established and we adopt an adaptive hierarchical sliding mode controller. The adaptive law is derived for the uncertainty estimation. Then, the system stability is proved based on Lyapunov function. Finally, the simulation of the controller for the system is presented to prove the effectiveness.

Keywords: Tail-like mechanism · Underactuated system · Sliding mode control

1 Introduction

The research of bionics is of great significance to robot design. Inspired by animals, researchers have made kinds of robots, such as snake robots in [1, 2], fish robots in [3, 4], cheetah robots [5, 6] and so many others. In this paper, the animal tail is one of the main issues to be discussed. As we known, in nature, the tails of different animals have different functions. Common birds such as chickens, ducks, geese, and most birds use their tails to maintain body balance, fish use their tails to swim, peacocks use their tails to attract the opposite sex, dragonflies use their tails to lay eggs, and scorpions use their tails to attack and prey. Bees use their tails to bite enemies, dogs use their tails to express their emotions, deer use their tails to alert, geckos use their tails to defend themselves, squirrels and foxes use their tails to keep warm in winter, and horses, cows and donkeys use their tails to drive away mosquitoes. We will bring the tail-like mechanism into our self-balance robot design.

In terms of the self-balance vehicle, two wheels are usually driven by two separate motors as shown in [7–9]. The main strategy to keep the two-wheeled vehicle balance, i.e. to keep an inverted pendulum model balance, is to collect the body attitude data by inertial measurement unit (IMU). After the controller board reads the data, a close loop feedback controller can be carried out. As for the controller, PID/PD controller [10], fuzzy controller [11], LQR controller [12], robust control [13], and LMI-based classical method [14] have been tried and

improved. And, the corresponding experiments show effective results. However, we rethink this two-wheeled vehicle design that adjust the tilt angle using the wheel motor to control the balance and find that if the vehicle is expected to track a specific speed, the wheel motor will be under heavy pressure. It means the wheel motor need adjust the two-wheeled vehicle balance and, meanwhile, follow the expected velocity curve. In order to further improve the structure and control of the two-wheeled vehicle, we got inspiration from the animal tails to design a self-balance robot with tail-like mechanism used for uncertainty compensation and balance control.

Actually, tail-like or bionic tail mechanism have been designed in recent years, like the fish robot with 3-DOF tail in [15], the hopping robot Salto in [16] and the bio-robot with primates-inspired active tail in [17]. This paper proposes a self-balance robot with tail-like mechanism and corresponding dynamic model. Unlike the self-balance vehicle which adjust the tilt angle using the wheel motor to control the balance, the balance of the self-balance robot in this paper is realized by controlling the tail-like mechanism. In Sect. 2, the schematic diagram of our self-balance robot with tail is presented and the underactuated system model is established. In Sect. 3, we adopt an adaptive hierarchical sliding mode controller to control the tail and then keep the robot balance. Adaptive law is derived for the uncertainty estimation. Then, the system stability is proved based on Lyapunov function. In Sect. 4, the simulation results are shown to prove the effectiveness.

2 Motivation and Problem Formulation

Figure 1 shows the schematic diagram of the self-balance robot with tail-like mechanism. Here are 3 motors in the new designed self-balance robot, one of which is used to drive the tail to keep the robot body balance and the other two are used to drive the wheels. We assume that the designed tail mechanism is rigid. The tail can only rotate in a plane parallel to the wheel. As shown in Fig. 1, ϕ is the title angle of the robot body relative to the vertical plane and ζ is the rotation angle from the start position (i.e. on the body plane).

Let $q = [\phi \ \zeta]^T \in \mathbb{R}^2$ denote the state vectors. As mentioned in Sect. 1, structural uncertainty should be taken into account when the dynamic equation is established and then the dynamic equation can be derived as follows:

$$M(q)\ddot{q}(t) + V(q(t), \dot{q}(t)) = \tau \quad (1)$$

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \ddot{\phi} \\ \ddot{\zeta} \end{bmatrix} + \begin{bmatrix} V_{11} \\ V_{21} \end{bmatrix} = \begin{bmatrix} \tau \\ \tau \end{bmatrix} \quad (2)$$

where

$$\begin{aligned} M_{11} &= \frac{1}{3}(5M_b + 3m_p)R^2 \\ M_{12} &= M_{21} = m_p R l \cos(\zeta) \\ M_{22} &= m_p l^2 \\ V_{11} &= -m_p R l (\dot{\zeta})^2 \sin(\zeta) + \Xi \\ V_{21} &= m_p g l \sin(\zeta) \end{aligned} \quad (3)$$

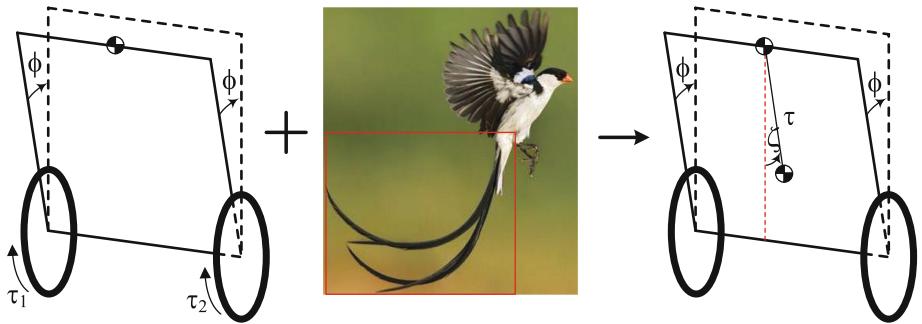


Fig. 1. Schematic diagram of self-balance robot.

And l is the length of the tail, M_b is the mass of the robot body, m_p is the mass of the tail, Ξ stands for the uncertainty, R is the radius of rotation of the body, g is the acceleration of gravity. Let $x_1 = \phi, x_2 = \dot{\phi}, x_3 = \zeta, x_4 = \dot{\zeta}$ represent the state variables of the system, respectively. The state space equation of the self-balance robot can be written as follows:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f_1 + \beta_1 \tau + b_1 \Xi \\ \dot{x}_3 = x_4 \\ \dot{x}_4 = f_2 + \beta_2 \tau + b_2 \Xi \end{cases} \quad (4)$$

where

$$\begin{aligned} f_1 + b_1 \Xi &= \frac{1}{\det M(q)} (M_{12} V_{21} - M_{22} V_{11}) \\ f_2 + b_2 \Xi &= \frac{1}{\det M(q)} (-M_{11} V_{21} + M_{21} V_{11}) \\ \beta_1 &= \frac{1}{\det M(q)} (M_{22} - M_{21}) \\ \beta_2 &= \frac{1}{\det M(q)} (-M_{21} + M_{11}) \end{aligned} \quad (5)$$

It can be seen that the system is an underactuated system. The purpose of the controller in this paper is to control the robot to keep balance. Here, we define the tracking error e as

$$e = \begin{bmatrix} e_1 \\ e_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_{1d} \\ x_3 - x_{3d} \end{bmatrix} \quad (6)$$

3 Controller Design

Balance controller using adaptive hierarchical sliding mode approach is introduced in this section. Based on the assumptions in Sect. 2, the balance controller need compensate the torque Ξ caused by the uncertainty. The first layer of error function is defined as follows:

$$\begin{aligned} S_1 &= c_1 e_1 + \dot{e}_1 \\ S_2 &= c_2 e_3 + \dot{e}_3 \end{aligned} \quad (7)$$

whose the first-order time derivative is

$$\begin{aligned}\dot{S}_1 &= c_1 \dot{e}_1 + \ddot{e}_1 = c_1 \dot{e}_1 + \beta_3 \tau + f_1 + b_1 \hat{\Xi} - \dot{x}_{2d} \\ \dot{S}_2 &= c_2 \dot{e}_3 + \ddot{e}_3 = c_2 \dot{e}_3 + \beta_4 \tau + f_2 + b_2 \hat{\Xi} - \dot{x}_{4d}\end{aligned}\quad (8)$$

where $\hat{\Xi}$ is estimation of the uncertainty, like the eccentric moment. From the moment balance equation $\Xi = m_p g l \sin(\zeta)$, we can get the expectations of system

$$(x_{1d}, x_{2d}, x_{3d}, x_{4d}) = \left(0, 0, \arcsin(\hat{\Xi}/(m_p g l)), 0\right) \quad (9)$$

The equivalent controllers were derived from $\dot{S}_1 = 0$ and $\dot{S}_2 = 0$:

$$\begin{aligned}{}^1\tau_{xeq} &= \frac{1}{\beta_1} \left(\dot{x}_{2d} - c_1 \dot{e}_1 - f_1 - b_1 \hat{\Xi}_x \right) \\ {}^2\tau_{xeq} &= \frac{1}{\beta_2} \left(\dot{x}_{4d} - c_2 \dot{e}_3 - f_2 - b_2 \hat{\Xi}_x \right)\end{aligned}\quad (10)$$

Then, the second layer of error function is obtained by combining S_1 and S_2 as follows:

$$S_x = cS_1 + dS_2 \quad (11)$$

where c and d are selected positive constants. The controller τ also consists of equivalent controllers and switch controller. The switching controller and final input are computed as

$$\tau_{xsw} = -(c\beta_1 + d\beta_2)^{-1} (d\beta_2 {}^1\tau_{xeq} + c\beta_1 {}^2\tau_{xeq} + k_1 \text{sign}(S_x) + k_2 S_x) \quad (12)$$

$$\tau = (c\beta_1 + d\beta_2)^{-1} (c\beta_1 {}^1\tau_{xeq} + d\beta_2 {}^2\tau_{xeq} - k_1 \text{sign}(S_x) - k_2 S_x) \quad (13)$$

The adaptive laws of (11) and (13) are derived as follows, which are based on Lyapunov function $V = \frac{1}{2}S_x^2 + \frac{1}{2\lambda_1} \tilde{\Xi}_x^2$:

$$\dot{\tilde{\Xi}} = -\lambda_1 (cb_1 + db_2) S_x \quad (14)$$

Since $\tilde{\Xi} = \Xi - \hat{\Xi}$ and $\dot{\tilde{\Xi}} = -\dot{\hat{\Xi}}$, we can get the following result after substituting (10) (13) (14) into the Lyapunov function:

$$\begin{aligned}\dot{V} &= S_x \dot{S}_x - \frac{1}{\lambda_1} \tilde{\Xi} \dot{\tilde{\Xi}} \\ &= S_x \left[c \left(c_1 \dot{e}_1 + \beta_1 \tau + f_1 + b_1 \hat{\Xi} - \dot{x}_{2d} \right) + d \left(c_2 \dot{e}_3 + \beta_2 \tau + f_2 + b_2 \hat{\Xi} - \dot{x}_{4d} \right) \right] - \frac{1}{\lambda_1} \tilde{\Xi} \dot{\tilde{\Xi}} \\ &= S_x \left[c \left(c_1 \dot{e}_1 + f_1 + b_1 \hat{\Xi} - \dot{x}_{2d} \right) + d \left(c_2 \dot{e}_3 + f_2 + b_2 \hat{\Xi} - \dot{x}_{4d} \right) + (c\beta_1 + d\beta_2) \tau \right] - \frac{1}{\lambda_1} \tilde{\Xi} \dot{\tilde{\Xi}} \\ &= S_x \left[-k_1 \text{sign}(S_x) - k_2 S_x - (cb_1 + db_2) \tilde{\Xi} \right] - \frac{1}{\lambda_1} \tilde{\Xi} [-\lambda_1 (cb_1 + db_2) S_x] \\ &= -k_1 |S_x| - k_2 S_x^2 \leq 0\end{aligned}\quad (15)$$

So, according to the Lyapunov theorem, the controller (12) (13) and adaptive laws (14) can guarantee the asymptotic stability for the self-balance robot system in this paper.

4 Simulation

In order to verify the proposed method for controlling the self-balance robot with tail-like mechanism, we assumed that the eccentricity $\varepsilon = 47/375$, i.e. $\Xi = 3.6 \text{ N} \cdot \text{mm}$ in Fig. 1, $M_b = 3 \text{ kg}$, $m_p = 8 \text{ kg}$, $l = 94 \text{ mm}$. The parameters of the controllers are $c = 1.45$, $d = 1.6$, $c_1 = 0.01$, $c_2 = 5$, $k_1 = 0.01$, $k_2 = 0.55$. The parameters of the adaptive law are $\lambda_1 = 0.6$. The initial values of the states and estimated torque are $x_i(0) = 0$ ($i = 1, 2, 3, 4$) and $\Xi(0) = 0$.

The simulated results of balance control are shown in Fig. 2. One can find that the robot body loses balance due to structural uncertainty at the beginning. During the unbalanced period, the overshoot of ϕ is 0.777 rad. It can be seen that the unbalanced period lasts nearly 3.5s, and then the states converge to a small set around the desired values, i.e. $[\phi \dot{\phi} \zeta \dot{\zeta}]^T = [0.0149 -4 * 10^{-6} 0.5236 -1.7398 * 10^{-5}]^T$. Figure 3 shows the estimation of Ξ based on adaptive law in Eq. (14). The balance controller τ is presented in Fig. 4, which have good ability for disturbance rejection and robustness.

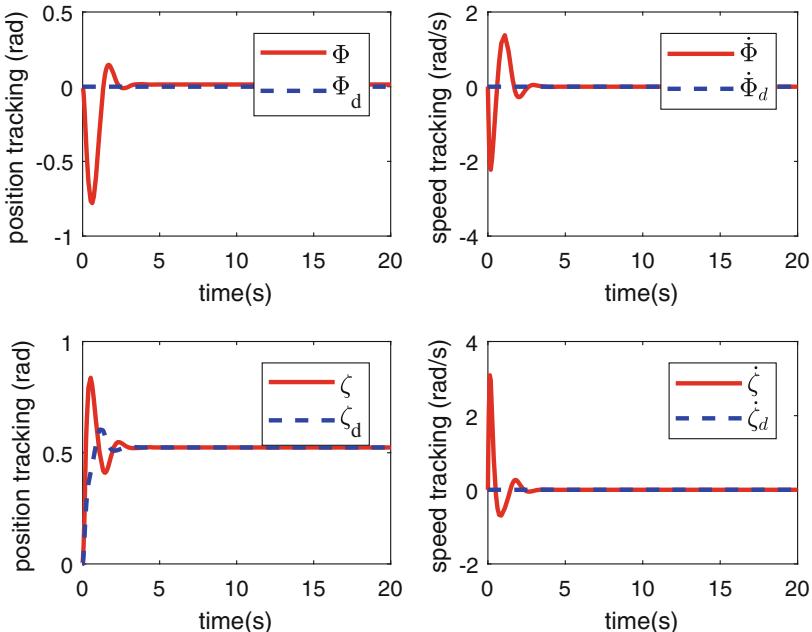


Fig. 2. Simulation results of balance controller for the self-balance robot.

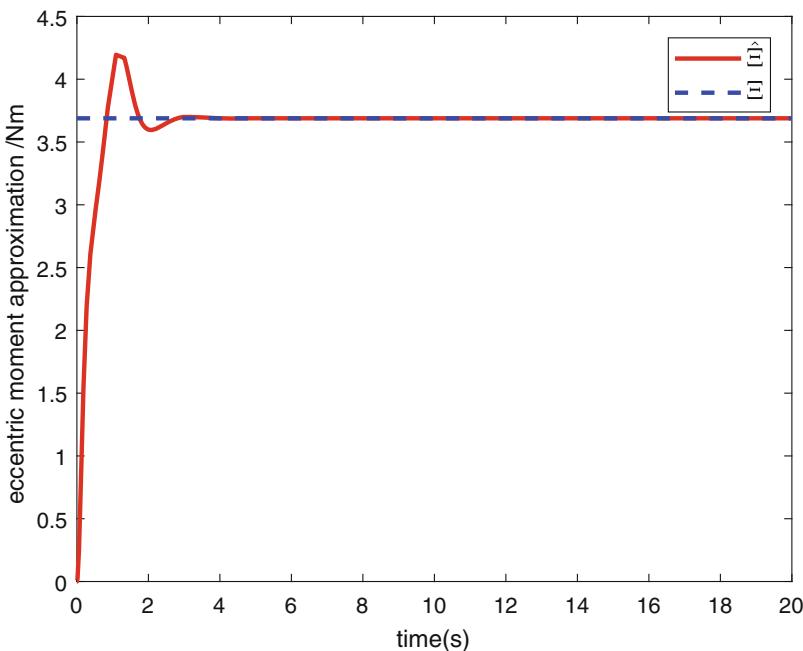


Fig. 3. Uncertainty estimation results.

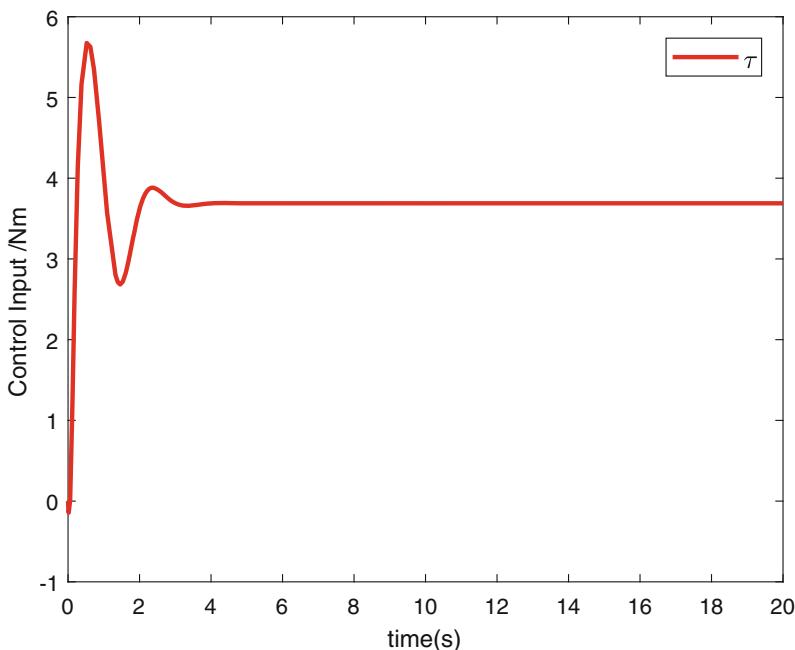


Fig. 4. The proposed control action.

5 Conclusion

This paper proposes a self-balance robot with tail-like mechanism and corresponding dynamic model. The balance of the self-balance robot in this paper is realized by controlling the tail-like mechanism. To control the underactuated system model, we adopt an adaptive hierarchical sliding mode controller. The adaptive law is derived for the uncertainty estimation. Then, the system stability is proved based on Lyapunov function. Finally, the simulation of the controller for the system is presented to prove the effectiveness. And, we will make prototype based on this design concept and finish the experiments.

References

1. Crespi, A., Badertscher, A., Guignard, A., Ijspeert, A.J.: AmphiBot I: an amphibious snake-like robot. *Robot. Autonom. Syst.* **50**(4), 163–175 (2005)
2. Dowling, K.J.: Limbless locomotion: learning to crawl with a snake robot. Ph.D Thesis, The Robotics Institute, Carnegie Mellon University 5000 (1996, Unpublished)
3. Xie, F., Li, Z., Ding, Y., Zhong, Y., Du, R.: An experimental study on the fish body flapping patterns by using a biomimetic robot fish. *IEEE Robot. Autom. Lett.* **5**(1), 64–71 (2020)
4. Zhong, Y., Song, J., Yu, H., Du, R.: Toward a transform method from lighthill fish swimming model to biomimetic robot fish. *IEEE Robot. Autom. Lett.* **3**(3), 2632–2639 (2018)
5. Wang, X., Li, M., Guo, W., Wang, P., Sun, L.: Design and development of a cheetah robot under the neural mechanism controlling the leg's muscles. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2749–2755 (2012)
6. Seok, S., Wang, A., Chuah, M.Y., Otten, D., Lang, J., Kim, S.: Design principles for highly efficient quadrupeds and implementation on the MIT cheetah robot. In: 2013 IEEE International Conference on Robotics and Automation, pp. 3307–3312 (2013)
7. Ciężkowski, M., Pawłuszewicz, E.: Determination of interactions between two-wheeled self-balancing vehicle and its rider. In: 2015 20th International Conference on Methods and Models in Automation and Robotics (MMAR), pp. 851–855 (2015)
8. Sun, H., Zhou, H., Li, X., Wei, Y., Li, X.: Design of two-wheel self-balanced electric vehicle based on mems. In: 2009 4th IEEE International Conference on Nano/Micro Engineered and Molecular Systems, pp. 143–146 (2009)
9. Wang, S., Lu, H., Hou, F.: An improved computing model for a two-wheeled self-balancing vehicle's state determination. In: 2015 8th International Congress on Image and Signal Processing (CISP), pp. 1379–1384 (2015)
10. Abdullah Bin Azhar, M., Hassan, W., Rahim, U.: PID control behavior and sensor filtering for a self balancing personal vehicle. In: 2012 International Conference of Robotics and Artificial Intelligence, pp. 7–10 (2012)
11. Chen, C., Jean, J., Xu, D.: Application of fuzzy control for self-balancing two-wheel vehicle. In: 2011 International Conference on Machine Learning and Cybernetics, vol. 3, pp. 1204–1209 (2011)

12. Meng, X., Li, Y., Liu, X., Xie, F., Wang, Q.: Research on the controller of two-wheeled self-balanced vehicle based on the sensitivity analysis. In: Proceedings of the 10th World Congress on Intelligent Control and Automation, pp. 2833–2837 (2012)
13. Jia, Y.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000)
14. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003)
15. Suebsaiprom, L.: Fish-tail modeling for fish robot. In: International Symposium on Computer (2012)
16. Haldane, D.W., Yim, J.K., Fearing, R.S.: Repetitive extreme-acceleration (14-g) spatial jumping with salto-1P. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2017)
17. Li, X., Jiang, Z., Li, H., Mo, Y., Zou, M., Huang, Q.: Dynamic stability control for a bio-robot with primates-inspired active tail. In: 2015 IEEE International Conference on Mechatronics and Automation (ICMA), pp. 2035–2040 (2015)



Intermittent Fault Detection and Isolation for Discrete System with Unknown Disturbance Based on Interval Observer

Hao Li, Jianfeng Qu^(✉), Xiaoyu Fang, Jinzhuo Liu, and Hongpeng Yin

School of Automation, Chongqing University, Chongqing 400044, China
qujianfeng@cqu.edu.cn

Abstract. In this paper, the problem of intermittent fault diagnosis for linear discrete systems subjected to unknown disturbance is studied. An interval observer for fault diagnosis is designed, residual evaluation and threshold setting are not required, and the upper and the lower bounds of the residual generated by the interval observer can be directly used as adaptive thresholds for the intermittent sensor fault detection. In order to locate the faults, a set of interval isolation observers is designed to isolate sensor faults. Finally, a simulation example is presented to demonstrate the effectiveness of the proposed diagnosis scheme.

Keywords: Intermittent fault diagnosis · Interval observer · Adaptive thresholds

1 Introduction

The rapid development of modern industrial technology has increased the safety requirements for equipment. In most industrial processes, one of the most widely existing faults is the intermittent fault (IF), while most of the current research on fault diagnosis focus on permanent faults [1–3]. The appearing time and disappearing time of IF are uncertain, and the amplitude, frequency and duration of IF have cumulative effect. If an intermittent fault diagnosis (IFD) is not performed, the frequency and the amplitude of the IF of the system will gradually increase over time and will eventually become permanent fault, leading to system failure [4]. Moreover, the IF will lead to the replacement or repair of good components, thus increase the maintenance costs. In this context, the IFD becomes significantly important.

At present, the on-line IFD techniques mainly include data-driven and model-based methods. Model-based IFD method use the system model as prior knowledge to monitor the consistency between the actual process and the behavior predicted by model, and it has been an effective method of detecting faults. In [5–7], the methods based on interval or filter are respectively proposed for IFD, while these methods are applicable to IFs with long duration and long interval. If the duration and intervals of the IFs are too short, the above approach may lead to false alarms and missing alarms.

Recently, the development of interval observer theory and its application in control have provided a new idea for IFD [8,9]. The interval observer can deal with the uncertainty and calculate the upper and the lower bounds of the residuals at each time step for fault detection [10]. Moreover, the threshold selection of IFD is different from PF. Due to the intermittent characteristics of IF, the detection threshold setting is extremely difficult and there is no research result for the selection of IFD threshold. If the interval observer is applied to IFD, the interval computation can avoid the residual evaluation and the threshold selection of the traditional observer-based method, and the estimated output interval can be directly used for fault diagnosis.

The rest of this paper is organized as follows. In Sect. 2, the preliminaries and the problem statement are given. The design of the interval observer and an IFD scheme based on interval observer are described in Sects. 3 and 4, respectively. The simulation results are presented in Sect. 5 to prove the effectiveness of the proposed method. Finally, the conclusions are provided in Sect. 6.

2 Preliminaries and Problem Statement

Notation: In the following statements, the matrices and the vectors are indicated by bold italic while the variables are presented in italic. R^n and $R^{m \times n}$ denote the n -dimensional Euclidean space and the set of $m \times n$ real matrices, respectively. 0 denotes the zero matrix or vector. For two matrices $\mathbf{A}, \mathbf{B} \in R^{m \times n}$, $\mathbf{A} > \mathbf{B}$ indicates that each element in \mathbf{A} is larger than its counterpart in \mathbf{B} . \mathbf{A}^+ stands for $\max\{0, \mathbf{A}\}$, \mathbf{A}^- stands for $\mathbf{A}^+ - \mathbf{A}$ (such that $\mathbf{A}^+ \geq 0$ and $\mathbf{A}^- \geq 0$ are satisfied). Matrix $\mathbf{M} \in R^{m \times n}$ is Schur stable if the modulus of all its eigenvalues is less than one and it is non-negative if all its elements are non-negative.

Lemma 1 [11]: Given a system described by $\mathbf{x}_{k+1} = \mathbf{Ax}_k + \mathbf{u}_k$, the solutions of the system with $\mathbf{x}_k \in R^n$ and a non-negative matrix $\mathbf{A} \in R^{n \times n}$, are non-negative for all $k \geq 0$ provided that $\mathbf{u}_k \geq 0$ and $\mathbf{x}_0 \geq 0$, and the system is called cooperative.

Consider the following discrete time-invariant system subjected to unknown disturbance as:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{Ax}_k + \mathbf{Bu}_k + \mathbf{Dd}_k \\ \mathbf{y}_k &= \mathbf{Cx}_k + \mathbf{F}_i f_k \end{aligned} \quad (1)$$

where $\mathbf{x}_k \in R^{n_x}$ and $\mathbf{y}_k \in R^{n_y}$ are the state and the output vectors, respectively. $\mathbf{u}_k \in R^{n_u}$ is the known input. $\mathbf{d}_k \in R^{n_d}$ denotes the external disturbance or the unknown input. f_k stands for the intermittent sensor fault. $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are constant matrices with appropriate dimensions. $\mathbf{F}_i = (0, \dots, 1, \dots, 0)^T$ denotes i th sensor is in fault case, where the i th element is 1 and the others are 0.

Assumption 1: Assume that the initial state is $\underline{\mathbf{x}}_0 \leq \mathbf{x}_0 \leq \bar{\mathbf{x}}_0$, and there exist the known bounds $\underline{\mathbf{d}}, \bar{\mathbf{d}} \in R^{n_d}$ such that, for all time step $k \geq 0$, $\underline{\mathbf{d}} \leq \mathbf{d}_k \leq \bar{\mathbf{d}}$.

Assumption 2: Assume that the pair (\mathbf{A}, \mathbf{C}) in (1) is detectable.

3 Interval Observer Design

The main idea in this section is to design an interval observer that can estimate the lower and the upper bounds of the state vector \mathbf{x}_k , and the state \mathbf{x}_k is guaranteed to be bounded by the estimates of the interval observer under the consideration of the external disturbance in fault-free case.

The interval observer for system described in (1) is designed as follows:

$$\begin{aligned}\bar{\mathbf{x}}_{k+1} &= (\mathbf{A} - \mathbf{LC})\bar{\mathbf{x}}_k + \mathbf{B}\mathbf{u}_k + \mathbf{Ly}_k + \mathbf{D}^+\bar{\mathbf{d}} - \mathbf{D}^-\underline{\mathbf{d}} \\ \underline{\mathbf{x}}_{k+1} &= (\mathbf{A} - \mathbf{LC})\underline{\mathbf{x}}_k + \mathbf{B}\mathbf{u}_k + \mathbf{Ly}_k + \mathbf{D}^+\underline{\mathbf{d}} - \mathbf{D}^-\bar{\mathbf{d}}\end{aligned}\quad (2)$$

$$\begin{aligned}\bar{\mathbf{y}}_k &= \mathbf{C}^+\bar{\mathbf{x}}_k - \mathbf{C}^-\underline{\mathbf{x}}_k \\ \underline{\mathbf{y}}_k &= \mathbf{C}^+\underline{\mathbf{x}}_k - \mathbf{C}^-\bar{\mathbf{x}}_k\end{aligned}\quad (3)$$

where $\bar{\mathbf{x}}_k \in R^{n_x}$ and $\underline{\mathbf{x}}_k \in R^{n_x}$ are the upper and the lower estimates of state \mathbf{x}_k , respectively. $\bar{\mathbf{y}}_k \in R^{n_y}$ and $\underline{\mathbf{y}}_k \in R^{n_y}$ denote the upper and the lower bounds of output \mathbf{y}_k , respectively.

Theorem 1: Given a system described by (1), if the *Assumption 1* is satisfied, then (2) is an interval observer for system (1), and the state \mathbf{x}_k can be limited by the interval estimation in a fault-free case.

Proof: By defining the estimation error as $\bar{\mathbf{e}}_k = \bar{\mathbf{x}}_k - \mathbf{x}_k$ and $\underline{\mathbf{e}}_k = \mathbf{x}_k - \underline{\mathbf{x}}_k$, using (1) and (2), the following estimation error dynamics can be obtained when no IF occurs:

$$\begin{aligned}\bar{\mathbf{e}}_{k+1} &= (\mathbf{A} - \mathbf{LC})\bar{\mathbf{e}}_k + \mathbf{D}^+(\bar{\mathbf{d}} - \mathbf{d}_k) + \mathbf{D}^-(\mathbf{d}_k - \underline{\mathbf{d}}) \\ \underline{\mathbf{e}}_{k+1} &= (\mathbf{A} - \mathbf{LC})\underline{\mathbf{e}}_k + \mathbf{D}^+(\mathbf{d}_k - \underline{\mathbf{d}}) + \mathbf{D}^-(\bar{\mathbf{d}} - \mathbf{d}_k)\end{aligned}\quad (4)$$

According to *Lemma 1* and *Assumption 1*, $\bar{\mathbf{e}}_k \geq 0$ and $\underline{\mathbf{e}}_k \geq 0$ can easily be obtained if matrix $\mathbf{A} - \mathbf{LC}$ is designed non-negative, such that $\underline{\mathbf{x}}_k \leq \mathbf{x}_k \leq \bar{\mathbf{x}}_k$ holds for all $k \geq 0$ providing that $\underline{\mathbf{x}}_0 \leq \mathbf{x}_0 \leq \bar{\mathbf{x}}_0$. The design task of the interval observer is completed, the next step is to diagnosis the IF.

4 Intermittent Sensor Fault Diagnosis Scheme

Most of the existing observer-based fault detection methods use the measured output and the estimated output provided by the observer to generate the residual, then evaluate the residual and calculate the threshold, and indicate a fault if the residual is greater than the threshold. However, the interval observer-based approach propagates the parameter uncertainty to the residual, and the estimates can be used to detect the fault directly without residual evaluation and threshold calculation.

4.1 Fault Detection

By defining the upper and the lower residuals $\bar{r}_k = \bar{y}_k - y_k$ and $r_k = \underline{y}_k - y_k$, respectively, using (1) and (3), $\bar{e}_k = \bar{x}_k - x_k$ and $e_k = \underline{x}_k - x_k$, the residuals can be obtained as follows in a fault-free scenario:

$$\begin{aligned}\bar{r}_k &= C^+ \bar{e}_k + C^- e_k \\ r_k &= -C^+ e_k - C^- \bar{e}_k\end{aligned}\quad (5)$$

From (5), $\bar{r}_k \geq 0$ and $r_k \leq 0$ for all $k \geq 0$ can easily be obtained based on *Theorem 1*, then $\underline{y}_k \leq y_k \leq \bar{y}_k$ can also be derived. Thus, the output detection and the residual detection conditions for IFD shown in (6) can be obtained:

$$y_k \in [\underline{y}_k, \bar{y}_k] \text{ or } 0 \in [r_k, \bar{r}_k] \quad (6)$$

The faults can be indicated if the above equation is not satisfied. Considering that both output detection and residual detection conditions can detect IFs, the former is selected for convenience.

4.2 Fault Isolation

Fault isolation is one of the central tasks of fault diagnosis in order to find the location of the fault in the process system. If the number of fault is limited and their distribution structure is reasonable, fault isolation can be achieved without complex residual evaluation [12].

Fault isolator design requires that each component of the residual vector corresponds to a sensor fault. Assume that each output component is measured by a sensor, which means that n_y fault isolators need to be designed, and each of the isolators is driven by only one output.

The fault isolator for the i th sensor is designed as follows:

$$\begin{aligned}\bar{x}_{k+1} &= (\mathbf{A} - \mathbf{L}_i \mathbf{C}_i) \bar{x}_k + \mathbf{B} u_k + \mathbf{L}_i y_{i,k} + \mathbf{D}^+ \bar{d} - \mathbf{D}^- d \\ x_{k+1} &= (\mathbf{A} - \mathbf{L}_i \mathbf{C}_i) x_k + \mathbf{B} u_k + \mathbf{L}_i y_{i,k} + \mathbf{D}^+ d - \mathbf{D}^- \bar{d}\end{aligned}\quad (7)$$

$$\begin{aligned}\bar{y}_{i,k} &= C_i^+ \bar{x}_k - C_i^- x_k \\ y_{i,k} &= C_i^+ x_k - C_i^- \bar{x}_k\end{aligned}\quad (8)$$

where $i \in [1, n_y]$, \mathbf{L}_i and \mathbf{C}_i denote the i th column of \mathbf{L} and i th line of \mathbf{C} , respectively. $y_{i,k}$ denote the i th component of the output. C_i^+ and C_i^- denote the i th line of \mathbf{C}^+ and \mathbf{C}^- , respectively.

The residual dynamic equations can be obtained using Eqs. (1), (7) and (8):

$$\begin{aligned}\bar{e}_{k+1} &= (\mathbf{A} - \mathbf{L}_i \mathbf{C}_i) \bar{e}_k + \mathbf{D}^+ (\bar{d} - d_k) + \mathbf{D}^- (d_k - \underline{d}) + \mathbf{L}_i \mathbf{F}_i f_k \\ e_{k+1} &= (\mathbf{A} - \mathbf{L}_i \mathbf{C}_i) e_k + \mathbf{D}^+ (d_k - \underline{d}) + \mathbf{D}^- (\bar{d} - d_k) - \mathbf{L}_i \mathbf{F}_i f_k\end{aligned}\quad (9)$$

$$\begin{aligned}\bar{r}_{i,k} &= C_i^+ \bar{e}_k + C_i^- \underline{e}_k - F_i f_k \\ \underline{r}_{i,k} &= -C_i^+ \underline{e}_k - C_i^- \bar{e}_k - F_i f_k\end{aligned}\quad (10)$$

Obviously, the residual signal of i th isolator is only associated with the i th sensor fault signal. Moreover, interval observer is calculated in each time step, which eliminates the impact of historical fault data. The design of interval isolation observers has been completed, this set of isolators can be used to isolate the sensor fault. So far, the overall design of fault diagnosis has been completed, and the scheme of IFD is shown in Fig. 1.

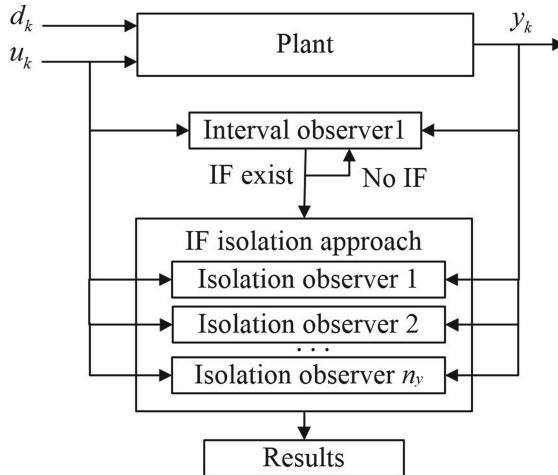


Fig. 1. Architecture of the IFD scheme

5 Simulation

In order to demonstrate the effectiveness of the proposed method in IFD, an example is presented in this section.

$$\boldsymbol{A} = \begin{bmatrix} 0.15 & 0.04 & 0.06 \\ 0.15 & 0.25 & 0.15 \\ 0.13 & 0.17 & 0.63 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 1 \\ -0.5 \\ 0.5 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{D} = \begin{bmatrix} 0.5 \\ 0.5 \\ -0.5 \end{bmatrix}, \boldsymbol{F}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Taking the sampling time $t = 0.01s$ and the system input $u_k = \sin 0.2k + \cos 0.05k$, the disturbance is described as:

$$0 \leq d_k = 0.15 + 0.15 \sin 0.25k \leq 0.3$$

and the bounded initial condition:

$$\begin{pmatrix} -0.1 \\ -0.2 \\ -0.3 \end{pmatrix} \leq \mathbf{x}_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \leq \begin{pmatrix} 0.1 \\ 0.2 \\ 0.3 \end{pmatrix}$$

According to previous discussion, a Schur and non-negative matrix needs to be designed to make the error dynamics cooperative. For convenience, following is selected:

$$\mathbf{A} - \mathbf{L}\mathbf{C} = \begin{bmatrix} 0.6 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$$

and then observer gain matrix \mathbf{L} can be obtained:

$$\mathbf{L} = \begin{bmatrix} -0.45 & 0.04 & 0.47 \\ 0.15 & 0.05 & -0.05 \\ 0.13 & 0.17 & 0.13 \end{bmatrix}$$

Then, the interval observer (2) is designed using the observer gain matrix \mathbf{L} . In order to verify the effectiveness of the proposed detection scheme, a single intermittent sensor fault is added to the system, which is described as:

$$f_k = \begin{cases} 2.4 & 125 \leq k < 135 \\ 2.5 & 145 \leq k < 155 \\ 2.6 & 165 \leq k < 175 \\ 2.7 & 190 \leq k < 197 \\ 2.8 & 210 \leq k < 220 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

By adding IF, the detection results using the proposed scheme are shown in Fig. 2. Obviously, the IF can be well detected, and all appearing times and disappearing times of the IF can be detected as well.

Moreover, the detection results are listed in Table 1, where p_j and q_j are the actual appearing and disappearing time steps of j th IF, respectively. $p_{j,1}$ and $q_{j,1}$ denote the detection time steps using output y_1 , $p_{j,2}$ and $q_{j,2}$ denote the detection time steps using output y_2 , $p_{j,3}$ and $q_{j,3}$ denote the detection time steps using output y_3 , respectively. It can be clearly seen that each appearing (disappearing) time step of the IF can be successfully detected before the next IF disappears (appears) in several time steps. Hence, the proposed IF detection method can effectively detect IFs.

In order to verify the performance of the designed fault isolators, the fault shown in (11) is added to each sensor separately. The isolation results using the proposed scheme are shown in Figs. 3, 4 and 5.

From Figs. 3, 4 and 5, each isolator is sensitive only to the corresponding sensor fault, which indicates that the isolators designed can isolate the sensor fault and achieve the purpose of fault location.

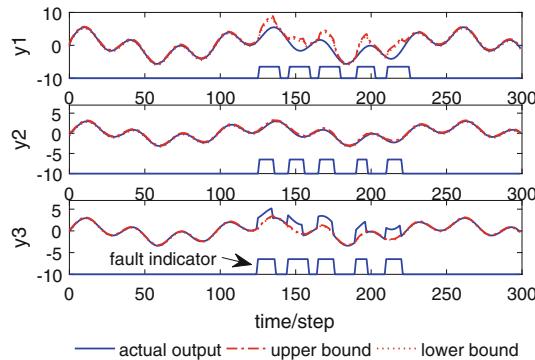


Fig. 2. Detection results using the output detection condition

Table 1. Detection results by using the proposed IF detection scheme

j	p_j	$p_{j,1}$	$p_{j,2}$	$p_{j,3}$	q_j	$q_{j,1}$	$q_{j,2}$	$q_{j,3}$
1	125	127	127	126	135	141	137	136
2	145	147	147	146	155	161	157	156
3	165	167	167	166	175	181	177	176
4	190	192	192	191	197	204	199	198
5	210	212	212	211	220	227	222	221

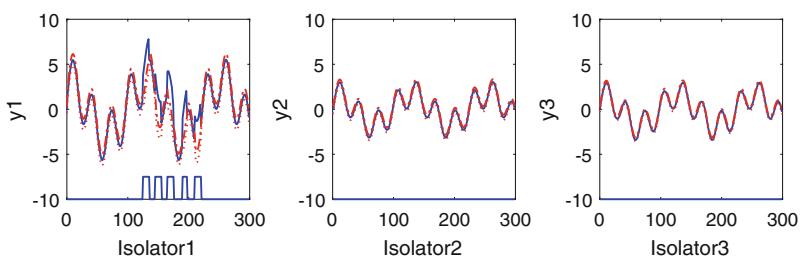


Fig. 3. Isolation results with sensor 1 in fault state

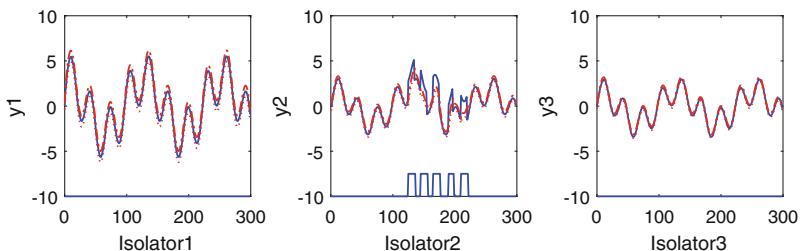


Fig. 4. Isolation results with sensor 2 in fault state

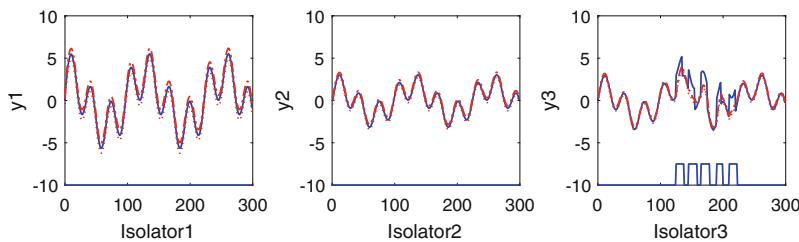


Fig. 5. Isolation results with sensor 3 in fault state

6 Conclusion

In this paper, an intermittent fault diagnosis method based on interval observer for discrete systems with unknown disturbance is proposed. An interval observer is designed to detect the intermittent sensor faults and a set of isolators is designed to isolate the faults. Then an example is provided to verify the effectiveness of the proposed method.

References

1. Zhou, D., Zhao, Y., Wang, Z., et al.: Review on diagnosis techniques for intermittent faults in dynamic systems. *IEEE Trans. Ind. Electron.* **67**(3), 2337–2347 (2019)
2. Syed, W.A., Perinpanayagam, S., Samie, M., et al.: A novel intermittent fault detection algorithm and health monitoring for electronic interconnections. *IEEE Trans. Compon. Packag. Manuf. Technol.* **6**(3), 400–406 (2016)
3. Zhang, J., Christofides, P.D., He, X., et al.: Intermittent sensor fault detection for stochastic LTV systems with parameter uncertainty and limited resolution. *Int. J. Control* **93**(4), 788–796 (2020)
4. Correcher, A., García, E., Morant, F., et al.: Intermittent failure dynamics characterization. *IEEE Trans. Reliab.* **61**(3), 649–658 (2012)
5. Bennett, S.M., Patton, R.J., Daley, S.: Sensor fault-tolerant control of a rail traction drive. *Control Eng. Pract.* **7**(2), 217–225 (1999)
6. Sedighi, T., Foote, P.D., Sydor, P.: Feed-forward observer-based intermittent fault detection. *CIRP J. Manuf. Sci. Technol.* **17**, 10–17 (2017)
7. Huang, J., He, X.: Detection of Intermittent fault for discrete-time systems with output dead-zone: a variant Tobit Kalman filtering approach. *J. Control Sci. Eng.* (2017)
8. He, Z., Xie, W.: Control of non-linear systems based on interval observer design. *IET Control Theory Appl.* **12**(4), 543–548 (2017)
9. Gu, D.K., Liu, L.W., Duan, G.R.: Functional interval observer for the linear systems with disturbances. *IET Control Theory Appl.* **12**(18), 2562–2568 (2018)
10. Zhang, Z., Yang, G.: Interval observer-based fault detection in finite frequency domain for discrete-time fuzzy systems. *Neurocomputing* **310**, 38–45 (2018)
11. Efimov, D., Raïssi, T.: Design of interval observers for uncertain dynamical systems. *Autom. Remote Control* **77**(2), 191–225 (2016)
12. Ding, S.X.: Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms and Tools. Springer, London (2012)



Research on Simulation Technology of Cooperative Guidance for Aircraft Formation

Jiahui Tong^{1,2(✉)}, Jin Liu^{1,2}, Jing Li^{1,2}, and Tianming Zhan³

¹ Science and Technology on Complex System Control and Intelligent Agent Cooperation Laboratory, Beijing, China
tjh80825@126.com

² Beijing Electro-Mechanical Engineering Institute, Beijing, China
³ TAIJI Computer Corporation Limited, Beijing, China

Abstract. This paper analyzed the structure and characteristics of cooperative guidance system of aircraft formation, puts forward a multi-level simulation system architecture of centralized/distributed driving to solve the cooperative guidance for aircraft formation, realized the modeling of aircraft formation members and interaction by using multi-agent method, discussed the simulation methods of data link networking transmission and positioning, space-time registration of simulation system, and so on. The simulation is carried out based on a certain combat scenario, and the results show that the simulation system constructed in this paper can simulate the whole process of cooperative guidance of aircraft formation.

Keywords: Aircraft formation · Cooperative guidance · Multi-agent modeling · Data link simulation · Architecture of simulation system

1 Introduction

According to the requirements of combat mission, the cooperative guidance system of aircraft formation forms a formation with specific capabilities through support network, such as situation awareness, resource sharing, functional complementarity, tactical coordination, etc. through the implementation of formation decision-making and management, it can realize the coordination of fire-power allocation and echelon, task planning and target allocation, cooperative route planning and collaborative guidance control, and cooperative attack, so that the formation can complete combat tasks efficiently [1].

It is an important way to greatly improve the effectiveness of formation cooperative operation with simulation method to simulate and analyze the behavior and capability of multi aircraft formation cooperative guidance system. In this paper, combined with the characteristics of cooperative guidance system of aircraft formation, a centralized/distributed driven real-time simulation method for multi aircraft formation is proposed, which supports the realization, verification and optimization design of formation collaborative architecture design, dynamic ad hoc network, distributed intelligent decision-making and formation efficient control.

2 Overview of Cooperative Guidance System for Aircraft Formation

As shown in Fig. 1, the cooperative guidance system of aircraft formation mainly consists of five parts [1]: situation awareness system, formation decision system, formation flight control system, member flight control system, formation support network, etc.

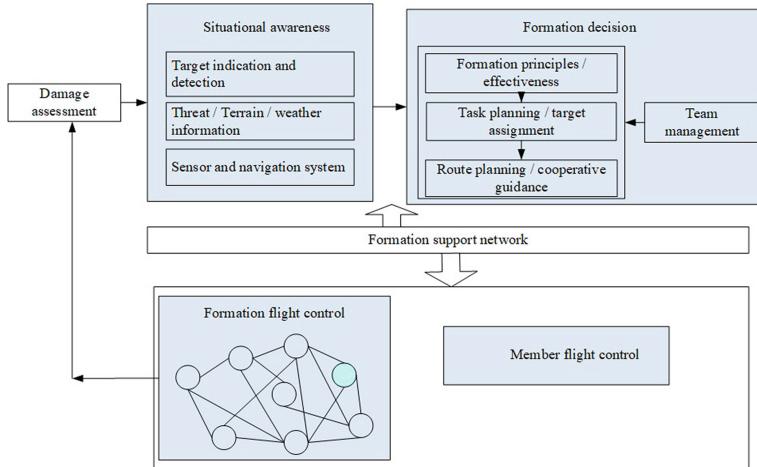


Fig. 1. Architecture of cooperative guidance system for aircraft formation

Situation Awareness System: It is used to obtain the information of each aircraft node, formation flight status and mission environment in the process of formation coordination. It includes the flight status of aircraft members, main sensor status, formation scale, networking status, binding combat tasks, target indication, battlefield environment and threat information, etc.;

Formation Decision System: It makes decision and management of mission planning and target assignment, route planning and cooperative guidance process. It can optimize formation, generate guidance information, conduct departure management, etc., formulate formation principles and effectiveness indicators according to situation awareness information, so as to maximize the comprehensive combat effectiveness of each aircraft member and formation;

Formation Flight Control System: It optimizes and maintains the formation in real time according to the formation decision-making results, such as formation generation, fault-tolerant control technology and so on. It can avoid the behavior conflict between the formation and the collision between the aircraft members in the formation, and ensure the completion of the formation task;

Member Flight Control System: It realizes the controlled flight of aircraft members themselves according to the requirements of formation flight control and mission planning;

Formation Support Network System: It is the key component of aircraft formation. It realizes the information network transmission between the aircraft members and the aircraft formation systems according to the communication protocol, and dynamically adapts to the continuous changes of the aircraft formation scale and network structure.

3 Simulation Method of Cooperative Guidance System for Aircraft Formation

In order to verify the cooperative guidance performance of aircraft formation, the architecture and solution of simulation system should also reflect the mission characteristics and actual performance of aircraft formation guidance system as much as possible [2].

3.1 Analysis of System Hierarchy and Design of Simulation System Architecture

The cooperative guidance system of aircraft formation is composed of a large number of aircraft members and their sensor components through the interaction to form a network, which has strong coupling. The centralized/distributed driven real-time simulation system can consider the global characteristics of the system. In addition, it can accurately assess the performance and operational adaptability of equipment and even components in a small scale. As shown in Fig. 2, it is implemented through a three-tier architecture:

Task Centralized Deployment Layer: It evaluates the complexity according to the input simulation task information, generates and distribute distributed operation deployment scheme according to the task deployment rules, and adjusts the dynamic task in real time according to the information of calculation node;

Resource Dynamic Scheduling Layer: It schedules distributed nodes to execute computing tasks according to the top-level deployment, monitors and collects the computing node information in real time through the node server, obtains the online state information of computing nodes for event scheduling, and synchronously uploads task deployment layer for centralized task deployment;

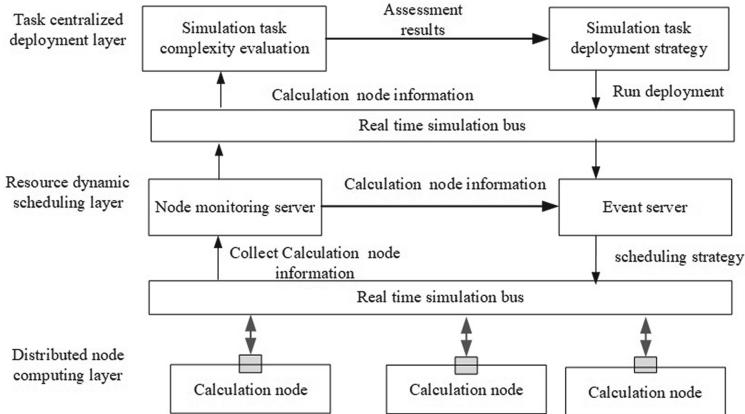


Fig. 2. Architecture of real-time simulation system with centralized/distributed driver

Distributed Node Computing Layer: It consists of several distributed multi-core processors, allocates an exclusive computing kernel for each model, so that the multi model collaborative simulation is independent of each other. It performs node calculation according to the upper scheduling strategy, and feeds back the process data and node status up in real time.

3.2 Multi Agent Modeling of Formation Members and Interaction

Agent is considered to be a physical or abstract entity, which can respond to the environment by collecting external information and thinking independently, and has the ability of knowledge, communication and decision-making planning [3]. In the cooperative guidance system of aircraft formation, an agent can be used to describe an individual member of the formation, and a multi-agent hierarchical structure or a full connection structure can be used to describe the cooperation between members. It can be described in the form of mixed structure, as shown in Fig. 3.

Formation Decision-Making Agent: It obtains target and battlefield situation information, carries out departure and entry management, mission planning and dynamic target assignment for different formation members. By comprehensively considering the system effectiveness of the formation members ($E = a * d * c$), it enables each member to give full play to its characteristics, and maximizes the comprehensive combat effectiveness of the formation. Note: a-availability, d-credibility, c-capability;

Formation Flight Control Agent: It obtains the requirements of formation optimization generated by formation decision agent, and forms formation guidance, control and maintenance instructions in real time, such as basic formation

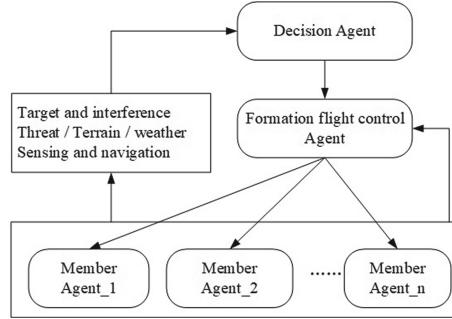


Fig. 3. Multi-agent hybrid structure for cooperative guidance of aircraft formation

of column, horizontal formation, wedge formation and diamond formation, or large-scale formation. The command is output to the member flight control system to keep a certain safe distance ($\Delta x, y, z$) and achieve high-quality formation.

Formation Member Agent(i): It obtains the optimized reference flight trajectory and attitude control commands input by the formation flight control agent, compares them with the actual flight trajectory and attitude information perceived by itself, forms real-time flight control commands, completes the actual flight control of the member through the rudder feedback loop, and outputs the relevant information through the formation support network.

3.3 Network Transmission and Positioning Simulation of Data Link

Data link simulation should not only support the verification of data link networking characteristics and end-to-end wireless link transmission capability, but also meet the verification of dynamic characteristics of data link and self-healing

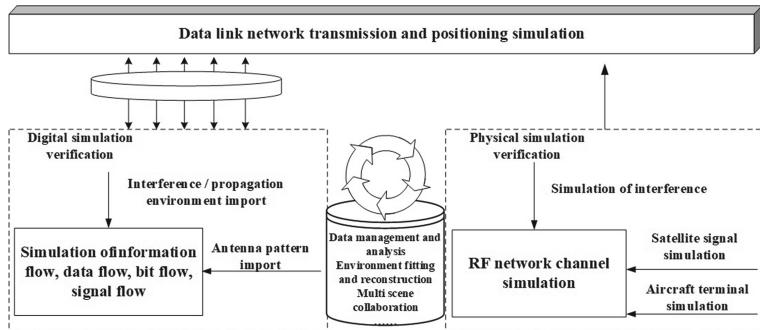


Fig. 4. Composition of simulation and verification system for data link network

capability of failure mode. As shown in Fig. 4, the data link networking transmission and positioning simulation can use mathematical and hardware in the loop joint simulation to achieve the application layer, network layer, link layer and physical layer simulation verification.

Digital Simulation Environment: It simulates the characteristics of data link terminal and data link by establishing mathematical model of data link system:

- Establish simulation nodes for multi aircraft members of simulation system;
- System simulation for data link system;
- Establish a multi-dimensional and multi-functional hierarchical collaborative simulation system;
- Establish simulation of communication environment.

Hardware in the Loop Simulation Environment: It establishes the link channel between aircrafts in the laboratory environment by the way of RF line wired link, and the simulation means of wireless channel and link interference are provided:

- Establish simulation nodes for each aircraft member;
- Simulation of various transmission channels such as satellites and aircrafts;
- Realize the simulation of multi jamming frequency band, pattern and intensity.

3.4 Time and Space Registration of Simulation System [4]

In the process of simulation, data fusion error may be caused when multiple aircrafts, sensors and detectors measure the same object. It may be caused by many reasons, such as the deviation of the sensor itself, the angle and distance deviation measured by each sensor, the position error of the sensor relative to the common coordinate system, the timing error and so on.

Time Registration Method of Simulation System: Time consistency is the key feature to ensure the working state and data interaction of simulation system. There are many task nodes involved in the simulation system of aircraft formation cooperative guidance. Many reasons will lead to serious time inconsistency in the simulation system, such as the inconsistency of clock sources in each node, the difference of simulation time promotion mechanism, the inconsistency of simulation time resolution, the difference of time acquisition or timing mechanism, etc.

In order to ensure that the simulation system can truly reproduce the cooperative guidance process of the aircraft formation, the time deviation estimation and compensation methods can be used to make the operation calculation of the simulation nodes advance orderly on the same time axis aiming at the time registration problem of the simulation system caused by the “system timing” and the inconsistent sampling period of the sensors. The method used is shown in Fig. 5:

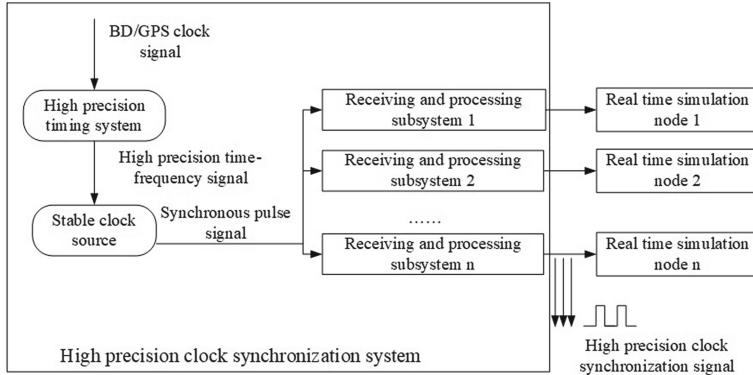


Fig. 5. Time registration method

- The BD or GPS signal received is used to obtain high-precision time-frequency signal;
- The corresponding time-frequency information is obtained by analyzing the stable clock source, and the synchronous pulse signal is output;
- The received synchronous pulse signal is used to compensate the clock signal of the current node by the time system signal receiving and processing subsystem;
- The high-precision clock synchronization signal is provided for the corresponding simulation nodes for data synchronization of real-time simulation system.

Space Registration Method of Simulation System: During the simulation process, some problems that may cause the simulation space inconsistency will seriously affect the accuracy of the sensitive information of the tested object, such as the inconsistent description of the formation member space information, the inconsistent description of the member internal space information, the inconsistent coordinates used in the measurement. There are three ways to solve this problem:

- Accurate modeling of the target with different observation angles;
- Calibration and error compensation of simulation system;
- Information description and measurement in a unified common coordinate system.

4 Design and Implementation of Cooperative Guidance Simulation System for Aircraft Formation

4.1 Operational Scenario

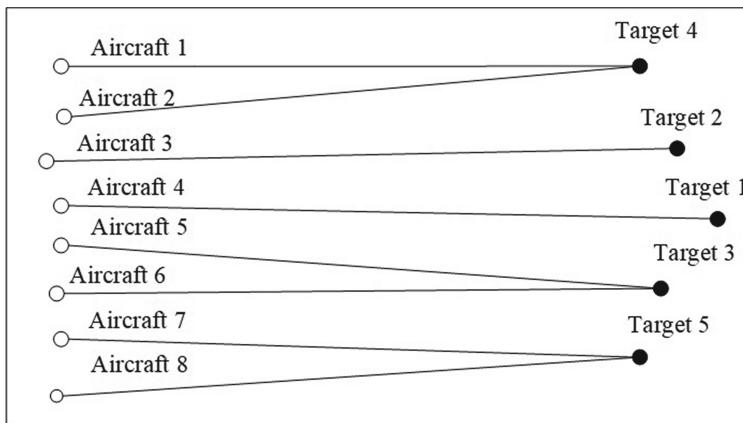
By designing typical combat scenarios, the simulation method of aircraft formation cooperative guidance proposed in this paper is used to verify the formation decision-making ability and the adaptability of cooperative guidance algorithm. See Table 1 for details.

Table 1. Mission requirements for cooperative guidance of aircraft formation

Task type	Description
Aggregation/networking	Enter the planning area, fly in a certain formation
Online planning	Plans new tracks in real time
Target search	The radar starts to search for the target
Task allocation	Assign team members according to combat tasks
Target strike	The formation carries out target attack

4.2 Simulation Process and Result Analysis

According to the scenario of prior simulation, the integrated information management system schedules the simulation operation of each node, and transmits the flight status, battlefield target information and jamming command information to the digital and hardware in the loop simulation environment; the aircraft formation collaborative decision-making system plans the flight route in real time according to the battlefield situation, and guides the formation members to fly according to the real-time planning route; the digital and hardware in the loop simulation node At a certain distance from the target, the radar starts to search, reports the target information, and carries out target strike according to the mission assigned by the command and control system.

**Fig. 6.** Target assignment results of formation attack from cooperative decision system of aircraft formation

At the end of the formation flight, according to the target information reported by the aircraft formation, the command and control system outputs the distribution result of the formation attack target, as shown in Fig. 6. The results of aircraft formation attacking targets are that the aircraft No. 1 and

No. 6 entered the water due to interference; the other aircrafts successfully hit the target and completed the attack task. The simulation results are consistent with the scenario analysis.

5 Conclusion

With the further development of complex simulation system represented by cooperative guidance of multi aircraft formation, various requirements are put forward for the simulation system to realize dynamic perception of mission parameters, flexibility, customization, agile construction, etc.; The number of nodes involved in simulation applications such as formation cooperation/unmanned cluster combat is large and constantly changing, and the network configuration and interaction are complex [5]. It is necessary to study advanced and efficient real-time network communication technology and communication architecture; In addition, the intelligent development of the system and the complex interaction between the system and the environment, based on Hybrid Petri net and multi-agent modeling will play a great role in promoting.

References

1. Sentang, W.: Missile Autonomous Formation Cooperative Guidance Technology. National Defense Industrial Press, Beijing (2015)
2. Zhang A.S.: Research on multi missile cooperative guidance method. Master's thesis of Beijing University of Aeronautics and Astronautics (2014)
3. Zhang, D., Li, W., Shen, S.: Modeling and analysis of mass incident evolution mechanism based on hybrid simulation. *J. Inf. Global Grid* **38**(7), 130–137 (2019)
4. Jin, C., Jin, L., et al.: Foundation and Application of Hardware in the Loop Simulation Technology. China Aerospace Publishing House, Beijing (2020)
5. Jin, W., Xiao, T., Hu, X., Ma, Y.: Distributed aggregated collaborative simulation. *Comput. Simul. Tech.* **22**(5), 120–123 (2005)



On Fatigue Driving Detection System Based on Deep Learning

Yasheng Yuan^{1,3}, Fengzhi Dai^{1,2,3(✉)}, Yunzhong Song⁴, and Jichao Zhao¹

¹ Tianjin University of Science and Technology, Tianjin, China
daifz@tust.edu.cn

² TianjinTiankeIntelligent and Manufacture Technology CO., LTD., Tianjin, China

³ Advanced Structural Integrity International Joint Research Centre,
Tianjin University of Science and Technology, Tianjin 300222, China

⁴ Henan Polytechnic University, Jiaozuo 454003, China
<http://www.tust.edu.cn/>

Abstract. Aiming at the present methods of detecting human fatigue, this paper proposes a new idea of fatigue detection based on deep learning. First of all, YOLOV3-Tiny algorithm is used to detect faces and open mouths in images. Compared with SSD, FCNN and other object detection algorithms, YOLOV3-Tiny has a higher detection and recognition rate for small object, and can also detect targets faster. Then a variant based on LeNet-5 network was used to detection the closed state of the eyes. Compared with the traditional hand-crafted human eye feature descriptor, the deep learning method adopted in this paper can identify the closure of eyes more accurately and has better robustness. Finally, the improved PERCLOS algorithm is used to judge fatigue.

Keywords: Deep learning · Fatigue testing · LeNet-5 network · Face recognition network

1 Introduction

References 1 points out that fatigue driving is the main cause of traffic accidents [1]. After studying a large number of traffic accidents, it is found that drunk driving, overspeed and fatigue driving are the three “killers” of traffic accident. If the three killers of driving behaviors can be well controlled, the traffic accident incidence can be effectively reduced. At present, people in drunk driving and speeding traffic detection has made a series of achievements. Driver fatigue makes the vehicle speed constantly changing, which leads to the vehicle steering insensitively [2]. However, the detection technology for fatigue driving have not been solved, so it is necessary to research and develop the detection technology for fatigue driving.

At present, the detection of fatigue driving is mainly divided into subjective and objective. Subjective detection refers to the evaluation of drivers' driving

behaviors based on their driving habits, driving time [3] and driving tasks, so it can further realize self-detection and self-restraint of driving behaviors [4,5]. Objective detection refers to the detection of drivers' physiological and psychological indicators. This detection mainly tests driver fatigue state [6,7]. However, both of these two methods have the disadvantages of long detection cycle and high detection cost. Therefore, it is of certain significance to use the method combining deep learning and image processing to judge the fatigue driving behavior.

2 Research Method

In this paper, deep learning is used to train and extract human faces from the image. Then extract the eyes and the open mouth from the face image. The open-closed eyes are identified by the improved LeNet-5 network.

2.1 Data Set Processing

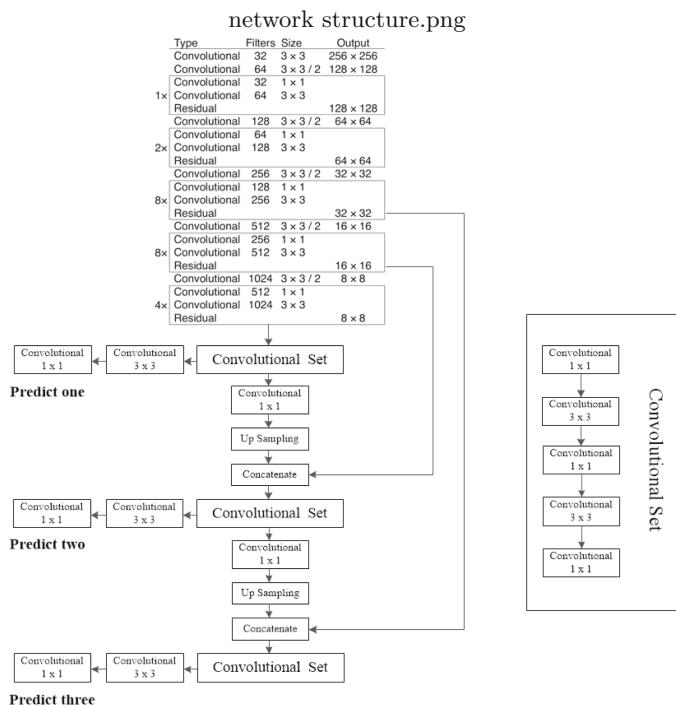
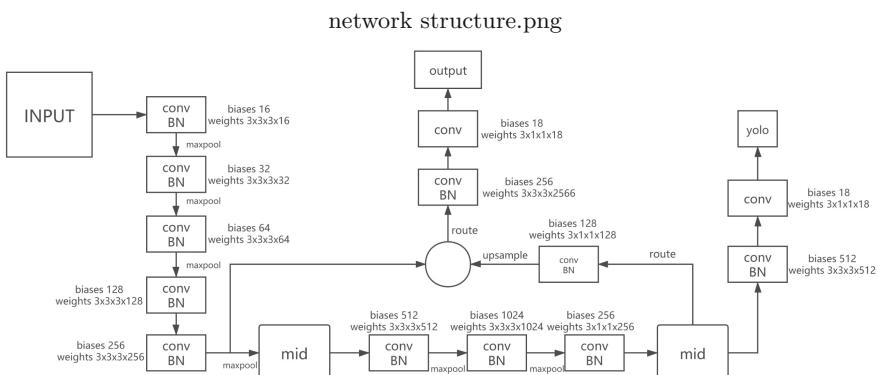
The data set used in this paper is the videos of some volunteers performing actions in the driving car, which is mainly divided into fatigue driving state and non-fatigue driving state. We use Python to decompose the video into pictures and extract every 25 frames which can ensure the recognition frequency and improve the accuracy. After that, marked the face of the decomposed picture and the open mouth, establish the training set and test set.

2.2 Net Selection

The detection algorithm selected in this paper is YOLOV3 [8] algorithm, and the algorithm structure is shown in Fig. 1.

The main reason why adopted the YOLOV3 in this paper is that it can complete all the prediction and detection tasks at one time, which cannot be achieved in the RCNN (Region Convolutional Neural Networks) structure. YOLOV3 uses the residual network structure to form a deeper network level and multi-scale detection to improve the effect of mAP (mean Average Precision) and small objects. In YOLOV3, each box uses multiple label classifications to predict a bounding box might contain which classes. This algorithm does not use Softmax because it has high performance, so YOLOV3 uses separate logical classifier. During the training, we use binary entropy loss to make category prediction. For some overlapping tags, the multi-tag approach can simulate the data better.

This paper chooses the YOLOV3-Tiny algorithm to detect and extract the image of the face and mouth, YOLOV3-Tiny is a simplified version of the YOLOV3. Compared with YOLOV3, the YOLOV3-Tiny version compressed the network a lot, and did not use the res layer, only used two different scales of YOLOV3 output layer, the backbone network using a 7 layer conv + max network to extract feature, structure as shown in Fig. 2. Although simplification of the network structure results in a decrease in the accuracy of YOLOV3-Tiny in detecting multiple objects and larger images, the accuracy of small objects

**Fig. 1.** YOLOv3 network structure**Fig. 2.** YOLOv3-Tiny network structure

is not reduced. At the same time, due to the small amount of calculation, the training speed and detection speed are greatly improved.

First, the training parameters of YOLOV3-Tiny should be modified [9], and the maximum number of iterations is set to 25,000. When the number of iterations reaches 15,000, the learning rate will decay by 10 times. When the number of iterations reaches 20,000, the learning rate will decay 10 times on the basis of the previous learning rate. To increase the training sample, we set the rotation angle to 25, adjust saturation to 1.3, and exposure to 1.6. The learning rate is set at 0.002.

2.3 Facial Image Processing

We use the OpenCV tool to divide the resulting face into left eye region, right eye region and mouth region. Then the object is detected by the eye state detection network. The eye state detection network is adapted from LeNet-5 network.

Although LeNet-5 [10] network is small, it contains the basic modules of deep learning: convolutional layer, pooling layer and full connection layer. LeNet-5 has seven layers, not including input. Each layer contains trainable parameters and multiple feature maps. The feature map extracts an input feature through one convolutional filter. LeNet-5 network structure is shown in Fig. 3.

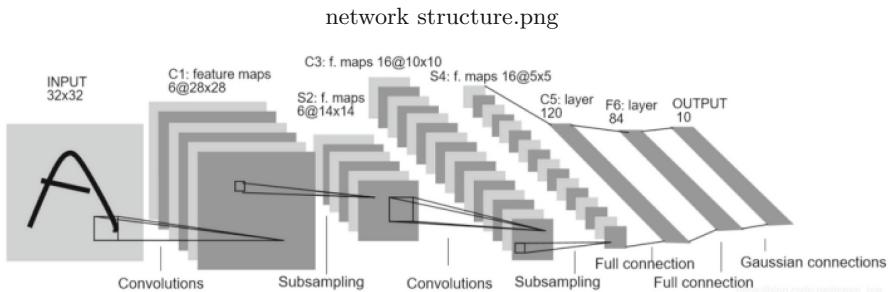


Fig. 3. LeNet-5 network structure

The improved LeNet-5 network structure is shown in Fig. 4.

The size of the network input image is uniformly normalized to 28281. After the first convolution calculation, 64 feature graphs of 2424 were obtained, the size of the convolution kernel is 55, and the slide step size of the convolution layer is 1. Then, through the maximum pooling layer with a sampling area of 22, 64 feature graphs of 1212 were obtained, and the sliding step size of the pooling layer was 2. After the second convolution calculation with the size of the convolution kernel 55, 32 feature graphs of 88 were obtained. The pooling layer adopted the maximum pooling layer of 22. The connection between the first full connection layer and the last convolutional layer results in a 256-dimensional

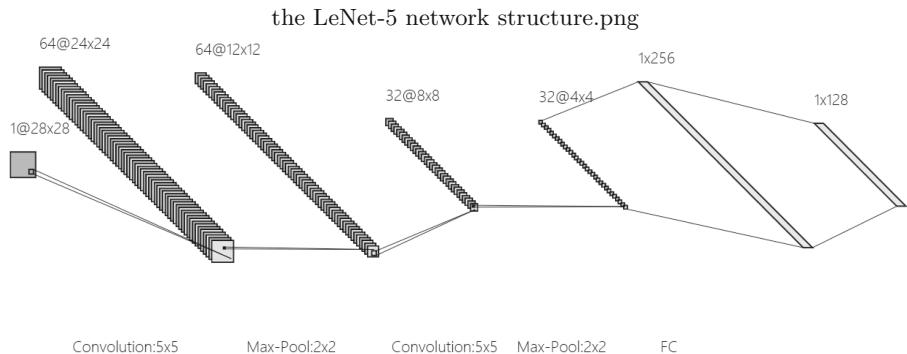


Fig. 4. The improved LeNet-5 network structure

Table 1. Different learning rates in network

Base_lr	State of the eye	Test number	Mistake number	Accuracy	Loss
10^{-3}	Open	2462	1	99.96%	0.00766
	Close	2834	2	99.94%	
10^{-4}	Open	2462	113	95.39%	0.13245
	Close	2834	133	95.30%	

vector. The second full connection layer and the upper connection result in a 128-dimensional vector. The final output is dichotomy (open and close), where open is represented by 1 and close by 0. The excitation function used for the excitation layer is ReLu (Rectified Linear Unit). The accuracy rate and loss value of the test network are shown in Fig. 5 and Fig. 6.

At the same time, this paper also discusses different results under different learning rates, as shown in Table 1.

3 Improved PERCLOS Algorithm

PERCLOS (Percentage of Eyelid Closure Over the Pupil over time) is a set of criteria commonly used in fatigue driving research, closing time of the percentage in the total time is closely related to a person's level of fatigue. The ratio of the length of time in which the human eye is in a closed state and the total length of time can determine whether the person is fatigued or not. The original standard of eye closure used in PERCLOS algorithm is the degree of eye closure. In this paper, due to the clarity of the data set picture, we directly adopted the eye closure frequency as the standard, and also added the variable of mouth opening A in the standard of fatigue. The formula is shown below.

$$P = \frac{\text{Frame Number of Closed Eyes}}{\text{One Period Frames}} + A \quad (1)$$

accuracy.png

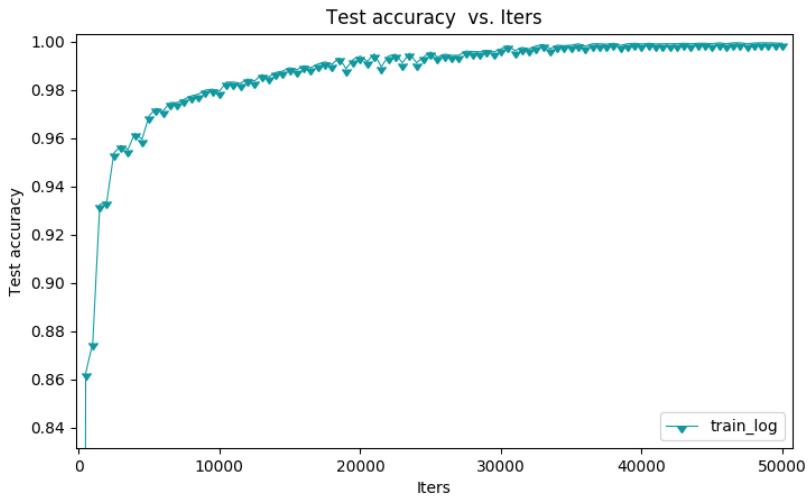


Fig. 5. Test accuracy

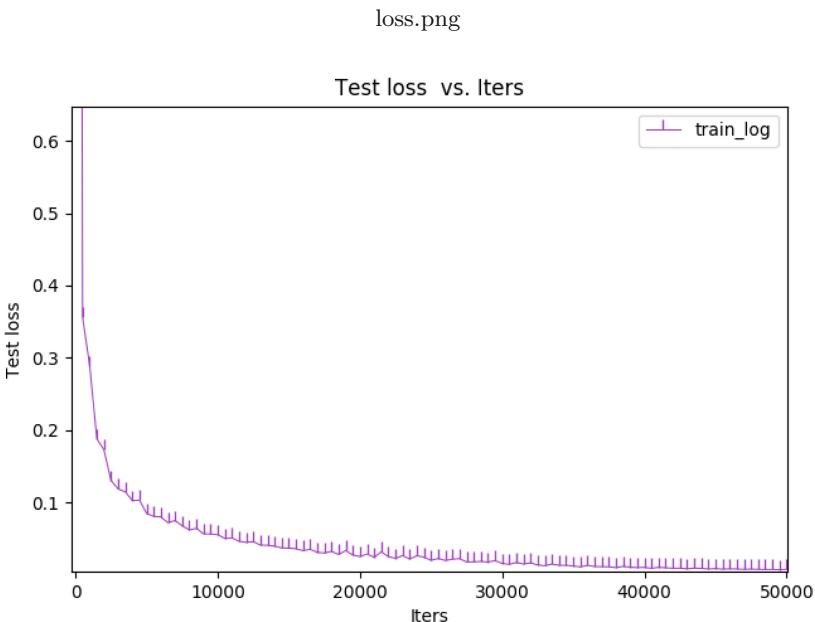


Fig. 6. Test loss

$$A = \log \frac{OnePeriodFrames + 3 * FrameNumberofOpenMouth}{OnePeriodFrames} \quad (2)$$

Table 2. Fatigue detection in the fatigue video

		Fatigue 1	Fatigue 2	Fatigue 3
Part 1	Fatigue	✓	✓	✓
	Awake			
Part 2	Fatigue	✓	✓	✓
	Awake			
Part 3	Fatigue	✓	✓	✓
	Awake			

Table 3. Fatigue detection in the awake video

		Awake 1	Awake 2	Awake 3	Awake 4	Awake 5
Part 1	Fatigue	✓		✓		
	Awake		✓		✓	✓
Part 2	Fatigue		✓	✓		✓
	Awake	✓			✓	
Part 3	Fatigue		✓	✓		
	Awake	✓			✓	✓

We choose 60 pictures for one period. The blink time of people is less than 1s, it is about 0.2 to 0.4 s. Most of the people blinks dozen times in one minute, and we extract the video every 25 frames (average 1s to capture a picture). That is to say, without considering the state of the mouth, the frequency of closing the eyes of person in one cycle should be 0–7 pictures, so we specify the value of P in the fatigue state is [0.5, 1.6], the value of P in the awake state is [0, 0.4].

4 Conclusion

In this paper, we randomly select 3 fatigue videos and 5 non-fatigue videos to test the results, and each video was divided into three parts for testing. When more than two parts in the same video were fatigue, the driver in this video is judged fatigue, and the test results were shown in Table 2 and Table 3.

The chart shows that this method in judging the fatigue state effect is more significant, but the probability of miscalculation is bigger when determine the fatigue state. Because of the video selected for this paper is shot from the side, when the driver looks down or heads to one side, this method misidentifies the open eye as closed eye. If driver have small eyes or the light is dim, it will also misidentify the open eye as the closed eye. Due to this paper use the low-resolution data set, and the PERCLOS algorithm used in this paper is based on the blink frequency rather than the degree of eye open and close, which may be the main reason for a few misjudgments.

In this paper, we use YOLOV3-Tiny and an improved LeNet-5 for fatigue identification. Compared with the traditional hand-crafted human eye feature descriptor, the method proposed in this paper can identify fatigue state more accurately, improve the speed and accuracy, and have better robustness. Compared with traditional model using one depth model, two depth models were used to extract the face and the eyes respectively, so that more fatigue features could be identified. Improved PERCLOS algorithm is a reasonable improvement on the traditional PERCLOS algorithm, joining the mouth impact factor to increase the robustness of the algorithm. The actual effect of the algorithm proposed in this paper is not very good due to the influence of camera resolution and camera position, and this paper still needs to be improved.

References

1. Sussman, D., CoPlen, M.: Fatigue and alertness in the United States railroad industry. *Transp. Res. Part: F* **3**, 211–220 (2000)
2. Jia, Y.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000)
3. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003)
4. Hoddes, E., Dement, W., Zarcone, V.: The development and use of the Stanford sleepiness scale. *Psychophysiology* **9**, 150 (1972)
5. ZhiqiangLiu, Y.W.: Driving fatigue detection method based on machine vision. *China Manuf. Inf.* **03**, 63–66 (2006)
6. Fan, X., Sun, Y., Yin, B., et al.: Gabor-based dynamic representation for human fatigue monitoring in facial image sequences. *Pattern Recogn. Lett.* **31**(3), 234–243 (2010)
7. Murphy-Chutorian, E., Trivedi, M.M.: Head pose estimation in computer vision: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(4), 607–626 (2009)
8. Yi, Z., Yongliang, S., Jun, Z.: An improved tiny-yolov3 pedestrian detection algorithm. *Optik-Int. J. Light Electron Opt.* **183**, 17–23 (2019)
9. Yasheng, Y., Fengzhi, D., Lingran, A., et al.: Research on fatigue detection method based on deep learning. In: Proceedings of 2020 International Conference on Artificial Life and Robotics, Oita, Japan, pp. 640–643 (2020)
10. LeCun, Y., Bottou, L., Bengio, Y., et al.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)



On Improvement of Target Recognition eCCA Method Based on SSVEP

Di Yin^{1,3}, Fengzhi Dai^{1,2,3(✉)}, Yunzhong Song⁴, Jichao Zhao¹,
and Dejin Wang¹

¹ Tianjin University of Science and Technology, Tianjin, China
daifz@tust.edu.cn

² Tianjin Tianke Intelligent and Manufacture Technology CO., LTD., Tianjin, China
³ Advanced Structural Integrity International Joint Research Centre, Tianjin

University of Science and Technology,
Tianjin 300222, China

⁴ Henan Polytechnic University, Jiaozuo 454003, China
<http://www.tust.edu.cn/>

Abstract. In recent years, the brain-computer interface (BCI) has made great progress, especially in steady-state visual evoked potentials (SSVEP) because of its high recognition rate and information transfer rate (ITR), which has received extensive research attention. This paper proposes a feature recognition based on extended canonical correlation analysis (eCCA) method, and adds the phase characteristic of canonical correlation analysis sine and cosine signal to electroencephalogram(EEG) signal for reference. Compared with task-related component analysis and eCCA methods, research shows that this method can strengthen the property of BCI based on SSVEP.

Keywords: EEG · Feature extraction · Spectrum analysis · eCCA

1 Introduction

SSVEP is the brains response to periodic visual blinking stimuli. It is widely used in BCI system research. Compared with other paradigms (motor imaging, P300) in BCI systems, it is characterized by no training or less training, high signal-to-noise ratio (SNR) and high information transmission rate [1–3, 6]. SSVEP detects the selected frequency of the subject by identifying the specific SSVEP component in the subjects EEG signal.

Earlier studies mainly analyzed the frequency domain characteristics of EEG signals. Commonly used frequency identification algorithms include single channel power spectral density analysis (PSDA) [4] and multi-channel frequency identification algorithms. However, because the SNR of the single-channel EEG signal is relatively low, the recognition performance is significantly different from that of the multi-channel signal. Therefore, full use of multi-channel EEG data

for recognition algorithms has high accuracy and robustness. For example, canonical correlation analysis (CCA) between SSVEP and sine and cosine signals can improve the SNR of SSVEP, thereby improving classification accuracy [5–8]. Therefore, the SSVEP system also needs to consider the fact that the input, output and reference signal of the actual system are in the same dimension [9]. CCA also uses the projection principle to reduce the dimension and maximize the correlation coefficient. The system is like a vehicle control problem. The selection and optimization of feature extraction parameters affect its steering performance. For this reason, the proposed feature recognition algorithm should make SSVEP-BCI robust and stable [10].

In the past few decades, multi-channel EEG signals obtained weight vectors through projection and linear combination to obtain spatial filters. Zhang et.al proposed multi-channel CCA [11]. Wong et al. proposed eCCA [12], Masaki et al. introduced the task-related component analysis (TRCA) method to SSVEP identification [8]. We found that there is no EEG signal involved in the sine and cosine reference signal. On the basis of eCCA, we tried to add the reference signal to some features of the EEG signal, and obtained this method extended canonical correlation analysis-Y (eCCA-Y). By comparing with the eCCA and TRCA methods, the recognition rate and ITR have been improved to a certain extent.

2 SSVEP Principle

When a person looks at a stimulus block that periodically flashes at a certain frequency, an EEG response containing the same fundamental frequency component and its higher harmonic component can be recorded on his scalp. The process involves data collection, feature extraction, target recognition, and command output corresponding to the controlled device. As shown in Fig. 1.

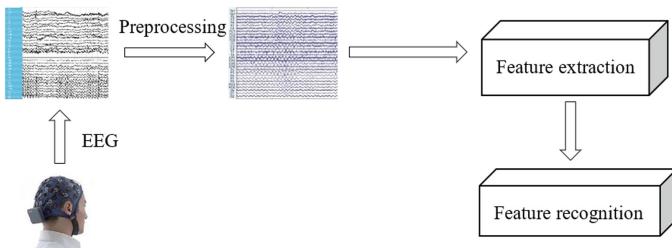


Fig. 1. System block diagram[13]

3 Experiment

3.1 Experimental Data Description

This dataset consists of 35 healthy subjects with an average age of 22, including 17 women. The main stimulus paradigm is 40 stimulus target characters of dif-

ferent frequencies (the interval between two stimuli is 0.2 Hz, from 8 to 15.8 Hz). For each subject, there are six blocks of experiments, each of which contains 40 target recognition. The subjects gaze at 40 visual stimuli (40 stimulus frequencies) in turn, and each stimulus lasts for 6 s. In order to prevent visual fatigue, each block can rest for 2–3 after the completion of the experiment.

3.2 Data Preprocessing

Sixty-four channels covered the whole scalp of the subject and were aligned according to the international 10–20 system. Here nine channels Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2 are used for feature extraction. Because of the size of the data set, the data set is down-sampled to 250 Hz, and then IIR filter (6–90 Hz) is used for band-pass filtering. Since the SSVEP stimulation is a transient component within 0.14 s, the data is intercepted for 1.5 s after 0.14 s.

4 Feature Extraction

4.1 Task-Related Component Analysis (TRCA)

TRCA is a method of maximizing the covariance between tasks to optimize the weight coefficients and extract the relevant components of the task [14,15]. A multi-channel EEG signal $x(t) \in R^{N_c}$ is composed of task-related signals and task-independent signals. We hope that only task related components can be obtained from the EEG signal, so that the SNR of the signal can be increased and the EEG signal can be weighted and summed:

$$y(t) = \sum_{j=1}^{N_c} w_j x_j(t) = W^T X(t) \quad (1)$$

Where j is the index of the channels, $y(t) \in R$ is the optimized signal, $x(t) \in R^N$ is the EEG signal after multi-channel preprocessing, and W is the optimized spatial filter. The EEG signal of let k -th trial and the weighted and optimized task-related component are $x^{(k)}(t)$ and $y^{(k)}(t)$, similarly, the signal from L -th are $x^{(l)}(t)$ and $y^{(l)}(t)$, and the correlation function between them:

$$\hat{c}_{k,l} = Cov(y^{(k)}(t), y^{(l)}(t)) = \sum_{i,j=1}^N w_i w_j Cov(x_i^{(k)}(t), x_j^{(l)}(t)) \quad (2)$$

The sum of covariance among multiple trials is the objective function, that is:

$$\sum_{k,l=1, k \neq l}^K \hat{c}_{k,l} = \sum_{k,l=1, k \neq l}^K Cov(y^{(k)}(t), y^{(l)}(t)) = W^T S W \quad (3)$$

Constrain the coefficients and use normalization processing to make $y(t)$ to 1, that is:

$$\text{Var}(y(t)) = \sum_{i,j=1}^N w_i w_j \text{Cov}(x_i^{(t)}, x_j^{(t)}) = W^T Q W = 1 \quad (4)$$

Therefore, the constraint problem is transformed into the eigenvalue problem of Rayleigh-Ritz, that is:

$$\hat{W} = \arg \max \frac{W^T S W}{W^T Q W} \quad (5)$$

The best coefficient vector \hat{W} can be obtained by solving the eigenvector of $Q^{-1}S$. Finally, the correlation coefficient can be obtained.

4.2 Extended Canonical Correlation Analysis (eCCA)

The application of CCA in SSVEP has achieved good results. On this basis, Chen et al. [2] proposed an extension method based on CCA, the input signal obtained by the average training data of the subjects: $\hat{X}_k^{N_c * N_s}$. When the subjects train the data, we can get three kinds of multichannel signals, test data $X(t) \in R^{N_c * N_s * N_t}$, Average of training data $\hat{X}^{N_c * N_s}$, and the constructed sincos reference signal Y_{fk} , Where $N_c N_s N_t$ represents channels, points and test set trial respectively. Any two signals can be used to calculate the spatial filter according to CCA.

Wong et al. [12], through permutation and combination, 6 kinds of spatial filter forms are obtained, and then 10 typical variables will be generated, and then correlation coefficient can be obtained between each two typical variables, so there are 45 correlation coefficients in total. According to the correlation coefficient formula obtained by the two people, Chen et al. were chosen as the reference, and only three filter forms were selected: (1) $W_X(X\hat{X}_k), W_{\hat{X}_k}(X\hat{X}_k)$ test data and average training data; (2) $W_X(XY_{fk})$ test data and reference signals; (3) $W_{\hat{X}_k}(\hat{X}_k Y_{fk})$ average training data and reference signals.

We can get the combination of six correlation coefficients. For the consideration of the final accurate recognition rate and ITR, we choose the combination of four correlation coefficients to explain, namely: eCCA formula (6). As shown in Fig. 2.

$$r_k = \begin{bmatrix} r_k(1) \\ r_k(2) \\ r_k(3) \\ r_k(4) \end{bmatrix} = \begin{bmatrix} \rho(X^T W_X(XY_k), Y_k^T W_Y(XY_k)) \\ \rho(X^T W_X(X\hat{X}_k), \hat{X}_k^T W_X(X\hat{X}_k)) \\ \rho(X^T W_X(XY_k), \hat{X}_k^T W_X(XY_k)) \\ \rho(X^T W_{\hat{X}_k}(\hat{X}_k Y_k), \hat{X}_k^T W_{\hat{X}_k}(\hat{X}_k Y_k)) \end{bmatrix} \quad (6)$$

If only the first item is considered, the algorithm becomes the standard CCA. If the correlation coefficient is fused, the correlation coefficient under k-th stimulation can be obtained. Finally, the classification can be completed by confirming the maximum correlation coefficient for the corresponding stimulation, that is (7):

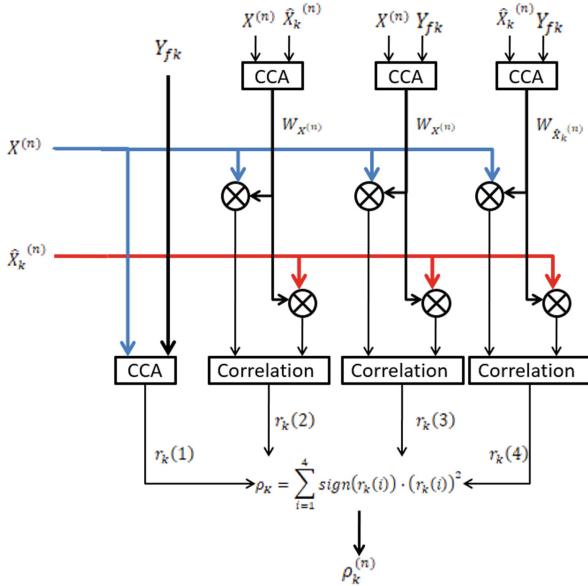


Fig. 2. Diagrams of the proposed methods

$$\rho_K = \sum_{i=1}^4 sign(r_k(i)) * (r_k(i))^2 \quad (7)$$

In the process of cosine reference signal, only the parameters we construct are independent of the EEG signal, so we try to add the phase characteristics of the EEG signal into the reference signal. Y_{f_k, θ_k} is the reference signal containing phase information θ_k for the k -th frequency stimulation (If $\theta_k = 0$, it is the reference signal of CCA):

$$Y_{f_k, \theta_k} = \begin{bmatrix} \sin(2\pi f_k t + \theta_k) \\ \cos(2\pi f_k t + \theta_k) \\ \vdots \\ \sin(2\pi N_h f_k t + N_h \theta_k) \\ \cos(2\pi N_h f_k t + N_h \theta_k) \end{bmatrix} \quad (8)$$

40 targets are encoded by different frequency phases. The frequency phases of each target can be calculated by formula (9):

$$\theta_k = \theta_0 + \Delta\theta * [(k_y - 1) * 5 + (k_x - 1)] \quad (9)$$

Among them k_x and k_y represent the rows and columns of the experimental paradigm matrix respectively, respectively, namely $k_x = 1, 2, \dots, 5$ and $k_y = 1, 2, \dots, 8$. The initial phase (θ_0) is 0, and phase interval ($\Delta\theta$) is 0.5π .

5 Classification Results

At present, TRCA and eCCA are the mainstream algorithms in SSVEP recognition algorithm, so we will compare the accuracy and ITR between the new algorithm(eCCA-Y) and the two algorithms. As shown in Figs. 3 and 4.

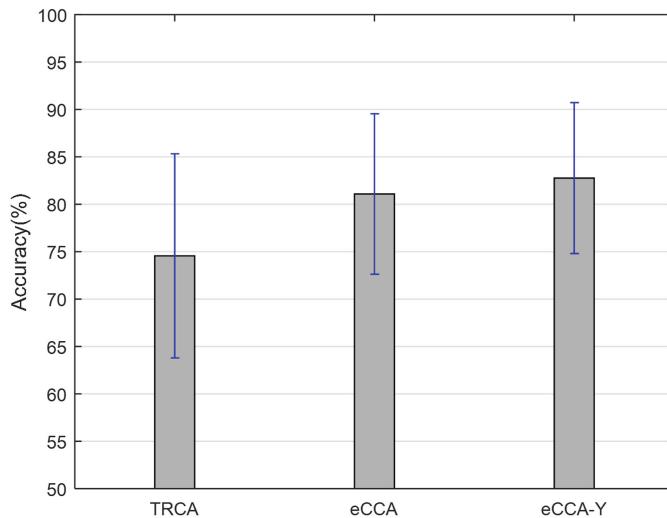


Fig. 3. Average accuracy across of 35 subjects

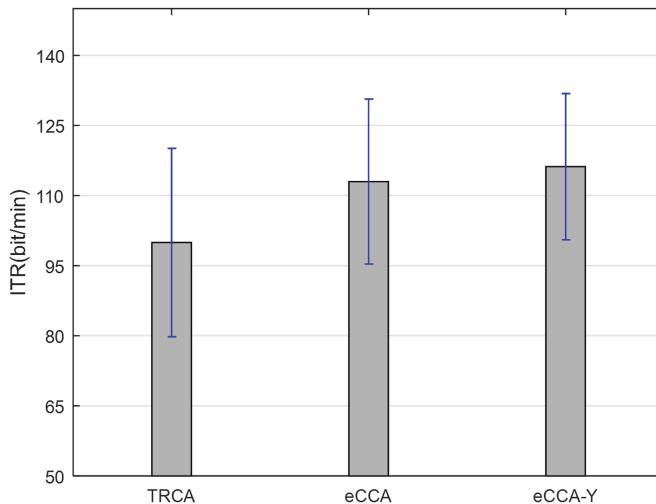


Fig. 4. Average ITR across of 35 subjects

The average error of the correct recognition rate and its results of 35 subjects were analyzed. It was found that the recognition rate of eCCA was significantly higher than that of TRCA by 11%. The method proposed in this paper (eCCA-Y) is slightly higher than eCCA, because it is improved on the basis of eCCA, its recognition rate is 2.1% higher than eCCA on average, and its ITR is also increased by 2.8%.

6 Conclusion

When three different methods are used for recognition, it is found that the recognition rate and ITR of different subjects will be greatly different, because the response degree of each person's EEG signal to the stimulation frequency is different. At the same time, compared with the eCCA method, although the method is slightly improved, there is not a relatively obvious progress, because the spatial filter has not changed, if the filter is added, the recognition rate will be greatly reduced, so far there is no relatively suitable filter combination to improve the recognition rate. The change of the weight coefficient of the correlation coefficient is also a part of the influencing factors, which will be further studied in the future.

References

1. Gao, S., Wang, Y., Gao, X., et al.: Visual and auditory brain-computer interfaces. *IEEE Trans. Biomed. Eng.* **61**(5), 1436–1447 (2014)
2. Chen, X., Wang, Y., Nakanishi, M., et al.: High-speed spelling with a noninvasive brain-computer interface. *Proc. Natl. Acad. Sci.* **112**(44), E6058–E6067 (2015)
3. Wang, Y., Gao, X., Hong, B., et al.: Brain-computer interfaces based on visual evoked potentials. *IEEE Eng. Med. Biol. Mag.* **27**(5), 64–71 (2008)
4. Hakvoort, G., Reuderink, B., Obbink, M.: Comparison of PSDA and CCA detection methods in a SSVEP-based BCI-system. *J. Centre Telemat. Inf. Technol. Univ. Twente TR-CTIT-11-03*, 1–12 (2011)
5. Bin, G., Gao, X., Yan, Z., et al.: An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method. *J. Neural Eng.* **6**(4), 046002 (2009)
6. Chen, X., Chen, Z., Gao, S., et al.: A high-ITR SSVEP-based BCI speller. *J. Brain-Comput. Interfaces* **1**(3–4), 181–191 (2014)
7. Li, Y., Bin, G., Gao, X., Hong, B., Gao, S.: Analysis of phase coding SSVEP based on canonical correlation analysis (CCA). In: 2011 5th International IEEE/EMBS Conference on Neural Engineering, Cancun, pp. 368–371 (2011). <https://doi.org/10.1109/NER.2011.5910563>
8. Nakanishi, M., Wang, Y., Chen, X., et al.: Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis. *IEEE Trans. Biomed. Eng.* **65**(1), 104–112 (2017)
9. Jia, Y.: Alternative proofs for improved LMI representations for the analysis and the design of continuous-time systems with polytopic type uncertainty: a predictive approach. *IEEE Trans. Autom. Control* **48**(8), 1413–1416 (2003)

10. Jia, Y.: Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity-varying motion. *IEEE Trans. Control Syst. Technol.* **8**(3), 554–569 (2000)
11. Zhang, Y., et al.: Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIs. In: International Conference on Neural Information Processing, pp. 287–295. Springer, Heidelberg (2011)
12. Wong, C.M., Wan, F., Wang, B., et al.: Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs. *J. Neural Eng.* **17**, 016026 (2020)
13. Yin,D., Dai, F.Z., et al.: Research on emotion classification based on EEG. In: 2020 International Conference on Artificial Life and Robotics, Oita, Japan, vol. 1, pp. 656–659 (2020)
14. Tanaka, H., Katura, T., Sato, H.: Task-related component analysis for functional neuroimaging and application to near-infrared spectroscopy data. *NeuroImage* **64**, 308–327 (2013)
15. Tanaka, H., Katura, T., Sato, H.: Task-related oxygenation and cerebral blood volume changes estimated from NIRS signals in motor and cognitive tasks. *NeuroImage* **94**, 107–119 (2014)



Review of Relative Navigation for Noncooperative Spacecraft in Close Range

Dayi Wang¹, Qiyang Hu^{2(✉)}, Wenbo Li², Haidong Hu², and Kebei Zhang²

¹ Beijing Institute of Spacecraft System Engineering, Beijing 100094, China

² Beijing Institute of Control Engineering, Beijing 100190, China

hqy_1991@163.com

Abstract. On-orbit Service (OOS) has demonstrated great potential in future space mission. The target in the OOS mission are often noncooperative spacecrafts. The unavailability of the artificial retroreflectors and communication link makes it a great challenge to acquire the state of motion of the noncooperative target through relative navigation during the rendezvous in close range. Therefore the relative navigation represents one of the key techniques required for the success for the noncooperative rendezvous. In this article, the main research achievement about the relative navigation in close range for noncooperative spacecraft is reviewed. First, the state-of-art developments of noncooperative OOS projects and corresponding relative navigation scheme are reviewed. Second, the principle and application of different types of the electro-optical sensors are summarized. Third, the pose measurements techniques based on the 2D images and 3D point cloud are investigated respectively. Then, research on the relative navigation filter algorithm are summarized. Finally, the conclusion of this article and recommendation for the further development are presented.

Keywords: Noncooperative spacecraft · Relative navigation · Attitude estimation

1 Introduction

The On-Orbit Service (OOS) is referred to as the execution of assembly, repair, refuel, rescue, retrieval and maintenance of a spacecraft in-orbit [1]. During the OOS mission, Rendezvous between the servicing spacecraft and the target has to be realized before any operation. As regards the rendezvous, not only the position and orientation (pose), but also the state of motion (velocity and angular velocity) has to be determined through relative navigation. To date, the relative navigation has become routine operation in cooperative mission through identification of reflectors attached on the target. In addition, the measurements from the sensors on the target provide a stable basis for the Estimation results. However, most of the targets in OOS mission are noncooperative spacecrafts, which are not designed for the rendezvous originally. In this

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 750–772, 2021.

https://doi.org/10.1007/978-981-15-8458-9_81

case, no artificial reflectors to be recognized or the communication link between two spacecrafts are available, making the relative navigation a big challenge for noncooperative OOS.

According to investigation, the research of close-range relative navigation for noncooperative spacecrafts is based on the employment of electro-optical sensors, which can measure the superficial information of the target sufficiently. Different from far-range navigation problems, the ‘pose measurement’ is indispensable in most cases, which offer a preliminary pose of the target. Then the results can be implemented in navigation filter to realize the relative navigation.

This article will give a review of the relative navigation for noncooperative spacecraft. The structure is arranged as follows. The typical noncooperative OOS projects are firstly summarized, followed by the analysis of the feature and structure of the noncooperative relative navigation scheme in detail. Then, the research progress of the three key components of the relative navigation i.e. electro-optical sensor, pose measurement and navigation filter, are concluded respectively. Finally, the conclusion and perspective are proposed.

2 Current Research Status

2.1 State-of-Art of Noncooperative OOS Projects

In the past years, researchers from different countries and institutes have proposed many projects for the OOS for the noncooperative spacecrafsts.

The US is one of the first countries that carry out research on the unmanned OOS. A plan for the repairing of Hubble Space Telescope (HST) with space robot was proposed by NASA. The HST was regarded as noncooperative because of the lack of artificial retroreflectors attached to it [2], and the corresponding pose measurement algorithm was designed [3]. Although the plan was denied at last due to the immaturity of technology, it lay a solid foundation for further research. In 2005, the experimental small satellite, XSS-11 [4], realized the acquisition of the point cloud and pose measurement through the LIDAR it carried without relying on the artificial makers. In 2007, Defense Advanced Research Projects Agency (DAPRA) proposed the Front-end Robotics Enabling Near-Term Demonstration (FRRND) [5] for the demonstration of capture of noncooperative target in space. Its electro-optical system is composed of LIDAR and stereo vision, the measurement mode of which can switch automatically according to the measurement range. The prototype was developed successfully. The PHOENIX plan was also established by DAPRA in 2011 in place of FREND. One of its aim was the verification of noncooperative rendezvous and docking technology.

The research for the noncooperative OOS in Europe is mainly in the context of GEO defunct spacecrafts. As early as in 1990, European Space Agency (ESA) proposed the geostationary servicing vehicle (GSV) [6] scheme aiming at providing on-orbit surveillance and maintenance for GEO satellites. Experimental Servicing Satellite (ESS) [7] was planned in 1994 by DLR for the exploration of the implement of robot for OOS. In that program, the conical nozzle of apogee

engine, which is commonly installed on the GEO satellites, serves as the target to be recognized and captured. In 2002, the Robotic Geostationary Orbit Restorer (ROGER) [8], also proposed by ESA, was a scheme for the capture of noncooperative targets in GEO orbit with net. However, the projects mentioned above were terminated after the concept design for some reasons. In 2007, DLR initiated the Deutsche Orbital Servicing Mission (DEOS) [9], which was designed to acquire the relative pose through stereo vision. The project is expected to be verified in low orbits first. Besides, Smart Orbital Life Extension Vehicle (SMART-OLEV) [10] was developed jointly by German, Spain and Sweden, the purpose of which is to service and refuel the defunct communications in GEO. Its solution for the relative pose measurement was based on the combination of stereo vision and laser ranger finder. Corresponding algorithm has been proposed [11].

Japan is one of the few countries that has realized the rendezvous and docking technology on orbit. The Engineering Test Satellite-7 (ETS-7), launched in 1997 by National Space Development Agency of Japan (NASDA), represented the first free-flying space robotic system in the world. The navigation to a cooperative target from far, mid and close range was tested respectively on orbit with corresponding measurement pose strategies and sensors. The success in ETS-VII then became the milestone in OOS research. As regard the noncooperative cases, JAXA (Japan Aerospace Exploration Agency) proposed the pose measurement method based on the recognition of the docking rings instead of artificial makers. In addition, A test system that offer the simulation for target motion and imaging in space was developed, which enabled the verification of the point cloud-based method for noncooperative pose measurement on ground [12]. Over the past years more attention was paid to the space debris removal in Japan. In 2005, Japan's National Institute of Information and Communications Technology (NICT) proposed the Orbital Maintenance System (OMS) [13], a program aiming at the verification of rendezvous and docking and capture technology of noncooperative target, e.g. space debris. Binocular camera and eye-hand camera were designed for the pose measurement and state detection of capture mechanism respectively. Another project for the research on space debris removal is Space Debris Micro-remover (SDMR) [14], designed by JAXA in 2006. In this program, electro dynamic tether (EDT) as well as folded arm were planned for the capture in the final phase.

The main noncooperative OOS projects are collected in Table 1.

Table 1. Summary of noncooperative OOS projects

Object	Mission	Sensors	Time	Status
GSV	Vision Inspection, Robot operation and Debris	Vision camera	1994–1996	Concept
ESS	Vision Inspection, Satellite Capture, Docking/Release	Laser Range finder and stereo vision cameras	1994–1997	Concept
ROGER	Web Capture Satellite, Space Object and Debris	Stereo cameras, Laser range finder, Zoom camera	2003	Concept
OLEV	Life extension for servicing commercial GEO Spacecrafts	Vision camera and LIDAR	2003	Demonstration
DEOS	Orbit Maintenance System	Vision camera and LIDAR	2004	Ongoing Mission
OMS	Orbital Maintenance System	SVGA resolution COTS CMOS Image/ARM 9 processor and star sensor	2005	Concept
XSS-11	Autonomous Rendezvous and Proximity Manuvres	Space-borne scanning LIDAR and combined vision camera and star tracker by SAIC	2005	On-Orbit Demonstration
HRV	Servicing Hubble telescope	Vision camera and RADAR	2005	Concept
SDMR	Testing debris using tether for small satellite	GPS, Vision sensors and star tracker	2006	Demonstration
FREND	Non-cooperative Capture Serving and recycling	Solid state LIDAR and stereo vision cameras	2007	Demonstration Ground Testing
ACTS	Autonomous capture and servicing of satellites	LIDAR	2006	Concept

2.2 Analysis of Relative Navigation Mission

According to the projects and designs mentioned above, a typical noncooperative OOS mission mainly includes series of operations [15]: far-range rendezvous, close-range rendezvous (is usually classified into closing and final approach), target capture, and repairing. (removal for debris). In this paper, the interest is limited to the close-range rendezvous involving relative range from a few meters to hundreds of meters. The main object of far-range rendezvous is relatively simple—the reduction of relative trajectory dispersion, i.e. acquisition of position and velocity. Thus, Light-of-sight (LOS) and range measurements or only LOS information is used for the navigation filter in far range. While the measurement of relative attitude is of no possibility and necessity. However, during the close phase, the measurement of full 6 DOF pose of target is feasible with proper sensors owing to the availability of more information of target due to the reduction of relative range. In addition, the mass distribution of the target as well as the velocity and angular velocity of the target has to be estimated through dynamic and kinetic model, based on which the further strategy of approaching,

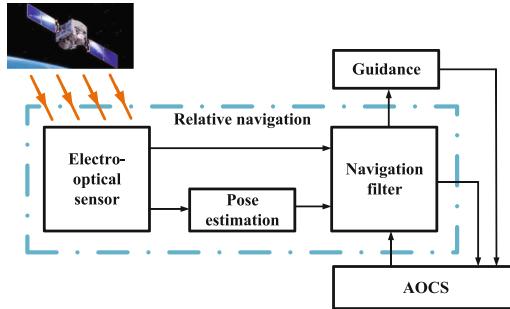


Fig. 1. Scheme of relative navigation for noncooperative spacecraft

detumbling or capturing can be designed. Then, when the chaser is in final approach phase, the structure to be captured on the noncooperative target should also be recognized. In conclusion, the goals of the measurement is summarized:

1. Acquisition of the state of motion, i.e. position, attitude, velocity and angular velocity.
2. Identification of dynamic parameter, i.e. centre of mass, moment of inertial (as much as possible)
3. Recognition of parts to be captured.

Consequently, relative navigation has to be implemented for the accomplishment of these aims. As regards the sensor, the angle-only or range-only sensors can't fulfil the requirement for 6 DOFs measurements in noncooperative case. While electro-optical sensors like monocular/stereo cameras or LIDAR can acquire enough information about the target in the sensor Field of View (FOV), thus representing the optimal sensors relative navigation.

The schematic overview of the GNC-system for noncooperative OSS is depicted in Fig. 1. While only the navigation part (within the dotted box) is what this paper focus on. As has been mentioned, a typical navigation scheme to a noncooperative spacecraft consists of three components: electro-optical sensors, pose measurement and navigation filter, working the following way: Firstly, these electro-optical sensors measure the target and provide 2D images or 3D point cloud (according to the selection of sensor) of the target in real time. Then pose measurement algorithm use these data to calculate the pose of the noncooperative target with respect to the chaser. These raw pose estimations will be fed to the relative navigation filter, e.g. an Extended Kalman filter, as an initial guess and the measure inputs. (The data from the sensor can be directly fed into filter without pose measurement according to some schemes) At last, the estimation of the state of motion and inertial parameter with relative high accuracy is realized through proper filters, which will be used in the control or guidance parts. In the rest of this paper, research progress on the three components will be introduced in detail respectively.

3 Electrical-Optical Sensor

During relative navigation, the property of the target and the illumination conditions have great impact on the measurement. On the one hand, the target is noncooperative spacecraft, which could be tumbling and is not stabilized. And no information about its states and no retroreflectors attached to the surface are provided. What's more, the target's surface is often composed of different materials, like strongly reflecting Multi-layer insulation (MLI) or other diffuse materials. On the other hand, the servicing spacecraft has to visually navigate under different kinds of illumination conditions. On one extreme the target can be completely dark and on the other, it can be in full sight. The sun can be located behind the chaser or behind the target or it may hit the target from the side, leading to sharp shadows, leaving a part of the target in darkness. The harsh measurements condition mentioned above as well as the limited recourse on orbit may bring the following difficulties: the change of image scale in large range resulting from the target's tumbling; the influence the illumination condition has on the images; the delay caused by the algorithm complexity; the miniaturization of sensors, all of which must be solved. Thus, one type of sensor with low mass, volume and energy consumption, low sensitivity to illumination and low algorithm complexity is required [16, 17].

Generally, electro-optical sensors [18], the device capable of collecting radiations reflected and/or emitted by the surrounding in the optical spectrum, is the optimal option for the noncooperative spacecraft pose measurement and navigation. According to chapter and table, the visible/infrared camera, laser range finder and LIDAR are electro-optical frequently used throughout the noncooperative on-orbit mission. However, the image from infrared, though enabling orientation measurement, are too poorly textured for the measurement. Similarly, laser range finder is able to compute only the range and LOS of the target (3 DOF) by measurement of the distance the laser beam travel. As a result, only the monocular/stereo camera working in the visible band and the LIDAR are qualified candidates for the 6DOF Full pose measurement. They can be further classified into passive sensor and active sensors based on whether they have energy source to emit radiation. The concise summary of the property of these sensors as well as the introduction of representative sensors tested on the ground or on orbit will be provided below (Fig. 2).

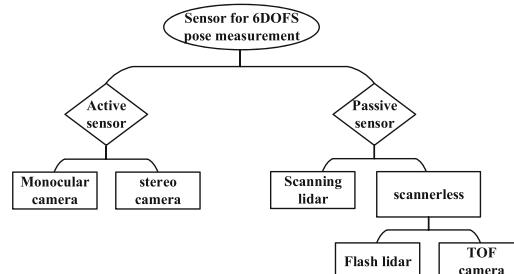


Fig. 2. Taxonomy of electro-optical sensor for relative navigation

3.1 Passive Sensors

In general, the passive sensors, i.e. monocular and stereo cameras, have lower hardware complexity, mass, volume and power cost. However, one noticeable drawback of the passive sensors is the sensibility to the harsh illumination conditions in space environment as well as the potential presence of other celestial body in the field of view, which lead to serious problems for its performance in poorly illumination scenario because of the unavailability of other light sources. Monocular camera is the most common electro-optical sensor used on orbit. The theory behind its application is simple (pinhole imaging), but the lack of depth information limits its usage. On the other hand, stereo camera can acquire the 3D information about the target by matching the feature in image planes. While the relevant image processing steps, e.g., image denoising and feature extraction, will inevitably result in the increase of algorithm complexity bringing measurement delay. What's more, the maximum range at which the depth resolution can be guaranteed is strongly limited by the baseline of stereo configuration. In that sense, the stereo camera appears to be less flexible and less convenient than the monocular one.

CCD/CMOS camera technology has a long flight heritage. Many camera-based measurement systems have been developed and tested on-orbit for the pose measurement. Proximity Operation Sensor (POX) [19], a monocular-based system designed by the National Space development Agency of Japan (NASDA), has been used for the ETS-VII mission successfully. Advanced video Guidance Sensor (AVGS) [20] is the main relative navigation sensor during the DART mission and contribute to the Orbit Express mission. Although these systems are designed for the measurement of cooperative target equipped with makers originally, in most cases they fit the noncooperative mission as well [21]. For instance, the Visual Based system (VBS), a stereo-based system design for the cooperative fly formation mission, also helps for the test of the noncooperative pose measurement algorithm. The Relative Navigation Sensor (RNS) system [22], which is composed three monocular cameras, is assessed during the STS-125 mission. Two of the three cameras are also integrated into the Argon, a test-system designed for the noncooperative target rendezvous and docking. Besides, the examples of camera tested through simulated environment can be seen in [23–25].

The main technology specification of the representatives monocular and stereo cameras is illustrated in Table 2.

Table 2. Technology specification of representative cameras

	Measurement principle	Pixel size (um)		Pixel number		FOV (°)		IFOV (°)		Focal length (mm)	Baseline (mm)
VBS	Monocular	8.6	8.3	753	582	22.3	16.8	0.03	0.03	16.4	—
RNS	Monocular	6.4	6.4	1024	1024	23	23	0.02	0.02	16	—
[23]	Monocular	7.4	7.4	640	480	49.1	37.8	0.08	0.08	5.2	—
[24]	Stereo	27.5	27.5	512	512	22.3	16.8	0.11	0.11	13.5	0.4
[25]	stereo	3.2	3.2	2048	2048	14.9	14.9	0.01	0.01	25	0.5

3.2 Active Sensor

The active sensor, i.e. LIDAR-based sensor, is characterized by relatively higher structure complexity, cost, and power consumption, with respect to the passive one. Different from stereo camera, it can directly acquire the 3D point cloud from the surface of target through the emission and detection of laser beam without any processing. The most advantage it has over the visible camera is its insensitivity to the poor illumination condition. It can also discriminate the target from background. According to technology solution, the LIDARs can be divided into scanning LIDAR and scannerless one. The scanning LIDAR has only one detector so that the laser beam has to move to scan the target in the FOV based on certain scanning pattern, e.g. Lissajous, tosette and spirals. Even though the calibration is simple, some issues make the use of scanning devices as main pose measurement navigation sensors extremely challenging. First, the moving part is more likely to result in hardware failure. Second, the corresponding processing rate is limited by the time needed to mechanically scan the FOV. Last but not least, the tumbling of the target may lead to the distortion of the point cloud acquired during time needed for scanning. While the scannerless LIDARs use broad laser beam to illuminate the entire scene and measure the laser TOF to the pixel by detector array instead of one detector at one time. Consequently, the resulting 3D point cloud are not subject to the undesirable motion blur. Another great advantage the detector array LIDAR systems has over scanning ones in on orbit application is that they have no moving part, which means the removal of a key source of failure from the scanning LIDAR. Nevertheless, detector array, which consist of more detector, are relatively challenging to calibrate. And the implementation consideration constraint of the size of these detector array limits the size of point cloud, leading to poorly and fixed spatial resolution. Two types of the LIDAR with detector array suitable for the noncooperative pose measurement are flash LIDAR and TOF camera. A flash LIDAR send out the laser pulse ('flash') and detect the TOF of the flash at each pixel. It can effectively restrain the illumination in the background and have relatively longer detection range. However, the requirement for the measurement accuracy of time limit its implement. In contrast to the flash LIDAR, TOF measure the phase difference between the laser emitted and then detected at each pixel, the intensity of which has been modulated. Its structure is relatively more compact. It has low mass, volume and power consumption relatively. However, the possibility of phase integer ambiguity limits its employment to only close range.

Many LIDAR-based systems are under development for the noncooperative pose measurement, some of which have be implemented on orbit successfully. The laser camera system (LCS) [26], developed by Neptec, is a triangulation scanning LIDAR exploiting the Lissajous scanning pattern. Two operation models are designed for inspection mission or identification of discrete objects respectively. The Triangulation LIDAR sensor (Tridar) [27], also developed by Neptec, combines the close-range triangulation system (similar with LCS) with a TOF LIDAR, which enables the pose measurement from mid rage to close range with the LASSO software. It represents the first active sensor system tested in space

for measure noncooperative target and has underwent updates between flight experiments. The ACS's Dragon Eye [28] is a Flash LIDAR, employing a InGaAs Avalanche Photo-Diode detector array. The performance of its prototype has been assessed during the STS-127 and STS133. Another representative flash LIDAR system is Vision Navigation Sensor (VNS), developed by Ball. It has flown on STS-134 for sensor test and is integrated in the Argon system for non-cooperative measurement. To the author's knowledge, the TOF cameras have never been used in space due to the lack of technology maturity. However, the Swiss ranger (SR4000) [29], a TOF camera developed by MESA, has been used for the algorithm verification in the laboratory test. A summary of information about the active sensor is given in Table 3.

Table 3. Technology specification of representative LIDARs

	Operational range	Operational range (m)	Range accuracy (cm)	Mass (kg)	Power consumption (W)	FOV (°)
LCS	Scanning LIDAR	1–10	0.01–0.5	12.1	75 (max)	30 × 30
Tridar	Scanning LIDAR	0.5–2000	N/F	15.9	70	30 × 30
Dragon Eye	Flash LIDAR	< 1500	10	3	35	45 × 45
VNS	Flash LIDAR	< 5000	10–20	12	30	20
SR4000	ToF camera	0.1–5	0.01	0.47	9.6–24	69 × 55

Table 4. Comparison of electro-optical sensors

	Monocular vision	Stereo vision	Scanning LIDAR	Flash LIDAR	TOF camera
Active/passive	Passive	Passive	Active	Active	Active
Need scan or not	Not need	Not need	Need	Not need	Not need
Depth information	Not	Yes	Yes	Yes	Yes
Movement distortion	Inexistence	Inexistence	Existence	Inexistence	Inexistence
Sensitivity to light	High	High	Low	High	High
Cost	Low	Low	High	High	Medium
Algorithm complexity	Medium	High	Medium	Low	Low
Compact	High	Low	Low	Medium	High
Maturity	Very mature	Mature	Mature	Immature	Immature

The comparison of main property of measurement sensors for noncooperative spacecraft in close range is summarized in Table 4 [30]. As can be seen from the table, each type of sensor has its own superiority and weakness. As a result, it is impossible to fulfil the measurement requirement relying on one sensor or even one type of sensors. Instead, sensors of different types should work together,

complementing one another and compensating each other, to realize the measure with high accuracy and reliability. For instance, Argon [31], the vision system developed for the noncooperative navigation, integrate two RNS cameras and a VNS flash LIDAR, thus taking the most advantage of the two sensors. The combination and infusion of sensors of different type is the new development trend.

4 Pose Measurement

Pose measurement, as its name implies, is the acquisition of relative position and attitude between two reference frames defined at the servicing spacecraft and the noncooperative target respectively. It's worth noting that during the pose measurement phase, only the optical measurement data is used without employment of any dynamic or kinematic property and model. Thus, the obtained pose doesn't necessarily correspond to the actual centre of mass or principal inertial axes of the target. Actually, its establishment is merely based on superficial geometry of the body in most cases. The discussions of pose measurement methods can be mainly classified into image-based method and point cloud-based method corresponding to the implement of monocular/stereo cameras and LIDAR respectively.

4.1 Image-Based Method

To some extent, the pose measurement with cameras is fully dependent on the features extracted from the images. In this case, the components on the surface of the noncooperative spacecraft often serve as the target to be recognized with priority. These components, e.g. solar panel, engine nozzle, solar panel support, are common even in a noncooperative target. Their projection to the image planes are often simple geometrical shapes, e.g. rectangle, circle, ellipse and triangle, which are easy to identify. Consequently, the recognition of aforementioned components can favour the design of pose measurement scheme in the following ways: Firstly, it offers the benchmark for the establishment of reference frame at target. Secondly, it helps to acquire the features for pose measurement algorithm. What's more, the geometrical constraint contributes to solving the pose. Last but not least, the component itself is sometimes designed to be capture, e.g. apogee engine nozzle. Once these components attached in the target are recognized, a close form solution can be trustfully used over some constraints especially through stereo vision. These analytical solutions exist for 3 points, 4 coplanar points, circle etc., allowing fast computation of the pose. In [24], the pose is measured based on the recognition of the rectangular structure in the antennas by collaborative cameras. The four points are extracted as intersection of the four sides of rectangle. The origin of the reference is at the centre of the rectangle. The equation related to the four points is then established according to the geometrical property that the side in the rectangle is parallel and perpendicular to the opposite and adjacent side respectively, the solution to which can

be reorganized to acquire the pose of target. A similar scheme is proposed in [32] 2016. However, only the partial rectangle is recognized owing to the limitation of FOV in close range. The structured light system composed with a monocular and a point light is used instead of two cameras. In [15], the solar panel support is recognized, the vertex of which are acquired simultaneously. The centre of the solar panel support in the two images is then calculated as the centroid of the triangle at which the reference frame of target is defined. Similarly, the three vertexes as well as the centre of the inscribed circle of the solar panel triangle structure are extracted in [33,34]. The coplanar points are then fed to the homogenous equation to solve the monocular-based pose measurement. Besides, Peng proposed a binocular-based method based on the recognition of launch vehicle interface ring and apogee engine nozzle [34]. The acquisition of pose is realized based on the calculation of the centre position, radius and normal vector of the circle structure on the aforementioned components, the projection of which are ellipses on the two image planes.

However, the employment of aforementioned methods may have some constraints. On one hand, it highly depends on some components on the target. The algorithm may not work without the availability of corresponding components. On the other hand, the algorithm only adopts fixed number and type of features, failing to take full advantage of the features from the images. Consequently, the result is more prone to suffer lack of robustness in the presence of noise and outlier. In general, the pose measurement not relying on very components is accomplished by using natural features e.g. as corners, lines, contours or feature descriptors. Among them, the feature points (corners or feature descriptor) are the most common features used. At least 3 feature points can fulfill the requirements for the establishment of reference frame on target: the line connecting two points is set as the first axis, and one of these points serve as origin. The second axis is perpendicular to the plane composed by the three points while the third one follows the righthand law [35]. The pose measurement based on the stereo vision is much easier. Because the 3D information of the feature points can be acquired, the pose can be obtained directly after the definition of the reference frame. However, when monocular sensors are used, the 3 DOF position of the target can't be fully measured due to the lack of depth information. One solution is the aforementioned structure light system, the combination of a monocular camera and a point light. The projection of point light on the image serve as the reference for the monocular camera to acquire absolute position information according to the known configuration of the sensor system. In addition, the life-sized model or at least the distribution of features (the coordinates of the features in the target coordinate frame) often serve as the a priori knowledge in most monocular-based pose measurement algorithm, which enables the acquisition of full 6 DOF pose. When the computation of pose is based on the 2D feature points of the target's image and corresponding 3D features on the model, the pose problem is referred to as perspective-n-points (PnP) problem. In [1], numerical methods for the PnP problem, e.g. Pose from Orthography and Scaling with Iterations (POSIT), Derivative Based Estimation (DBE) and

Horaud-Phong (HP), are summarized, each of which differs in the choices for the mathematical representation of the pose, the cost function to be minimized and the optimization method from each other. Then a new dual quaternion-based modified Horaud-Phong (MHP) algorithm is proposed. Numerical simulation and ground truth experiments are performed analyzing the convergence, noise, number of correspondences and estimation of motion of different aforementioned PnP solvers, showing the advantage the MHP algorithm has over other classical methods in its high accuracy and the reduction of iteration number. Because the solutions of the PnP problems are obtained by optimization of a defined object function, the iteration may suffer the problem of not guaranteeing convergence and becoming trapped in the local minima of the object function. Thus, a good starting estimation is required during the initial acquisition step. In this respect, the comparison of the initial pose estimation techniques for PnP problems, i.e. POSIT, coplanar POSIT, Efficient PnP (EPnP) and Newton-raphson, is conducted in [36] taking into consideration the computational runtime, the 3D points re-projection error and estimation error from the simulation results.

4.2 Point Cloud-Based Method

Different from the image-based method, the raw point clouds data from LIDARs are processed directly without image processing or feature extraction in most cases, which reduces the sensitivity to the harsh illumination condition and inevitable occlusion. As regards the point cloud, the pose measurement problem equals to the acquisition of the transformation that maps the measured point cloud to the preexisting point cloud of the model of target (stored or built on board), which is also referred to as point cloud registration. The Iterated Closest Point (ICP) algorithm is one of the standard technique to address the problem, which iteratively find the closest point for registration and solve the pose through the iterative reduction of specified error metric. The main steps of the ICP algorithm are as following [37]:

1. For each point in the data set, search for the closest points that belong to the point cloud model
2. Compute the pose according to the aforementioned registration
3. Apply the calculated pose to the data
4. Calculate the divergence between the updated data set and corresponding points in the model.

The algorithm iterates over steps 2–4 until the divergence is within allowable threshold or the maximum allowed number of iteration is reached. In [37], the kd sorting is implemented to reduce the runtime of searching point correspondence in step 1. The establishment of the kd tree, a generalized version of binary search tree, of the point model logarithmically reduce the computation load when searching for nearest neighbours. In [38], the four pose estimation techniques for the aligned two pieces of point cloud, i.e. singular value decomposition (SVD), unit quaternions, orthonormal matrices and dual quaternions, are analysed. The

SVD method is selected eventually for its robustness and accuracy. Two forms of error metric (corresponding to the step 2), i.e. point-to-point and point-to-plane are also compared, the former of which is more commonly used.

During the navigation, once the navigation filter acquires the result from the ICP algorithm, the pose can be predicted through the dynamic model, which serve as the initial estimation for the ICP when the next measurement comes. In this sense, the continuous pose tracking is realized. However, the convergence of ICP algorithm is sensitive to the very first guess of pose, corresponding to moment when the first point data is acquired and no a priori information about pose is available. Inappropriate staring estimation may lead to the local minimal solution, as is the case similar with the numerical methods for PnP problems. The calculation of initial pose is also referred to as ‘pose acquisition.’ Many research on the pose acquisition methods have been developed. In [39], the direction of the major axis corresponding to the elongation of the point cloud is determined by the Principle Component Analysis (PCA) to the covariance matrix while the left 4 degrees of freedom, i.e. the rotation around the major axis and all three translations, are recovered through the exhaustive matching method. The PCA-based method is demonstrated to have robustness to outliers, but its employment is limited to elongated target. A template matching algorithm is executed in [40]. The databases corresponding to the possible attitude configurations sampled by fixed angular step are generated, among which the one minimizes the error metric in the IPC algorithm is selected as initial pose. The worst attitude configuration is rejected before the ICP, speeding up the template matching procedure. Besides, the Oriented Unique and Repeatable Clustered Viewpoint Feature Histograms (OUR-CVFH) algorithm is implemented in [41], which realize the object recognition and pose estimation simultaneously. The global OUR-CVFH descriptors of the point cloud are calculated respectively for measurement data and model, based on the matching of which the raw pose is then acquired. The comparison between OUR-CVFH and another feature-based pose acquisition technique (i.e. spin images) can be seen in [42].

5 Navigation Filter

As mentioned, only a coarse results can be obtained through the pose measurement, which can't be used directly in the GNC loop due to the following inner drawbacks: Firstly, it is easily contaminated by the noise of the sensor resulting from harsh sensing condition as well as the limitation of capacity of the sensor. Secondly, only when the outcomes of the sensor arrive is the state of motion of the target available. In other words, the pose estimation from the sensor is an instantaneous one, which can't meet the requirement for continuous measurement. Last but not least, the pose acquired from the sensor is the relative change of the coordinate system defined based on the geometrical attribute of the target, as has been mentioned. (It is commonly defined at the centroid or the centre of some notable geometrical structure). However, it is the centre of mass and the principal axis coordinate frame defined at it that have engineering

value and have to be estimated. As a consequence, the coarse pose information (or the features) measured by the sensor should be fed to a navigation filter so that a smoother and continuous outcome can be acquired. Navigation filter is a recursive estimation algorithm, which predicts the states to be estimated through process equation and then correct them through the outcomes from the sensor measurements modelled by measure equation. Based on Kalman filter algorithm and its extension, it makes full use of the dynamics and kinematics relation between states, thus realizes the estimation of the pose and even its rate of change, the latter of which can't be measured directly.

5.1 Process Modelling

The position, attitude as well as their rate of change have to be estimated when addressing the navigation in close range. In [42], two paralleled filters for the translational and rotational motion respectively are exploited. Sequences of range images generated by a team of cooperating sensors are fused to produce the knowledge of the position of the centroid and the attitude of the principle geometric axes of the target with respect to sensor reference while building its geometric shape in form of occupancy-grid-based map. To get the position of the centre of mass, the offset between the centre of mass and the centroid is regarded as constant parameter that has to be estimated. To the best of the author's knowledge, it is the first time that the estimation of the full state of motion and the mapping of the geometric shape of a tumbling spacecraft is realized simultaneously. However, the translational part of the process model is simplified, assuming that the target is under constant velocity. This may result in considerable error when applied to targets tumbling violently. According to investigation results, the well known Clohessy Wiltshire (CW) and TH equation, which are simplified equation under the assumption that the spacecrafnts are running on circular or elliptical orbit respectively, are the equations more commonly used to model the relative translational motion of two spacecraft, especially the former one. Because the CW equation is defined under the Local Vertical Local Horizontal (LVLH) coordinate system of the spacecraft, it will bring much differences defining the LVLH at the noncooperative spacecraft or the servicing spacecraft. In [43], the coordinate is defined at the servicing spacecraft so that all the orbit parameter can be regarded as known. In fact, to approach and inspect the target, the servicing spacecraft has to be maneuvered constantly, which makes it hard to fulfil the assumption of free motion during the inference of CW equation. While in [44], when the coordinate frame is defined at the unknown centre of mass of the noncooperative, the unknown orbit angular in the CW equation has to be estimated, which bring complexity to the filter algorithm. An eclectic solution is proposed in [45] by Sonnenburg. The LVLH coordinate system is set at the target, while the orbit parameter of noncooperative spacecraft is replaced by that of the servicing spacecraft supposing that the two spacecrafnts are close enough so that their orbits are approximate. On the other hand, in [46] Segal demonstrated that CW equation is not qualified for the formulation of motion

of the feature points on the surface of the target that are not located at the centre because of the coupling between the translational and rotational motion. To address this problem, he proposed the process model of the feature points taking the coupling relation into consideration, which is implemented for the design of the navigation filer in [43]. It is worth noticing that not only the state of translational and rotational motion but also the coordinate of the feature point of target under its body coordinate system are to be estimated because no a priori knowledge of the target is available. This above method, estimating the state of motion as well as the position of the feature geometrical simultaneously, is often called SLAM (simultaneous localization and mapping) method. The concept of SLAM inherits from the robotic community, referring to the robotic equipped visual odometry estimating its own pose while building the map of the unknown environment. In most cases of robotic estimation, the environment (landmark) is static while the robotic is moving. When the SLAM method is introduced to the relative navigation of noncooperative spacecraft, the position of features is regarded as landmarks. The only difference is that the environment (the noncooperative spacecraft) is also moving now. As regards the rotational part of the process model, the euler equation and the velocity angular differential equation according to the theory of mechanics of rigid body is always implemented. To simplify the model, the target is often assumed to be free of control torque and the disturbs is treated as white noise with zero mean. In [47], Dong present the attitude in the form of three euler angles. When conforming to the 3-1-3 rotation, these angles can correspond to the precession, nutation and spinning of the noncooperative target, which bring noticeable description of the rotational motion. While in [48] the Modified Rodrigues parameters are used. However, the expression of attitude mentioned above may suffer the singularity when the target is tumbling within a large range of angle, which is the inherent defect of the expression of attitude with three degrees of freedom. To solve the problem, [43] adapts quaternion instead in 2014. Because the attitude of a rigid body with three degrees of freedom is expressed by the four elements of the quaternion, the singularity can be avoided. Every time the new quaternion is calculated during the prediction or correction of the quaternion-based navigation filter, it has to be normalized [49]. This brute-force approach for preserving quaternions norm is not elegant, which may lead to considerable estimation errors. To make things worse, regarding the quaternion as a 4-dimension state may result in the rank defect of the covariance matrix because only three elements of the quaternion are independent, which may bring numerical problems for the navigation filter. In [50], Aghili adopts a scheme to choose the quaternion error instead of the quaternion itself as the state to be filtered, which is firstly proposed by Lefferts for the attitude determination of spacecraft [51]. The quaternion error is defined as the difference between the real quaternion and its estimation and the scalar part is ignored. Then the three degree of freedom attitude can be expressed by the quaternion error of three dimensions, avoiding the rank defect while preserving the quaternion norm. The filter algorithm dealing with the error quaternion will be illustrated later. It is worth noticing that in [50], the discrete process model as

well as the discrete form of Covariance matrix of process noise is derived, which enable the computer implementation of the navigation filter. For the euler equation, the target inertial matrix is necessary, which is not realistic when no a priori knowledge about the target is available. To overcome this contradiction, the inertial property of the target should be estimated. A method is proposed in [43] to choose the moment of inertia from a set of possible results by using the maximum posterior estimation and the estimated state filtered by paralleled filter modelled under the hypothesis of each moment of inertia in the above set. However, the lack of information about target makes it difficult to decide a set of possible moment of inertial in advance. [51] has proved that the magnitude of the three moment of inertial are unobservable based solely on vision information without any external continuous stimulation exerted on the target. On the other hand, because the euler equation can be propagated with the moment of inertial scaled to any non-zero ratio in torque free motion, one solution is to treat the ratio of the inertial of the principal axis as the constant parameters to be estimated, as can be seen in [43, 50].

5.2 Measurement Modelling

The navigation filter relies upon the measurement results from the electro sensor for the estimation. Once the coarse pose estimation is acquired, it can be fed to the filter algorithm directly. In [50], the pose of the coarse measurements result by the LIDAR through ICP algorithm is regarded as the pose of a nominal coordinate of system. It does not necessarily coincide with the principal-axis coordinate system and a constant bias is supposed to exist between them. Then the measured pose serves as the input to the filter in the form of quaternion and position vector. The covariance matrix of measurement noise in the form of the quaternion and euler angles is also derived, the former of which is state-dependent because of the constraint imposed by the unit-norm property of quaternion. Similarly, Li, in his master thesis [52, 53], feeds the quaternion measured from the stereo-vision-based attitude determination algorithm into the navigation filter algorithm directly. However, when pose measurement algorithm is available, the features themselves have to serve as the measurement input. In this case, the measure mechanism of the sensors has to be taken into consideration for the measurement modelling, especially when dealing with camera-based filter algorithm. In [43], the navigation filter tracks the 10 features on the surface of the target. The relationship between the coordinates of the features in the camera coordinates system defined at one of its focus and their projection on the two image planes is built as the measurement model presuming that the two image planes are paralleled and that the baseline and focal of the binocular camera have been calibrated. What's more, the optical flow, the velocity of the feature point projection on the two planes, are also consider measurable. As the three-dimension coordinates of the feature point can be reconstructed through the match of its projection in two planes, [48] in 2013 proposed a simplified measurement model, where the recovered coordinates can be implemented directly as the measurement of filter. One the other hand, monocular measurement model

is different from the binocular one because only two degrees of the feature points can be acquired because of the inherent scale ambiguity of monocular mechanism. In [45], the measurements result of the monocular vision is the normalized position vector of the feature point while it is the azimuth angle and the elevation angle that are measured in [54]. It's worth noting that the measure model of the inertial measurement unit, i.e. gyro and accelerator, and the star sensor to be used for the determination of the absolute state of motion of the servicing spacecraft are involved in the filter algorithm, the model of which are always simplified under the assumption that all the states of motion of the servicing spacecraft has been acquired and considered as known parameters. According to the two types of monocular measurement model mentioned above, only the line of sight of the feature points is available. As a result, the lack of depth information may result in the shortage of the observability of the translational part of the state to be estimated, which will be discussed in detail later. Besides, a new pseudo measurement equation is implemented to the filter in [50]. Actually it is not really the result of measurement from the sensor but the euler equation the target should conform to under the assumption of torque free motion. In this way, the constraint of the euler equation is imposed to state to improve the observability and boost the convergence of the filter.

5.3 Filter Algorithm

Since the process and measurement models introduced above have strong nonlinearity, the linearization-based EKF [51] is often a suitable candidate. The EKF dealing with error quaternion is called a Multiplicative EKF (MEKF). In the MEKF algorithm, the reference quaternion is used for prediction while the three-dimension error quaternion, which complies with assumption of small angle, is for the correction. The estimated unit-norm quaternion is then calculated through quaternion multiplication between them. However, sometimes the conventional EKF can't meet the accuracy requirement because only one-order of approximation is implemented during linearization. In 2017, the IEKF algorithm [43] is adopted to improve the filter accuracy. In contrast to standard EKF, the correction in the IEKF should be implemented more than one time until the error is within the threshold. As a result, the maximum posterior estimation is acquired. On the other hand, other types of filter algorithm designed for nonlinear model is used, such as the certainly-sampled-based Unscented Kalman Filter (UKF), Cubature Kalman Filter (CKF) and randomly sampled-based PF. In [55], a EKF-UKF-PF combination filter is designed, which make use of the respective advantages of the three filters algorithm, for the relative pose estimation and navigation in vision SLAM framework. In [56], Cao implemented the UPF filter algorithm, the combination of the UKF and Predictive filter, the latter of which determines the optimal model error using a one-step ahead control approach to provide accurate state estimations, leading to better performance than EKF dealing with dynamic model of low accuracy. Besides, the complex space environment brings to the navigation filter algorithm at least two following challenges that have to be surmounted. On the one hand, the harsh measurement

condition and the violent relative motion between the sensor and the target may have great influence on the statistical characteristics of the measure noise. On the other hand, the occlusion, the inadequate lighting condition and the shadow in the space may raise the possibility of outliers in measurements. As a result, it is imperative to design proper filter algorithm that is adaptive to measurement noise and robust to measurement outliers. To address the problem, Agalili in [50] adopts an adaptive EKF algorithm that regard the error and its covariance matrix as state-depend parameter to be estimated. In [57], a filter outlier detection method based on M-estimation was proposed by Yu. A judgement rule is then given about whether the observation data is valid according to the result of the M-estimation. Different from the above method, Hou uses the multiple fading factor to tune the Kalman filter gain to against the influences caused by the faulty measurement. The judgement is realized through the chi-square detection to a statistical function of the residue. The same chi-square-detection-based judgement method also appears in [58, 59], where the variational Bayesian technology is used for an adaptive response to the measurement noises. In addition, Xiong proposed an adaptive filter algorithm which consists of several paralleled IEKF with different numbers of iteration [60]. The estimation result is the sum of the weighted results from these paralleled IEKF and the weights are adjusted adaptively according to the measurements. The simulation result shows the superiority to any IEKF working alone.

5.4 Research on Observability

Investigating the observability plays a significant role in navigation filter design. It offers a way to evaluate the feasibility of estimating the navigation state from the measurements through the proposed filter algorithm. In this sense, observability serve as the premise for the implementation of a filter algorithm. Since the observability is completely based on the process and measurement model of the filter, no consensus has been reached on the observability of the problem of relative navigation of noncooperative spacecrafts because of the differences in building models based on different forms of the navigation states and sensors used. The degenerate conditions proposed by Lister for the LIDAR-based filter algorithm in [42] shows that there is no unique solution to the principal axes if two or more principal inertias are identical and that the inertia is not observable when the target is spinning around a single axis. The estimator convergence time according to the nutation angle which changes from 0 to 90°, corresponding to the spinning of target from about the major principal axis to about the minor principal axis, is also illustrated. However, the analysis mentioned above in [59] is absolutely based on the numerical simulation results. Strictly speaking, the observability analysis should precede the design. Only after the observability is verified can the filter algorithm be implemented for the simulation, which means that the conclusion in [59] can serve merely as a verification for the observability but not observability analysis itself. In fact, the observability analysis in theory for the relative navigation filter in close range is much harder than dealing with relative navigation in far range or other absolute navigation problem due to the

high dimensions of the system (at least 12 dimensions, 6 for translational and rotational part respectively) and the complexity of the process and measurement model. Mishra, in his Master thesis, design three LIDAR-based filters, one of which is identical with that in [42]. The observability analysis is carried out by the calculation of the Observability Rank Condition (ORC), a matrix which consists of different orders of lie derivatives. The conclusion similar with that in [59] about the unobservability conditions is reached. The degrees of observability of the three filters are also compared based on the local unobservability index and local condition number. However, only the simplified process is used for ease of analysis. The observability of a binocular-vision-based navigation filter is offered in. Different from [59], an observability analysis method designed for Piece-wise constant system (PWCS) is adopted, confirming the feasibility of estimation of the all the states by tracking the 6 feature points on the target once none of the components of the relative angular velocity vector is zero. But it doesn't research the influence the number and distribution of the feature points has on the observability system. Besides, a monocular-based filter algorithm is more likely to result in the observability problem than using binocular camera or LIDAR, as has been mentioned before. In [61], Song explores the observability of the filter designed using line-of-sight measurements. The ORC similar to that in [59] is calculated, the rank of which is calculated resorting to the Matlab Toolbox instead due to the high dimension of filter system designed (26 dimensions). The two premises for the observability, i.e. the camera offset and orbit maneuver, are then verified through the analytical method designed by Geller [62] for the angle-only navigation during orbital proximity operation. Even though preliminary conclusion can be made about the conditions which may impact the observability such as the ambiguity of the inertias and the pure spinning motion of the target according to the results from aforementioned numerical and analytical analysis, the research on observability for the close-range relative navigation for noncooperative spacecraft is still an open research area.

6 Conclusion and Perspective

The relative navigation for the noncooperative spacecraft play a significant role during the noncooperative rendezvous, which serves as the precondition for the success of an OOS mission. This article aims to review the relevant research progress, providing a reference for follow-up studies. Based on the research of the existing noncooperative OOS projects, the configuration of the relative navigation is firstly summarized. Then, principal and application of the electro-optical sensors suitable for the relative are illustrated after the analysis of the measurement condition. Additionally, the pose measurement methods corresponding to different measurements sensors are reviewed. Finally, the research on the navigation filter is introduced based on the analysis of every parts of a filter algorithm. It is worth noticing that this review put emphasis on the research about three key components of the noncooperative relative navigation problem, i.e. the electro-optical sensor, pose measurement and navigation filter. In fact, there are some

other important areas of research about this problem, e.g. the image processing technique for the sensors measurement and the experimental verification methods for the navigation algorithm and schemes (through software or hardware). Due to length limit, they are not in the scope of this article.

Based on the investigation of existing research progress and characteristic of the noncooperative close-range navigation in OOS mission, some suggestions are proposed for the future research, which can be classified corresponding to the three key components of a relative navigation scheme respectively:

1. The sensors suitable for noncooperative measurement should be developed. On one hand, the sensors ought to cope with the harsh measurement condition. On the other hand, it is necessary for the sensor to balance the measurement performance and the resource occupation on orbit.
2. The pose measurement method that applicable to targets suffering lack of a priori knowledge should be researched. The realization of the SLAM-based methods for the noncooperative navigation problems is worth further studying.
3. The navigation filter based on the information infusion should be designed. Which can realize the optimal estimation based on the measurement data from different types of sensors. Besides, the filter algorithm has to overcome the inevitable deficiency and delay of measurements caused by the occlusion and runtime needed for pose measurement respectively.

Acknowledgment. This work was financially supported by National Natural Science Foundation of China (61690215, 61640304, 61573060, 61203093) and National Science Fund for Distinguished Young Scholars (61525301). The author wishes to thank all these that contribute to this article.

References

1. Arantes, M.I.G.: Rendezvous with a non-cooperative Target. Master thesis, University Bremen (2011)
2. King, D.: Hubble robotic servicing: stepping stone for future exploration mission. In: AIAA 1st Space Exploration Conference: Continuing the Voyage of Discovery, pp. 2005–2524. AIAA, Reston (2005)
3. Thienel, J.K., Sanner, R.M.: Hubble space telescope angular velocity estimation during the robotic servicing mission. *J. Guid. Control Dyn.* **30**, 29–34 (2007)
4. Richards, R., Tripp, J., Pashin, S., et al.: Advances in autonomous orbital rendezvous technology: the XSS-11 LIDAR sensor. In: Proceedings of the 57th IAC/IAF/IAA (International Astronautical Congress), Valencia, Spain 2–6 October 2005
5. Thomas, D., Sean, D.: Overview and performance of the front-end robotics enabling near-term demonstration. In: AIAA Infotech Aerospace Conference, Seattle, USA, 6–9 April 2009
6. Depreter, W., Visentin, G., Fehse, W.: Satellite servicing in GEO by robotic service vehicle. *ESA Bull.-Eur. Space Agency* **7**, 22–25 (1994)
7. Hirzinger, G., Landzettel, K., Brunner, B., et al.: DLRs robotics technologies for on-orbit servicing. *Adv. Robot.* **18**, 139–174 (2004)

8. Bernd, B.; Kerstein, L. ROGER-robotic geostationary orbit restorer. In: 54th International Astronautical Congress of the International Astronautical Federation, Bremen, Germany (2003)
9. Rupp, T., Boge, T., Kiehling, R., et al.: Flight dynamics challenges of German on orbit servicing mission. In: International Symposium on Space Flight Dynamics, Toulouse, France, 27 September–2 October 2009
10. Kaisera, C., Sjöberg, F., Delcurac, J.M., Eilertsen, B.: SMART-OLEV an orbital life extension vehicle for servicing commercial spacecrafts in GEO. *Acta Astronaut.* **63**, 400–410 (2008)
11. Khongsab, P.: Signal processing and performance evaluation of a PMD camera for space docking. Ph.D. thesis. Lule University of Technology (2009)
12. Terui, F., Kamimura, H., Nisida, S.: Motion estimation to a failed satellite on orbit using stereo vison and 3D model matching. In: The 9th International Conference on Control, Automation, Robotics and Vision, Singapore, 5–8 December 2006
13. Kimura, S., Nagai, Y., Yamamoto, H., et al.: Approach for on orbit maintenance and experiment plan using 150kg class satellites. In: IEEE Aerospace Conference Big Sky, USA, 5–12 March 2005
14. Nishida, S.I., Kawamoto, S., Okawa, Y., Terui, F., Kitamura, S.: Space debris removal system using a small satellite. *Acta Astronaut.* **65**, 95–102 (2009)
15. Xu, W.F., Liang, B., Li, C., et al.: Autonomous rendezvous and robotic capturing of non-cooperative target in space. *Robotica* **28**(5), 705–718 (2010)
16. Rems, F., Rissee, E.-A., Benninghoff, H.: Rendezvous GNC-system for autonomous orbital servicing of uncooperative targets. In: GNC 2017: 10th International ESA Conference on Guidance, Navigation and Control System, Salzburg, Austria, 29 May–2 June 2017 (2017)
17. Liang, B., He, Y., Zou, Y., et al.: Application of time-of-flight camera for relative measurement of non-cooperative target in close range. *J. Astronaut.* **37**(9), 1080–1088 (2016)
18. Opronolla, R., Fasano, G., Rufino, G., et al.: A review of cooperative and uncooperative spacecraft pose determination techniques for close-proximity operations. *Prog. Aerosp. Sci.* **93**, 53–72 (2017)
19. Kasai, T., Oda, M., Suzuki, T.: Results of the ETS-7 mission—rendezvous docking and space robotics experiments. In: Proceedings of the 5th International Symposium on Artificial Intelligence, Robotics and Automation in Space, ESTEC, Noordwijk, The Netherlands, ESA SP-440, pp. 299–306, June 1999
20. Rumford, T.E.: Demonstration of autonomous rendezvous technology (DART) project summary. In: Proceedings of SPIE Space Systems and Technology Operations, vol. 5088, August 2003
21. D'Amico, S., Benn, M., Jorgensen, J.L.: Pose estimation of an uncooperative spacecraft from actual space imagery. *Int. J. Space Sci. Eng.* **2**(2), 171–189 (2014)
22. Naasz, B.J., Burns, R.D., Queen, S.Z.: The HST SM4 relative navigation sensor system: overview and preliminary testing results from the flight robotics lab. *J. Astronaut. Sci.* **57**(1C2), 457–483 (2009)
23. Liu, C., Hu, W.: Relative pose estimation for cylinder-shaped space crafts using single image. *IEEE Trans. Aerosp. Electron. Syst.* **50**(4), 3036–3056 (2014)
24. Du, X.D., Liang, B., Xu, W.F., et al.: Pose measurement of large non-cooperative satellite based on collaborative cameras. *Acta Astronaut.* **68**(11), 2047–2065 (2011)
25. Yu, F., He, Z., Qiao, B., et al.: Stereo-vision-based relative pose estimation for the rendezvous and docking of noncooperative satellites. *Math. Probl. Eng.* (2014)

26. Ruel, S., English, C., Anctil, M., et al.: 3DLASSO: real-time pose estimation from 3D data for autonomous satellite servicing. In: Proceedings of the 2005 International Symposium on Artificial Intelligence for Robotics and Automation in Space, Munich, Germany, September 2005
27. English, C., Zhu, S., Smith, C., et al.: Tridar: a hybrid sensor for exploiting the complementary nature of triangulation and LIDAR technologies. In: Proceedings of the 8th International Symposium on Artificial Intelligence, Robotics and Automation in Space, September 2008
28. Christian, J.A., Cryan, S.: A survey of LIDAR technology and its use in space-craft relative navigation. In: Guidance, Navigation and Control and Co-located Conference, Boston, MA, 19–22 August 2013
29. Lim, T.W.: Point cloud modelling using the homogeneous transformation for non-cooperative pose estimation. *Acta Astronaut.* **11**, 61–76 (2015)
30. Hao, G.T., Du, X.P.: Advances in optical measurement of position and pose for space non-cooperative target. *Laser Optoelectron. Prog.* **8**, 1–9 (2013)
31. Joseph, M., John, V., Matt, S., et al.: Pose measurement performance of the argon relative navigation sensor suite in simulated flight conditions, pp. 1–25. American Institute of Aeronautics and Astronautics (2012)
32. Gao, X.H., Liang, B., Le, P., et al.: A monocular structured light vision method for pose determination of large non-cooperative satellite. *Int. J. Control Autom. Syst.* **14**(6), 1535–1549 (2016)
33. Song, J.Z., Cao, C.X.: Pose self-measurement of noncooperative spacecraft based on solar panel triangle structure. *J. Robot.* (2015)
34. Peng, J.Q., Xu, W.F., Yuan, H.: An efficient pose measurement method of a space non-cooperative target based on stereo vision. *IEEE Access* **5**, 22344–22362 (2017)
35. Yu, F., He, Z., Qiao, B.: Stereo-vision-based relative pose estimation for the rendezvous and docking of noncooperative satellites. *Math. Probl. Eng.* (2015)
36. Sharma, S., DAmico, S.: Comparative assessment of techniques for initial pose estimation using monocular vision. *Acta Astronaut.* **123**(1), 435–445 (2016)
37. Jasibedzki, P., Se, S., Pan, T., et al.: Autonomous satellite rendezvous and docking using LIDAR and Model based vision. In: Proceedings of SPIE Spaceborne Sensor II, vol. 5798, June 2005
38. Sell, J.L.: Pose performance of LIDAR-based relative navigation for non-cooperative objects. Master thesis, West Virginia University (2015)
39. Ventura, J., Fleischner, A., Walter, U.: Pose tracking of a noncooperative space-craft during docking maneuvers using a time-of-flight sensor. In: AIAA Guidance, Navigation and Control Conference, AIAA Paper 2016–0875, January 2016
40. Woods, J.O., Christian, J.A.: LIDAR-based relative navigation with respect to noncooperative objects. *Acta Astronaut.* **126**, 298–311 (2016)
41. Rhodes, A., Kim, E., Christian, J.A., et al.: LIDAR-based relative navigation of non-cooperative objects using point Cloud Descriptors. In: Proceedings of the 2016 AIAA/AAS Astrodynamics Specialist Conference (2016)
42. Lichter, M.D., Dubowsky, S.: Estimation of state, shape, and inertial parameters of space objects from sequences of range images. *Proceedings of SPIE* (2003)
43. Segal, S., Carmi, A., Gurfil, P.: Stereovision-based estimation of relative dynamics between noncooperative satellites: theory and experiments. *IEEE Trans. Control Syst. Technol.* **22**(2), 568–584 (2014)
44. Zhang, L.J., Zhang, S.F., Yang, H.B., et al.: Relative attitude and position estimation for a tumbling spacecraft. *Aerospace Sci. Technol.* **42**, 97–105 (2015)

45. Sonnenburg, A., Tkocz, M., Janschek, K.: EKF-SLAM based approach for space-craft rendezvous navigation with unknown target spacecraft. In: 18th IFAC Symposium on Automatic Control in Aerospace, vol. 43, no. 15, pp. 339–344 (2010)
46. Segal, S., Gurfil, P.: Effect of kinematic rotation-translation coupling on relative spacecraft translational dynamics. *J. Guid. Control Dyn.* **32**(3), 1045–1050 (2009)
47. Dong, G.Q., Zhu, Z.H.: Autonomous robotic capture of noncooperative target by adaptive extended Kalman filter based visual servo. *Acta Astronaut.* **122**, 209–218 (2016)
48. Yu, H., Zhang, S.J., Liu, L.Y.: Relative dynamics estimation of non-cooperative spacecraft with unknown orbit elements and inertial tensor. *Chin. J. Aeronaut.* **29**(2), 479–491 (2016)
49. Pesce, V., Lavagna, M., Bevilacqua, R.: Stereovision-based pose and inertia estimation of unknown and uncooperative space objects. *Adv. Space Res.* **59**(1), 236–251 (2017)
50. Aghili, F., Parsa, K.: Motion and parameter estimation of space objects using laser-vision data. *J. Guid. Control Dyn.* **22**(2), 537–549 (2009)
51. Lefferts, E.J., Lefferts, F.L., Shuster, M.D.: Kalman filtering for spacecraft attitude estimation. *J. Guid.* **5**(5), 417–429 (1982)
52. Tweddle, B.E.: Computer vision-based localization and mapping of an unknown, uncooperative and spinning target for spacecraft proximity operations. Ph.D. thesis. Massachusetts Institute of Technology, Cambridge, MA (2013)
53. Li, Y.P.: Research on relative position and attitude estimation on non-cooperative spacecraft. Master thesis, Chinese Academy of Space Technology (2016)
54. Zhang, L.J., Qian, S., Zhang, S.F., et al.: Research on angles-only/SINS/CNS relative position and attitude determination algorithm for a tumbling spacecraft. *J. Aerosp. Eng.* **231**(2), 218–228 (2016)
55. Hao, G.T., Du, X.P., Chen, H., et al.: Scale-unambiguous relative pose estimation of space uncooperative targets based on the fusion of three-dimensional time-of-flight camera and monocular camera. *Opt. Eng.* **54**, 053112 (2015)
56. Cao, L., Chen, X.Q.: A novel unscented predictive filter for relative position and attitude estimation of satellite formation. *Acta Astronaut.* (2014)
57. Yu, X.T., Yu, F., He, Z.: Stereo vision based relative state estimation for non-cooperative spacecraft with outliers. In: Proceedings of the 33rd Chinese Control Conference, Nanjing, China, 28–30 July 2014
58. Hou, X.H., Ma, C., Wang, Z., et al.: Adaptive pose and inertial parameters estimation of free-floating tumbling space objects using dual vector quaternions. *Adv. Mech. Eng.* **9**(10), 1–17 (2017)
59. Mishra, H.: Designing robust pose estimator for noncooperative space targets for visual servoing during approach maneuvers. Master thesis, Technische Universität München (2016)
60. Xiong, K., Wei, C.L.: Adaptive iterated extended Kalman filter for relative space-craft attitude and position estimation. *Asian J. Control* **20**, 1–16 (2018)
61. Song, L., Li, Z., Ma, X.R.: A novel position and attitude estimation for space debris. *J. Astronaut.* **36**(8), 906–915 (2015)
62. Geller, D.K., Klein, I.: Angles-only navigation state observability during orbital proximity operations. *J. Guid. Control Dyn.* **37**(6), 1976–1983 (2014)



Anomaly Detection of Satellite Momentum Wheel Based on Fast Dynamic Time Warping

Pengjie Cheng^{1,2}, Jianshen Li¹, Yan Hu², Na Li², Zifeng Qiu³,
and Shuangxin Wang^{1(✉)}

¹ School of Mechanical, Electronic and Control Engineering,
Beijing Jiaotong University, Beijing 100044, China

Shxwang1@bjtu.edu.cn

² Key Laboratory of Aerospace Information Applications of CETC,
Shijiazhuang 050081, Hebei Province, China

³ The 54th Research Institute of CETC, Shijiazhuang 050081, Hebei Province, China

Abstract. As a part of the satellite attitude control system, the momentum wheel is the most prone to failure because its long time and high speed rotation. Moreover, the fault identification of the satellite in orbit is mainly determined by the threshold setting of telemetry data, so it cannot accurately reflect the actual condition of the satellite in a timely manner. Aiming at the above problems, this paper proposes to construct the pattern off-line by clustering. Then the similarity measure, which is based on Dynamic Time Warping (DTW), is used to detect the anomaly. Due to the high time cost of the DTW, the existing acceleration algorithm is analyzed and improved. So that the online system can meet the demand of real-time monitoring. Finally, through conducting simulation of the common faults of momentum wheel, and being verified by experiments, it is concluded that the accelerated algorithm can detect anomalies.

Keywords: Satellite · DTW · Pattern matching · Anomaly detection

1 Introduction

As one of the most widely used spacecraft at present, satellite may appear abnormal or malfunction due to its complex structure and uncertainties in the space environment. After classifying and analyzing the fault data of 28 Chinese satellites in orbit from 1988 to 2014, environment, design and device-related faults are the main types of on-orbit faults, and most of them can be solved by on-orbit self-diagnosis and ground operation [1]. However, if these faults cannot be dealt with in a timely and effective manner, satellites may not be able to complete the scheduled orbital mission, and they may even go out of control and eventually crash, bringing huge economic losses and negative social impact.

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 773–783, 2021.

https://doi.org/10.1007/978-981-15-8458-9_82

At present, fault identification of in-orbit satellites is mainly determined by monitoring telemetry data transmitted in real time and using simple threshold setting and expert experience. However, the simple threshold judgment cannot detect the anomalies effectively and timely, and cannot reflect the actual running state of the satellite comprehensively, accurately and objectively [2]. At the same time, satellite data has the characteristics of large scale and high latitude, and these large amounts of data contain the knowledge that can be used for satellite fault diagnosis. Therefore, based on the data of the satellite momentum wheel, this paper proposes the anomaly detection of the satellite momentum wheel based on the accelerated DTW algorithm. Abnormal detection of momentum wheel is divided into two parts: off-line mode construction and online abnormal detection. Off-line mode construction obtains the normal mode by clustering the normal data of the momentum wheel. According to the existing normal mode, online anomaly detection determines whether the current momentum wheel is abnormal by accelerating DTW algorithm.

2 Failure Analysis of Momentum Wheel

2.1 Momentum Wheel Working Principle

The momentum wheel is mainly composed of four parts: the wheel body bearing components housing component bearing component and driving motor [3]. Its structure is shown in Fig. 1 [4].

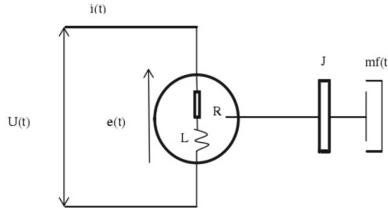


Fig. 1. Structure diagram of momentum wheel

Among them, The circle and the inner part of the circle are momentum wheel motors, the right rectangle is the moment of inertia of the momentum wheel flywheel, and the far right side is the momentum wheel output momentum. There are the following relations among the physical quantities related to the momentum wheel. In general, there are two control modes of momentum wheel, namely torque control and speed control. The two control modes have their own advantages and disadvantages. It is necessary to select the control mode according to the actual demand. There are flywheel control speed data. Therefore, it is speculated that the momentum wheel is under the speed control. It can be deduced that the output torque and the expected control torque of the momentum wheel

in the torque control mode are a proportional link with a gain of approximately 1. In the speed control mode, the control voltage of the momentum wheel is proportional to the expected angular momentum of the east two wheels, which can also be regarded as a proportional link approximately [5]. Therefore, both control modes can be regarded as linear systems, so fault data can be simulated according to the superposition of linear systems and existing normal data.

2.2 Failure Mode Analysis of Momentum Wheel

Peng et al. [6] made the statistics on the failure of each component of momentum wheel of a certain model and its similar products, and found that the failure of momentum wheel is mainly the failure of bearing components. These faults are all related to the rotation or output torque of momentum wheel flywheel. The common faults of momentum wheel can be summarized as: stuck fault, idling fault, friction increase fault, gain decrease fault and sudden change fault [7,8]. In Sect. 5, how to affect the parameters related to momentum wheel is discussed, and fault data is generated by simulation.

3 Research on Accelerating DTW Algorithm

3.1 Distance Measurement of Time Series

Distance measurement in time series is the most basic method in time series clustering, similarity search and pattern matching. The greater the distance between the two time series, the more dissimilar the two time series are, and vice versa. There are several methods to measure the distance between time series. Euclidean Distance and Dynamic Time Warping Distance (DTW) are used most widely. The expression of Euclidean Distance is shown in Eq. (1) and expression of DTW is shown in Eqs. (2)–(4).

$$ED(T_1, T_2) = \sqrt{\sum |t_{1i} - t_{2i}|^2} \quad (1)$$

$$DTW(P, Q) = \sqrt{D(n, m)} \quad (2)$$

Where $T = (t_1, t_2, \dots, t_n)$ represents a time series. t represents the value of a certain time. $P = (p_1, p_2, \dots, P_n)$ and $Q = (q_1, q_2, \dots, q_m)$ represent two series with different lengths. D is distance accumulation matrix. Its expression is shown in Eq. (3).

$$D(i, j) = \begin{cases} d(i, j) + 2D(i - 1, j - 1), \\ \min\{D(i - 1, j), D(i - 1, j - 1), D(i, j - 1)\} = D(i - 1, j - 1) \\ d(i, j) + \min\{D(i - 1, j), D(i - 1, j - 1), D(i, j - 1)\}, \\ others \end{cases} \quad (3)$$

$d(i, j)$ is shown in Eq. (4).

$$d(i, j) = \begin{cases} |p_i - q_j|^2, & 1 \leq i \leq n, 1 \leq j \leq m \\ \infty, & i = 0 \text{ or } j = 0 \end{cases} \quad (4)$$

Dynamic time warping is a global optimal algorithm, which is widely used to calculate the distance by finding a path that can make the distance between P and Q shortest. But the problem is also obvious. The time complexity of the algorithm is high, and its time complexity is $n \times m$ [9].

3.2 DTW Acceleration Algorithm

Early Abandonment. In DTW search, if a given sequence C is looking for the closest sequence in $Q = (q_1, q_2, q_3, \dots, q_n)$, we need to calculate the DTW distance between C and each sequence in Q . And then find the term with the smallest distance. The idea of early abandonment is to set a flag to store the minimum distance between C and each sequence in Q . The MATLAB pseudo code of algorithm flow can be summarized as follows.

Algorithm 1. Early abandonment of DTW

Input: C is the sequence given to search, Q is sequences in historical data, L_2 is the length of q_i

Output: (flag1, flag2), where flag1 is the matching distance, flag2 shows which historical data is matched with C .

1. Initialize flag1=inf, flag2=0

2. for($i = 1 : n$) {

$d(i) = 0$

for($j = 1 : L_2$) {

Calculate $cq(j)$, which is the distance between the j^{th} point in and the corresponding point in C

Add $cq(j)$ to $d(i)$

$d(i) = d(i) + cq(j)$

if($d(i) > \text{flag1}$) {

break

}

}

$\text{flag1} = d(i);$

$\text{flag2} = i;$

}

3. Output(flag1,flag2) pair

4. End

Set Warping Range. The main method to set the bending range indirectly is to calculate the lower bound function. Different bending range will result in different lower bound functions.

As shown in Fig. 2, the bending range represents DTW's tolerance range for lag or lead in data time, and will be dynamically aligned within the range. If the range setting is too small, the matching result will be close to the Euclidean

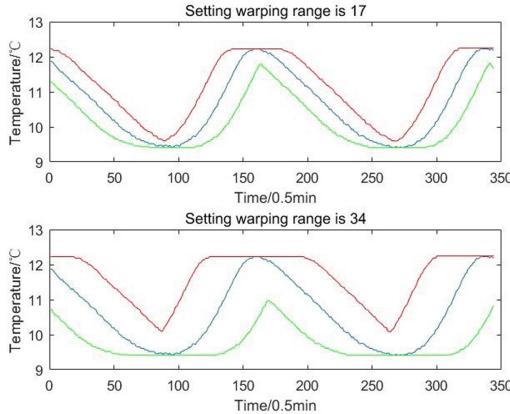


Fig. 2. Different warping range.

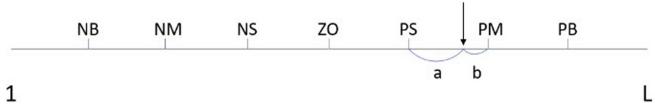
distance matching result; if the range setting is too large, the matching deviation may be large because of ignoring small fluctuations. Generally, the bending range is set at about 10% of the total length of the sequence.

3.3 Improvement of Acceleration Algorithm

Improve Early Abandonment with Fuzzy Algorithm. Rakthanmanon T, Campana B, Mueen A et al. [10] proposed to select two sequences from the data, arrange the distance between the two sequence points from large to small, and perform early abandonment in this order. However, there is contingency in only one sorting. Although the experimental results make early abandonment be executed earlier, there is contingency. Therefore, this paper proposes a fuzzy algorithm to find out the most likely location of the point with the largest distance between sequences, so as to perform early abandonment earlier and accelerate online pattern matching algorithm. Fuzzy algorithm is mainly divided into three steps: fuzziness, fuzzy reasoning and clarity.

Fuzzification. The process of fuzzification divides the time series into eight segments and seven fuzzy levels, which are NB, NM, NS, ZO, PS, PM, PB respectively. In Fig. 3, the position indicated by the arrow is the place with the maximum distance between two historical sequence points. The membership of the position to the fuzzy level PS and PM is $8b/L$ and $8a/L$ respectively.

For all historical series, the maximum distance between points is calculated, and the fuzzy membership degree is calculated. After the total sum is divided by the total calculation times, the membership degree of seven fuzzy levels is obtained. The membership values x_1, x_2, \dots, x_7 of each fuzzy level were obtained after doing fuzzification.

**Fig. 3.** Fuzzification.

Fuzzy Reasoning. Fuzzy reasoning is generally based on the fuzzy reasoning table specified in advance. The fuzzy reasoning table is made according to the actual needs. Table 1 shows the fuzzy reasoning values when different fuzzy level has the maximum membership.

Table 1. Fuzzy reasoning table

Maximum membership	Fuzzy reasoning values						
	NB	NM	NS	ZO	PS	PM	PB
NB	0	1	2	3	-3	-2	-1
NM	-1	0	1	2	3	-3	-2
NS	-2	-1	0	1	2	3	1
ZO	-3	-2	-1	0	1	2	3
PS	3	-3	-2	-1	0	1	2
PM	2	3	-3	-2	-1	0	1
PB	1	2	3	-3	-2	-1	0

According to the reasoning results shows in Table 1, and then according to the barycenter method, the defuzzification result will be near the maximum membership value.

Defuzzification. Defuzzification is the process of calculating the maximum position between points according to the membership results of fuzzification and the reasoning value of each fuzzy level. The commonly used method is the barycenter method. As shown in (5).

$$CRP = x_1NB + x_2NM + \dots + x_7PB \quad (5)$$

CRP is the position got by clarity. It's value should be between 0 to L . After the defuzzification results are obtained, the order of the accumulation process for calculating the distance between two time series during online matching is reordered, from CRP to L , and then from 0 to $CRP - 1$.

4 Abnormal Diagnosis System

4.1 Offline Mode

Data Preprocessing. In satellite data, part of the data may be missing or duplicated due to sensor problems or communication problems. Therefore, before analyzing the data, we should remove the obvious error but not the fault data. Moreover, due to the obvious periodicity of satellite data, the main frequency components can be found by Fourier transform. Then use the period as the basis of segmentation. The datapreprocessing result is shown in Fig. 4.

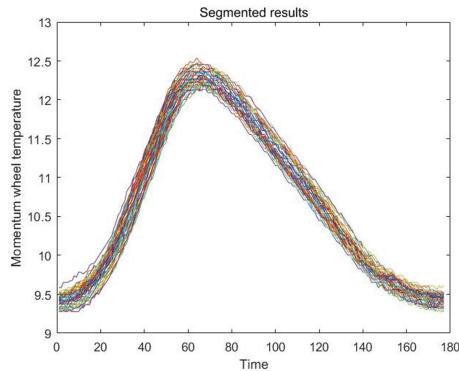


Fig. 4. Data Fourier transform results.

Clustering. The satellite historical data are classified by clustering. Then the distance between classes will be used as the basis for judging whether the current data is abnormal. The following figure shows the results of two-dimensional point set classification by randomly selecting cluster centers. It can be seen from the Fig. 5 that different initial cluster centers will get different clustering results.

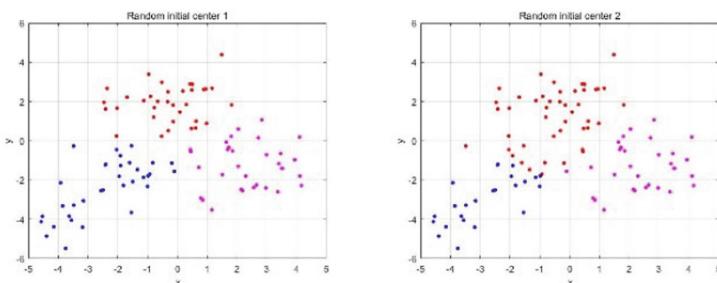


Fig. 5. Clustering results with different initial cluster center.

In order to avoid the bad clustering effect caused by some special centers, the initial clustering center is selected according to the density of a point. The point density can be defined as follows:

$$\rho_p = \frac{1}{D_i} \quad (6)$$

In Eq. (6), D_i is the distance which is i th nearest to the point. It is possible to compare the density of each point without setting the radius in advance. After the initial cluster centers are selected, the distance from each point to each center is calculated, and the point is classified into a group with its nearest center. After classification, according to the classification results, a new cluster center is obtained according to the center of gravity method, and the process of classification is repeated until the center converges or the number of iterations reaches the preset value.

4.2 Online Mode

Online anomaly detection is used to match the input time series with the historical series. The algorithm is evaluated by the algorithm time-consuming and the accuracy of data anomaly detection. By measuring the distance between sequences, the matching process finds the sequence with the smallest distance from the current sequence in the historical data as the matching result. Because DTW is better than other methods in time series distance measurement, and it can measure the distance between two time series with different length. Therefore, online matching measures the distance between two time series by using the acceleration algorithm in Sect. 3.

5 Fault Data Simulation of Momentum Wheel

The faults mentioned in Sect. 2 can lead to changes in temperature. In these faults, the temperature will change when stuck fault, idling fault, friction increase fault and sudden change fault occur. The effect on temperature is very small when gain decrease fault occur. Simulated fault data can be expressed as follows:

Step Fault

$$T_{out} = \begin{cases} T_{in}, & t < t_f, \\ T_{in}, & others \end{cases} \quad (7)$$

Stuck fault, idle fault, friction fault will lead to the decrease of momentum output of momentum wheel and the increase of internal energy when the input energy is constant. Therefore, at the time of fault occurrence, t_f , the temperature increases (step fault).

Abrupt Faults

$$T_{out} = \begin{cases} T_{in}, & t < t_f, \\ T_{in} + \Psi(t), & others \end{cases} \quad (8)$$

$\Psi(t)$ refers to a random fluctuation of temperature. A jump fault will increase or decrease the flywheel speed intermittently. Therefore, from the time of fault occurrence, t_f , the temperature will have a random fluctuation based on the normal temperature.

6 Experimental Analysis of Diagnosis System

6.1 Real Time Experiment of Pattern Matching

Pattern matching is the process of finding the most similar historical subsequence with the data to be searched online. Therefore, for each input sequence to be searched, all historical subsequences must be matched. Here, 10000 pieces of temperature data of are taken as historical data to verify the real-time performance of online pattern matching algorithm. Based on the experimental, we can calculate the time consumption of ED, DTW and DTW acceleration algorithm, is about 0.001s, 7.7s and 0.007s respectively. Therefore, the DTW algorithm after acceleration can meet the requirements of real-time.

6.2 Simulation Test Experiment of Temperature Fault of Momentum Wheel

In the off-line mode, according to the ED distance measurement and DTW distance measurement, all the historical subsequences in 10000 historical data are matched, and the average matching distance and the maximum matching distance are obtained.

Table 2. Measurement results of distance between historical data

Distance measurement	Average matching distance	Maximum matching distance
ED	5.13	18.16
DTW	1.44	10.91

As can be seen from the Table 2, the distance measured by DTW is always smaller than that measured by ED in normal historical data matching. If the matching distance is between the average distance and the maximum distance, it is considered as abnormal data that needs attention; if the matching result is greater than the maximum distance, it can be directly determined as fault data.

Test Experiment of Stuck Fault, Idle Fault and Friction Fault. According to the formula in Sect. 4, when the step temperature is 0.5–2.5°C, the matching results are shown in the Table 3.

It can be seen from the Table 3 that when a step of 0.5 °C occurs, data abnormality can be detected. With the increase of the step, the matching distance of ED and DTW increases. When the step increases to 2.5 °C, it can be determined as fault.

Table 3. Distance matching results of different step sizes

Step (°C)	0.5	0.75	1	1.5	2	2.5
ED Match distance	5.34	8.71	10.38	13.32	16.44	20.28
DTW Match distance	1.76	2.88	4.21	6.4	9	11.55

Test Experiment of Abrupt Faults. When the step temperature is 0.5–2.5 degrees, the matching results are shown in the Table 4.

Table 4. Distance matching results of different abrupt variables

Step (°C)	0.5	0.75	1	1.5
ED Match distance	7.84	10.19	12.19	20.65
DTW Match distance	3.21	4.68	5.59	9.86

It can be seen from the Table 4, the abnormality can be detected when the sudden variable reaches 0.5. When the mutation reaches 1.5, it can be determined as fault.

7 Conclusion

In this paper, the method of off-line clustering and on-line pattern matching is proposed to detect the anomaly of momentum wheel data. Because the online mode needs to meet the real-time requirements, the DTW algorithm is improved on the basis of studying the existing DTW acceleration algorithm. The real-time performance of online mode is verified by experiments. At the same time, by analyzing the common faults of momentum wheel, the fault data of momentum wheel are simulated. It is proved by experiments that when the temperature changes more than 0.5, the anomaly can be detected. This method can detect the data anomaly which can not be detected by simple threshold method.

References

- Pang, J., Liu, D., Peng, Y., Peng, X.: Intelligent pattern analysis and anomaly detection of satellite telemetry series with improved time series representation. *J. Intell. Fuzzy Syst.* **34**(6), 3785–3798 (2018)
- Liu, D., Pang, J., Song, G., Xie, W., Peng, Y., Peng, X.: Fragment anomaly detection with prediction and statistical analysis for satellite telemetry. *IEEE Access* **5**, 19269–19281 (2017). <https://doi.org/10.1109/ACCESS.2017.2754447>
- Liu, S.: Long life ground verification and evaluation of satellite flywheel products. *Navig. Control* **11**(1), 52–56 (2012)

4. Gu, P., Wang, D., Liu, C.: Reconfigurable design of zero momentum satellite wheel control system. *China Space Sci. Technol.* **33**(01), 7–14 (2013). <https://doi.org/10.3780/j.issn.1000-758X.2013.01.002>
5. Sun, D., Zhou, H., Li, S., Zhou, M.: Control moment gyroscope high precision frame control technology. *Electron. Meas. Technol.* **39**(11), 19–25 (2016). <https://doi.org/10.3969/j.issn.1002-7300.2016.11.006>
6. Liu, W., Liu, C.: Diagnostic evaluation of momentum wheel fault based on quantitative and qualitative methods. *China Space Sci. Technol.* **8**(4), 55–63 (2011). <https://doi.org/10.3780/j.issn.1000-758X.2011.04.008>
7. Chai, W., Dong, S., Chen, J., Zhang, C., Chen, T.: A fault tolerant control method for momentum wheel of remote sensing satellite. *Comput. Meas. Control* **27**(02), 79–83+88 (2019). <https://doi.org/CNKI:SUN:JZCK.0.2019-02-018>
8. Tian, K., Li, M.: Physical simulation of flywheel fault diagnosis based on adaptive observer. *Space Control Technol. Appl.* **36**(6), 27–32 (2010). <https://doi.org/10.3969/j.issn.1674-1579.2010.06.006>
9. Xie, Y., Lou, Y., Shang, Z.: An improved algorithm for improving the efficiency of DTW algorithm. *Comput. Digit. Eng.* **47**(03), 530–534 (2015). <https://doi.org/CNKI:SUN:JSSG.0.2019-03-009>
10. Rakthanmanon, T., Campana, B., Mueen, A., Batista, G., Westover, B., Zhu, Q., Zakaria, J., Keogh, E.: Searching and mining trillions of time series subsequences under dynamic time warping. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 262–270 (2012). <https://doi.org/10.1145/2339530.2339576>



Reliability Analysis of Flexible Test System for the Products of Aerospace Pyromechanical Device Based on Dynamic Bayesian Network

Jianxiong Zhou, Shanbi Wei^(✉), and Yi Chai

Chongqing University, Chongqing 400044, China
weishanbi@cqu.edu.cn

Abstract. The flexible test system for aerospace pyromechanical device products (PDPs) is an important mean to accurately assess the quality of aerospace-grade PDPs. Since aerospace PDPs are disposable products produced in small batch and multi-batch, only a limited number of samples can be tested to evaluate the quality level of the whole batch. Therefore, the reliability of ignition test system is concerned to ensure the reliability of limited test samples. The reliability of the test system is related to time and product batch task. Therefore, this paper proposes a dynamic Bayesian network method based on the combination of time and batch factors, and a test channel configuration optimization method is introduced to meet the reliability requirements of the batch change test, which provides an effective mean for evaluating the reliability of the test system. Numerical experiments show that the method can accurately reflect the reliability change of the flexible test system.

Keywords: Reliability · Real-time · Dynamic Bayesian network · Optimization

1 Introduction

The flexible test system for aerospace pyromechanical device products (PDPs) is a special test system, which is widely used to test and judge the product quality of aerospace-grade PDPs, and plays an important role in guaranteeing the quality and safety of PDPs. The credibility of the quality test results of PDPs depends on the reliability of the ignition test system. Once the test system fails, the wrong evaluation of the quality will lead to the failure of space launch mission, and even bring catastrophic consequences. Therefore, the reliability study of flexible test system for aerospace PDPs is of great significance [1].

Todays reliability assessments of test systems and similar systems are extensively studied. Eliassi et al. [2] proposed a method based on minimum cut set to extract Bayesian network (BN) structures and apply BN to the modeling, evaluation and analysis of reliability of composite power systems. Lin, Qi et al. [3]

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 784–792, 2021.

https://doi.org/10.1007/978-981-15-8458-9_83

discussed the most common reliability measurement methods, including reliability function and failure rate, and adopted two important reliability analysis methods: failure mode, effects and criticality analysis (FMECA) and fault tree analysis (FTA).

Due to a wide variety of aerospace PDPs, which belongs to the small batch and multi-batch of production mode. The reliability of the test system depends on the frequency of use and the time of use so the above method considering only the time factor is not applicable to the reliability evaluation of the test system of aerospace PDPs. Besides, due to the high reliability requirements of PDPs, the test system should not only meet the test requirements but also ensure that the reliability of the test system is extremely high and reaches the optimal level. Therefore, only by optimizing the device configuration selection in the test system and making the reliability of each device reach the local optimal, can the overall reliability of the system reach the optimal. The realization of the optimal reliability of the test system is mainly realized by the matrix switch, so the system needs to optimize the matrix switch. Many processing solutions for optimization problems have applied to reliability optimization configurations. Heungseob Kim [4] discusses the best reliability assessment for heterogeneous component systems, which are redundant allocation problems with mixed components (RAPMC). Ouyang et al. [5] used the improved particle swarm optimization algorithm (PSO) with random perturbation property to solve the reliability redundancy allocation problem (RRAP).

In order to fully guarantee the above requirements of the ignition test process, this paper makes the following contributions:

- 1) The batch change is introduced into the dynamic Bayesian network, and the reliability calculation process considers the change of time and batch simultaneously to adapts to the test switching of different batches of products.
- 2) For the reliability requirements and characteristics of batch product test, reliability prediction and the reliability before and after the impact of product test in the batch are introduced into reliability optimal configuration problem of channels, so as to ensure the reliability of the whole batch test process.

2 Flexible Test System of PDPs

PDPs are disposable products that filled with gunpowder and stimulated by the outside to detonate explosives or perform mechanical functions. Whether PDPs work normally directly affects the success of rocket launch, missile flight and other major missions. And as an important mean of quality inspection of PDPs, the flexible test system of aerospace PDPs is widely used in the quality test of various PDPs batches, the credibility of the PDPs quality test results is closely related to the reliability of the test system, so the reliability evaluation of the test system is an important way to guarantee the zero defects and absolutely reliable of the products. As shown in Fig. 1, the low resistance tester (LRT) is connected to the ignition circuit to obtain the resistance value of the product, and the current and voltage of the DC regulated power supply (DC) are set to

ignite and detonate by the detonation tester (DT). Photoelectric amplifiers (PA), target boxes (TB) and dynamic resistance strain gauges (DRSG) are connected to a special sealed explosion instrument to measure the time and thrust/pressure during the explosion, and the outputs are set to be connected to the input channels of the matrix switch (MS), which is configured with test channels. The measured signal is transmitted to the data acquisition instrument (DAI) through the circuit converter and displayed on the industrial personal computer (IPC).

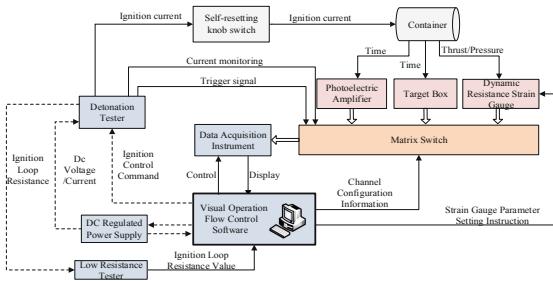


Fig. 1. Test system structure diagram of PDPs

3 DBN of Ignition Test System

3.1 Reliability Model of the Ignition Test System

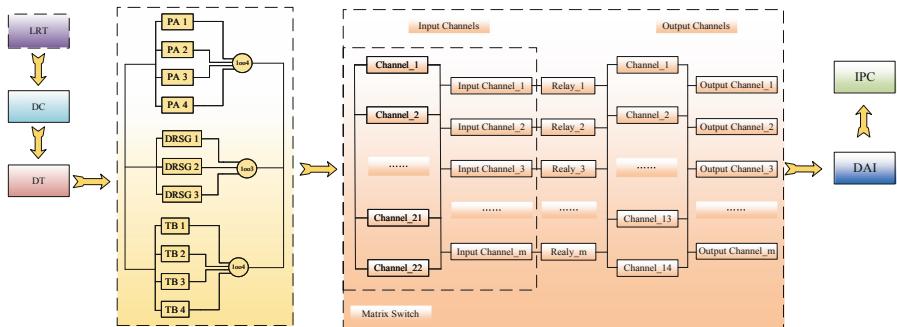


Fig. 2. Reliability block diagram of PDPs

As shown in Fig. 2, the whole system structure is a series structure. The DRSG subsystem, the PA subsystem and the TB subsystem are parallel systems, in which the selection of the channels depends on the specific product test

requirements, including the modes of multi-choice one and multi-choice two and so on. The test system selects the specific instruments of the different measuring instrument subsystems according to the characteristics of the product, in case of damage or failure of the measuring instruments, the system will automatically select the best measuring instruments, and the selected instruments are connected to the input channels of the MS subsystem. The IPC subsystem, the LRT subsystem, the DC subsystem and the DAI subsystem are single systems. The MS subsystem is a redundant system, which configures the output channels according to the closure of the relay controlled by the input channels of the measuring instruments, and the fixed two input channels are connected with the external trigger and current interface respectively. Due to the complexity of the channel selection mode, the reliability evaluation is more complicated.

The overall reliability of the ignition test is:

$$P = P_{LRT} * P_{DC} * P_{DT} * P_{DRSG} * P_{PA} * P_{TB} * P_{MS} * P_{DAI} * P_{IPC} \quad (1)$$

The DRSG has three channels. During each test, one or more channels are selected according to the test requirements of the product, so the reliability can be further expressed as:

$$\begin{aligned} P_{DRSG} &= P_{DRSG1} * P_{DRSG2} * P_{DRSG3} \\ P_{DRSGi} &= \begin{cases} P_{DRSGi} & \delta_{DRSGi} = 1 \\ 1 & \delta_{DRSGi} = 0 \end{cases}, i = 1, 2, 3 \end{aligned} \quad (2)$$

Where P_{DRSGi} is the reliability probability of the measuring instrument i , and $\delta_{DRSGi} \in [0, 1]$ is the component selection mark. Similar components include the PA and the TB.

The matrix switch contains input channels, relay switches and output channels. There are 22 input channels and 14 output channels, the reliability of the input channel m is expressed as $P_{MS}^{in,m}$, the reliability of the output channel n is expressed as $P_{MS}^{out,n}$, and the reliability of relay switch connected by channel m and output channel n is expressed as $P_{MS}^{relay,nm}$. The reliability of matrix switch is:

$$\begin{aligned} P_{MS} &= \sum_{m=1}^{22} \sum_{n=1}^{14} P_{MS}^{nm} \\ P_{MS}^{nm} &= \delta_m * P_{MS}^{in,m} * P_{MS}^{relay,nm} * \delta_n * P_{MS}^{out,n}, \delta_m \in [0, 1], \delta_n \in [0, 1] \\ \begin{bmatrix} \delta_{11} & \dots & \delta_{1m} \\ \vdots & \ddots & \vdots \\ \delta_{n1} & \dots & \delta_{nm} \end{bmatrix} &\in \Phi_{nm}, [\delta_1 \dots \delta_m] \in \{\Phi_{DRSG}, \Phi_{PA}, \Phi_{TB}\} \end{aligned} \quad (3)$$

Where Φ_{nm} represents the set of all relays located inside the matrix switch, Φ_{DRSG} , Φ_{PA} , Φ_{TB} represent the channel selection sets of the target box, the photoelectric amplifier and the dynamic resistance strain gauge, respectively.

3.2 DBN Parameters of Ignition Test System

In the above nine subsystems, the corresponding DBN model is developed under the condition of time-driven components and batch-driven components [6].

DBN Modeling Based on Time-Driven Components. The DBN model is considered in time for components other than TB, PA, DRSG and MS. The reliability of time-driven components is only related to time, and its time-varying process satisfies Markov property and Stationarity, meeting the assumptions of traditional dynamic Bayesian networks [7]. The failure rate and repair rate of components are represented by λ and μ respectively, assuming the current time is t and the time interval between two time slices is Δt , the node transition probability between two time slices is given as:

$$\begin{cases} P(X_i(t + \Delta t) = \text{Normal} | X_i(t) = \text{Normal}) = e^{-\mu \Delta t} \\ P(X_i(t + \Delta t) = \text{Fault} | X_i(t) = \text{Normal}) = 1 - e^{-\mu \Delta t} \\ P(X_i(t + \Delta t) = \text{Normal} | X_i(t) = \text{Fault}) = 1 - e^{-\lambda \Delta t} \\ P(X_i(t + \Delta t) = \text{Fault} | X_i(t) = \text{Fault}) = e^{-\lambda \Delta t} \end{cases} \quad (4)$$

Where ‘Normal’ represents the condition in which the component works properly, and ‘Fault’ represents the condition in which the component fails. As shown in Table 1, the posterior probability of each subsystem at any given time can be calculated and the influence of failure rate of individual components on system reliability can be evaluated.

Table 1. Failure rate and Repair rate of components for test system

Component	Failure rate	Repair rate
Low Resistance Tester	$2.305e^{-7}$	$1.097e^{-7}$
DC Regulated Power Supply	$2.745e^{-7}$	$1.427e^{-7}$
Detonation Tester	$1.646e^{-7}$	$1.789e^{-8}$
Data Acquisition Instrument	$1.976e^{-7}$	$2.386e^{-8}$
Industrial Personal Computer	$2.195e^{-7}$	$1.317e^{-7}$

DBN Modeling Based on Batch-Driven Components. The DBN model is considered in batch mode for components of TB, PA, DRSG and MS. As the matrix switch is a 22*14 redundant subsystem, it is very complicated in Bayesian network modeling, and the modeling effect is not very ideal, so the reliability analysis is carried out by constructing a single module model. On the basis of obtaining the field test data of multiple batches of sample products in the actual production process of a certain PDPs plant, by constructing histogram to solve the average and standard deviation of sample data, the Bernoulli distribution

model with the change of conditional transition probability and usage times is obtained.

$$\begin{cases} P(X(N + \Delta k) = \text{Normal}|X(N) = \text{Normal}) = P_i^{\Delta k} \\ P(X(N + \Delta k) = \text{Fault}|X(N) = \text{Normal}) = 1 - P_i^{\Delta k} \\ P(X(N + \Delta k) = \text{Normal}|X(N) = \text{Fault}) = 1 - P_j^{\Delta k} \\ P(X(N + \Delta k) = \text{Fault}|X(N) = \text{Fault}) = P_j^{\Delta k} \end{cases} \quad (5)$$

where Δk represents the interval of use times, P_i represents the probability of success of the next batch under the premise of success of the previous batch, P_j represents the probability of failure of the next batch under the premise of failure of the previous batch, which value are shown in Table 2.

Table 2. Bernoulli distribution

Component	$P_i(\%)$	$P_j(\%)$
Target Box	99.848	99.782
Dynamic Resistance Strain Gauge	99.852	99.731
Photoelectric Amplifier	99.856	99.819
Matrix Switch	99.864	99.722

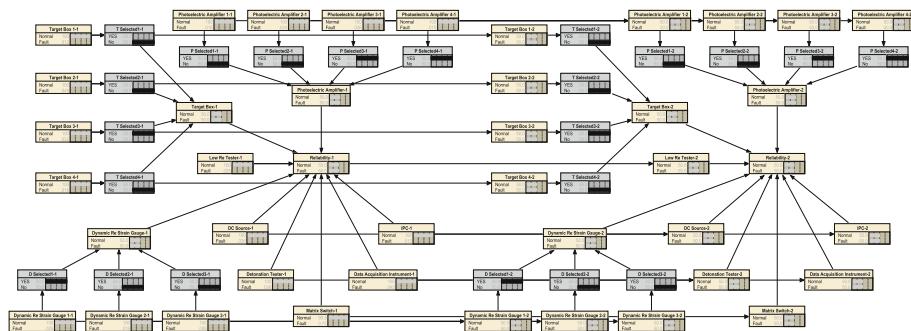


Fig. 3. DBN structure diagram of PDPs test system

DBN Model of Ignition Test System. The test system for aerospace PDPs is very complicated. According to the reliability diagram of the test system shown in Fig. 2, the Bayesian network model of the aerospace PDPs test system is shown in Fig. 3. The LRT subsystem, the DC subsystem, the DT subsystem, the DAI subsystem, and the IPC subsystem each contain only one component,

the TB subsystem, the PA subsystem and the DRSG subsystem form a system of the modes of multi-choice one and multi-choice two and so on, the MS subsystem establishes a model through a single module. Any failure of the nine nodes will result in a failure of the test system.

4 Optimal Selection and Configuration of Matrix Switches

During the test, the matrix switch needs to configure the output channels according to the input channels of the measuring instruments, so as to transmit the data acquired to DAI. Therefore, the optimal selection is based on the matrix switch to achieve the optimal reliability of the test system. The starting time of ignition test for a batch of products is set as t_s , and the test quantity for this batch of products is N_s . The reliability of the test system for this batch before the last product test is $P_{MS}(N_s)$. In order to meet the requirements of the overall test of the system, $P_{MS}(N_s)$ is satisfied:

$$\begin{aligned} P_{MS}(N_s - 1) &\geq P_{MS}^{min} \\ P_{MS}^{min} &= P^{min}/P_{LRT}(t_s) * P_{DC}(t_s) * P_{DT}(t_s) * P_{PA}(t_s) * P_{TB}(t_s) \\ &\quad * P_{DRSG}(t_s) * P_{DAI}(t_s) * P_{IPC}(t_s) \end{aligned} \quad (6)$$

The lower limit of the reliability requirement of the test system is P^{min} . The reliability of other components at the start of the ignition test is $P_{LRT}(t_s)$, $P_{DC}(t_s)$, $P_{DT}(t_s)$, $P_{DRSG}(t_s)$, $P_{PA}(t_s)$, $P_{TB}(t_s)$, $P_{MS}(t_s)$, $P_{DAI}(t_s)$ and $P_{IPC}(t_s)$ respectively, which changes dynamically with time and batch. The matrix switch input channels are determined by $[\delta_1, \dots, \delta_{22}]$ before the test. The channels constraint of the matrix switch is:

$$\begin{cases} \sum_{m=2k-1}^{22} \delta_m * \delta_{m+1} = 0, k \in \{1, 2, \dots, 11\} \\ \sum_{m=1}^{22} \delta_m * \delta_{m+1} \leq 1, n \in \{1, 2, \dots, 14\} \end{cases} \quad (7)$$

This constraint condition not only ensures that the input channel of a single component is unique, but also ensures that the input channels and the output channels are matched one by one, without sharing the output channels; Considering the reliability maximization of $P_{MS}(N_s)$, the optimization problem is constructed as follows:

$$\begin{aligned} J &= \max_{\Phi_{MS}} P_{MS}(N_s - 1) \\ s.t. (6)(7)(9)(10) \end{aligned} \quad (8)$$

$$P_{MS}(N_s - 1) = \sum_{m=1}^{22} \sum_{n=1}^{14} P_{MS}^{nm}(N_s - 1) \quad (9)$$

$$\begin{aligned}
P_{MS}^{nm}(N_s - 1) &= \delta_m * P_{MS}^{in,m}(N_s - 1) * P_{MS}^{relay,nm}(N_s - 1) * \delta_n * P_{MS}^{out,n}(N_s - 1) \\
&\quad \delta_m \in [0, 1], \delta_n \in [0, 1] \\
P_{MS}^{in,m}(N_s - 1) &= P_{MS,o}^{in,m} * \Delta P_{MS,o}^{in,m}(N_s - 1) \\
P_{MS}^{relay,nm}(N_s - 1) &= P_{MS,o}^{relay,nm} * \Delta P_{MS}^{relay,nm}(N_s - 1) \\
P_{MS}^{out,n}(N_s - 1) &= P_{MS,o}^{out,n} * \Delta P_{MS,o}^{out,n}(N_s - 1) \\
\begin{bmatrix} \delta_{11} & \cdots & \delta_{1m} \\ \vdots & \ddots & \vdots \\ \delta_{n1} & \cdots & \delta_{nm} \end{bmatrix} &\in \Phi_{nm}, [\delta_1, \dots, \delta_m] \in \{\Phi_{DRSG}, \Phi_{TB}, \Phi_{PA}\}
\end{aligned} \tag{10}$$

Where, $\Delta P_{MS,o}^{in,m}(N_s - 1)$, $\Delta P_{MS}^{relay,nm}(N_s - 1)$ and $\Delta P_{MS,o}^{out,n}(N_s - 1)$ are the transition probabilities of the matrix switch respectively.

5 Numerical Experiments

In the history records of multiple batches test samples of a certain PDPs plant, by integrating the type data of product samples, twenty batches of five different products were selected for numerical experiments. Through the matrix switch optimization algorithm calculation, the input channels of current product are 1,11,19,22, and the output channels are 2,5,8,10. Therefore, the connection of test channels is realized by closing the corresponding relay switches. After 20 batches of tests, the reliability change [8] of the test system is shown in Fig. 4.

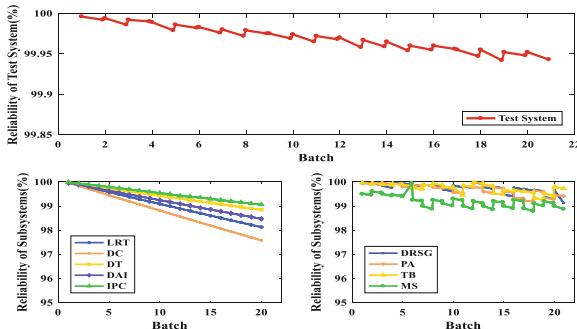


Fig. 4. DBN structure diagram of PDPs test system

In the dynamic Bayesian network considering time and batch, the reliability of time-driven components will gradually decrease with time, and the reliability of batch driven components will change with the number of use, with the increase of the use times of measuring instruments, the reliability of some batches of product testing has a certain improvement. In addition, in the process of reliability evaluation of the test system, with the change of time and product batch, the

reliability of the test system decreased from 99.996% to 99.952%, showing a downward trend as a whole, which has a high adaptability to the change of product batch. The whole test system is still at a high reliability level, meeting the needs of reliability evaluation of the test system. According to the reliability change of the test system, the operator can monitor whether the reliability of the test system meets the test requirements in real time, adjust and replace the measuring elements in time to avoid the unreliability of the test results.

6 Conclusion

In this paper, a new method based on time and batch variation dynamic Bayesian network is introduced to evaluate the reliability change of the aerospace PDPs flexible test system in the ignition test. This study shows how dynamic Bayesian network based on time and batch changes can be effectively used to analyze the level of reliability of test systems. unreliable components. In addition, a matrix switch channel optimization algorithm is proposed to ensure the optimization of the reliability of test system.

References

1. Guo, J., Li, Z.J., Keyser, T.: A Bayesian approach for integrating multilevel priors and data for aerospace system reliability assessment. *Chin. J. Aeronaut.* **31**(1), 41–53 (2018)
2. Eliassi, M., et al.: Application of Bayesian networks in composite power system reliability assessment and reliability-based analysis. *IET Gener. Transm. Distrib.* **9**(13), 1755–1764 (2015)
3. Lin, Q., Chen, Y.: FTA-based reliability measurement of test-bed system. *Adv. Mater. Res.* **791–793**, 922–925 (2013)
4. Kim, H.: Maximization of system reliability with the consideration of component sequencing. *Reliab. Eng. Syst. Saf.* **170**, 64–72 (2018)
5. Ouyang, Z.Y., et al.: An improved particle swarm optimization algorithm for reliability-redundancy allocation problem with mixed redundancy strategy and two heterogeneous components. *Reliab. Eng. Syst. Saf.* **181**, 62–74 (2019)
6. Chen, J., et al.: Risk analysis for real-time flood control operation of a multi-reservoir system using a dynamic Bayesian network. *Environ. Model Softw.* **111**, 409–420 (2019)
7. Rebello, S., Yu, H., Ma, L.: An integrated approach for system functional reliability assessment using dynamic Bayesian network and hidden Markov model. *Reliab. Eng. Syst. Saf.* **180**, 124–135 (2018)
8. Amin, M.T., Khan, F., Imtiaz, S.: Dynamic availability assessment of safety critical systems using a dynamic Bayesian network. *Reliab. Eng. Syst. Saf.* **178**, 108–117 (2018)



Research and Development of Vision System of SMT Machine

Xujie Hou^(✉), Yonghui Xie, and Lei Pan

Weifang College of Mechanical and Electrical Engineering, Weifang, Shandong, China
1016060817@qq.com

Abstract. As surface mount components are used more and more widely, the placement machines that can accurately and efficiently complete the placement tasks of chip components are becoming increasingly prominent in the electronics manufacturing industry. As the guarantee of mounting accuracy, vision system plays an important role in improving the performance and grade of SMT. In this paper, based on the needs of the placement machine for the identification of error factors such as the position and declination of electronic components, based on the research and analysis of the development status of the placement machine and its visual system at home and abroad, a new type of placement machine visual recognition is developed system.

Keywords: SMT · Vision system · FPGA · SOPC system · Digital image processing

1 Introduction

With the miniaturization of electronic products, surface mount devices are more and more widely used [1]. As an important production equipment, SMT machine is becoming more and more important in electronic manufacturing, which makes the welding and installation of SMT components more accurate and efficient [2]. At present, the development of automatic placement technology in China is not mature. As one of the key technologies of placement machine, vision system technology has a very important impact on the placement ability and accuracy of placement machine [3]. Therefore, the research on the vision system of Mounter has important theoretical significance and extensive use value.

2 Overall Design Scheme of the System

The vision system of mounter is a real-time control chip with micro control chip as the core, including camera, digital image processing and other structures. The camera converts the light signal of the target object into the analog electrical signal in the given visual range, then converts it into the digital image signal

through A/D, and finally collects the image by the computer. The digital image is divided into a certain number of pixels, each of which represents the average light intensity of a given unit. The computer processing module processes the obtained image, extracts the effective information, and sends it to the main control computer to control the operation of the mounter. Image acquisition and image processing are carried out separately, which will inevitably involve the situation of multi computer communication, and the speed will also decline. The biggest characteristic of FPGA is that it can run in parallel and embed in the kernel. So this paper adopts the scheme of FPGA. Image acquisition and processing are in the same chip, which avoids the problem of multi computer communication. In the overall design of the system, the EP3C16Q240C8 chip of Altera company is used as the main control chip of the system, and the image acquisition and processing system is constructed with OV7670 digital CMOS image sensor and HY57V641620 external SDRAM memory chip of HYUNGAU company. The display mode adopts black gold DB2C8TFT32 LCD screen, which can also be transmitted to the upper computer through serial port for processing and display. The overall block diagram of the system is shown in Fig. 1. The

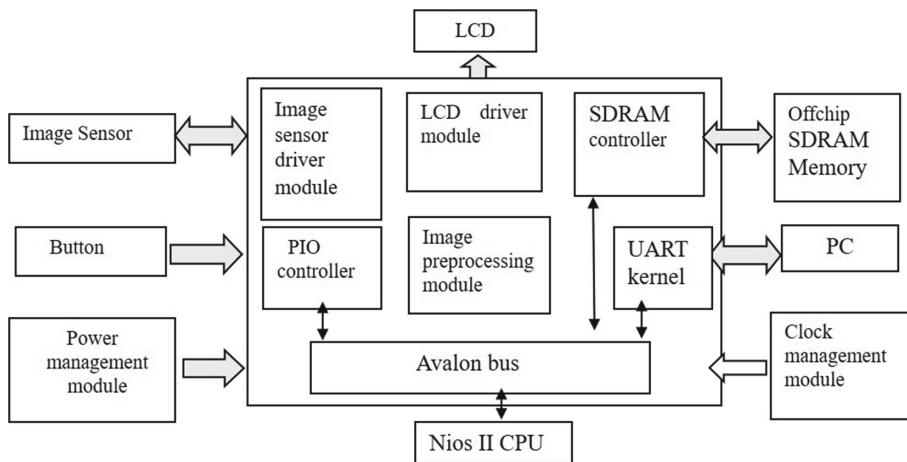


Fig. 1. The overall designer diagram of the system

main function modules of the system are power module, clock module, image acquisition module, image preprocessing module, image display module, memory module and serial port transmission. The working process of the system is as follows: After the start of the system, the FPGA first configures the function of the sensor, so that the sensor can work normally as required. When the configuration is completed, the sensor will return the configuration completion signal, and collect and output the image under the driving timing. There are two cache dual port RAM in the FPGA. The first line of the picture first enters one of the RAM. After the first line is finished, the RAM is full, and then another RAM is used

to receive data. The full RAM sends data to the SDRAM for temporary storage or directly enters the reality module under the control of the kernel. During pre-processing, the image is read out from SDRAM and transformed into a format that is conducive to image processing and then entered into the preprocessing module for processing.

3 Key Hardware Design of the System

3.1 Image Sensor OV7670

In the image acquisition module, OV7670 digital CMOS image sensor is used as the image acquisition device. This sensor is not only small in size, but also low in working voltage, including all functions provided by VGA camera. Through the SCCB bus, the sensor can be controlled to obtain some configuration data, such as the whole frame, window size, sub sampling and resolution data. The product can transmit up to 30 VGA images per second. Users can control sensors through SCCB bus, such as image quality, output data lattice and transmission mode.

3.2 Design of SCCB Bus Control Module

Design Principle of SCCB Bus Controller. Because the timing start signal and end signal of SCCB are realized by the cooperation of data line level and clock line level, and the data is transmitted at the rising edge of clock line. In order to meet the above requirements, the following design method is adopted. The clock and data are regarded as a group of output data. As long as these two groups of signals are serially output under the driving of qualified clock signals, the timing meeting the requirements of SCCB bus can be obtained at the output end.

SCCB Bus Controller Design. The interface module of SCCB bus controller is shown in Fig. 2.

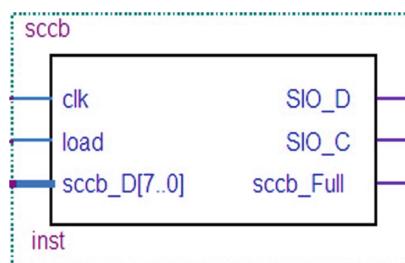


Fig. 2. The interface module of SCCB bus controller

The figure above is the schematic diagram of SCCB bus controller interface. CLK is the driver signal of the controller, load is the signal that the kernel sends data, SCCB_Full is controller busy signal, SIO_D and SIO_C is the output port of two groups of serial output data driven by clock CLK, SCCB_D [7:0] is the parallel data input.

3.3 Design of Image Acquisition Module.

The main function of the image acquisition module is to drive the image, and store the image information collected by the sensor in a certain format to the external SDRAM chip. The structure diagram of image acquisition and storage module is shown in Fig. 3.

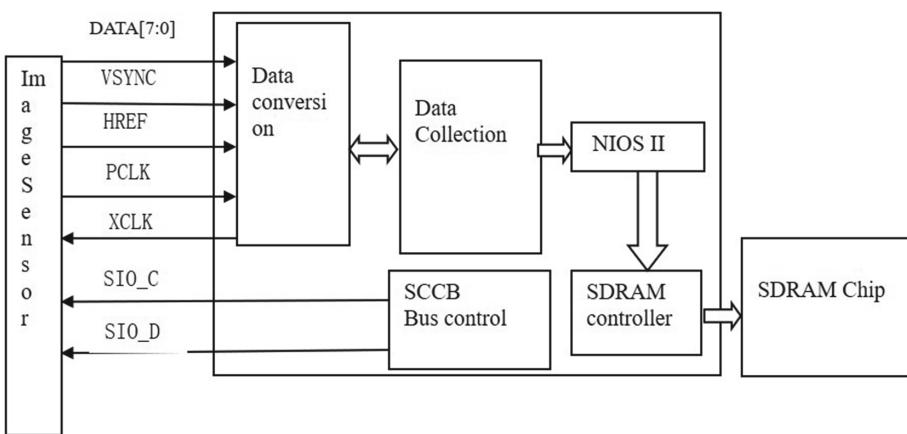


Fig. 3. The structure of image collect and storage module

The image sensor OV7670 is driven by FPGA, and its configuration is realized by SCCB bus module. SCCB module sends configuration information to image sensor, including image output format, row and column size, RGB gain, etc. After the configuration of image sensor is completed, the collected digital image information is returned to FPGA. Because it outputs digital information, it does not need AD conversion and can be directly input to FPGA.

3.4 Design of Image Display Module

Initialization of LCD Screen. Initialization is completed at the beginning of the system operation, mainly including the power setting of the screen, display size, display mode, backlight setting and the setting of the accumulated address AC. Since the data enters the display module directly without going through the kernel, it is necessary to set the AC accumulator to automatic accumulation

mode. We only need to display the data on the display, that is, we only need to design the data writing module without reading the data. The D/C port only needs to change the high and low levels of the port when initializing the register. After the initialization is completed, the port needs to be kept at a high level.

Display Control Module Design. It can be known from the timing of writing data that the data is written to the video memory at the rising edge of CS. During this period D/C and RD are kept high, and WR is kept low. Then when continuously writing data, as long as the data pulse signal is connected to CS, you can continuously write data to the video memory. Since the CS pin is also required during initialization, the two clock signals from the NiosII core and external hardware modules are separated from each other during design so that they do not affect each other. The design method is shown in Fig. 4 below:

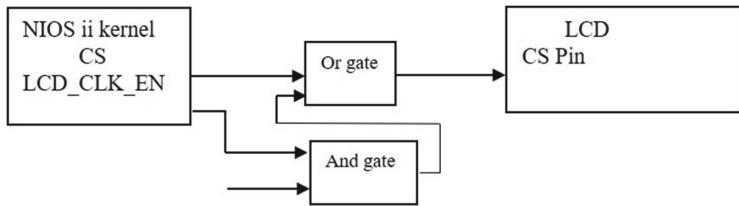


Fig. 4. The control method of CS pin

It can be seen from the figure that when the LCD is initialized, the pin LCD_CLK_EN is set to 0, then the LCD pin is only controlled by the NiosII core. When the initialization is complete and data arrives, then the NiosII core sets CS to 0 and sets LCD_CLK_EN to 1, so that the LCD CS pin is connected to the display drive clock, and the data will be written to the video memory when the clock falls. In the data acquisition and storage module, the data enters the dual-port RAM under the influence of pclk among the pixels. When the RAM is full, the NiosII core will be notified. The core only needs to write the corresponding program to implement the above process, and the collected data can be sent to the LCD display memory.

4 System Software Design

4.1 The Overall Design of the System Software

The hardware module driver is used to complete the collection, storage and display of images. In order to connect each module, it forms a whole system. Use the SOPU Builder tool to build the Nios II soft core system in the FPGA chip, which is mainly used to complete the configuration of the FPGA and the

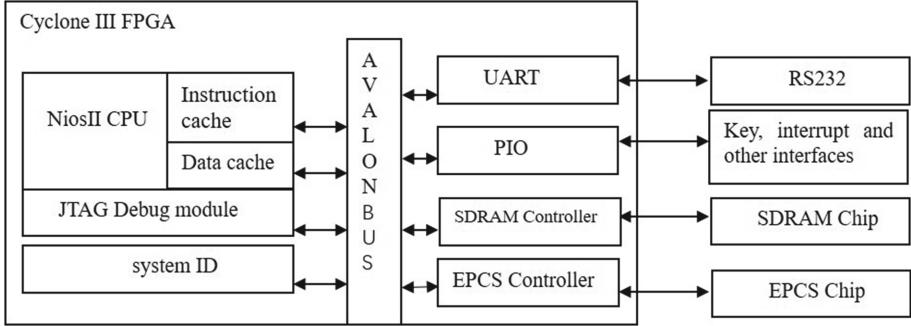


Fig. 5. The structure diagram of software of the system

read and write functions of the SDRAM off-chip memory. The overall scheme block diagram of the system software is shown in Fig. 5.

Each different module in the NiosII system has different tasks and functions. These modules are connected together through the Avalon bus, so that the system can be carried out in order. The roles and functions of the various modules in Fig. 5 are summarized as follows [4]: 1) The Nios II CPU is the core part of the SOPC system. It not only controls the operation of off-chip devices, but also executes the corresponding processing commands. When building the CPU, you must first set the instruction and data cache. 2) JTAG module is used to download logic hardware and upper layer software, and debug the system. 3) The system ID can be regarded as an ID of the system. When this module is added, the system has a unique identity, and the low-level hardware can be designed based on this ID to avoid errors. 4) The EPICS controller is used to control the off-chip EPICS chip, and later used to store downloaded programs and data. 5) The SDRAM interface is the control module of the external SDRAM chip. Due to the large storage capacity and low price of the SDRAM chip, the addition of the external chip prevents the RAM memory block inside the FPGA from being insufficient to run a larger system or store pending pictures. But SDRAM control timing is very difficult, so adding this module not only reduces the complexity of the design, but also achieves the purpose of increasing the storage unit. 6) UART is a serial port module, which is used to connect NIOS II CPU and PC, but it is only required that the level conversion chip RS232 is needed. 7) General I/O port is used to control IO devices, buttons or external interrupt sources, etc.

4.2 Flow Chart of SOPC System

As shown in Fig. 6, at the beginning, the entire system first initializes the sensors OV7670 and LCD, such as image resolution and scanning mode configuration information, so that the image collected by OV7670 can be displayed on the LCD normally. After the initialization is completed, the sensor starts to work. When a new frame of data starts, the sensor outputs signal VSYNC and generates an interrupt. The kernel sets the initial LCD address to zero and starts a new round of scanning. In order to carry out targeted image preprocessing, a certain frame of image needs to be stored in an external SDRAM chip, and the key enables the process. It can be known from the image acquisition module that the number of data stored in the dual-port ram is exactly one line of data, that is, 320, and the number of data is checked to determine whether the transmission of one line of data is completed. Finally, when the VSYNC signal arrives for the second time, it indicates the end of a frame of data, and the program returns to the beginning.

4.3 Construction of SOPC System

The construction of the Nios II soft core is shown in Fig. 7.

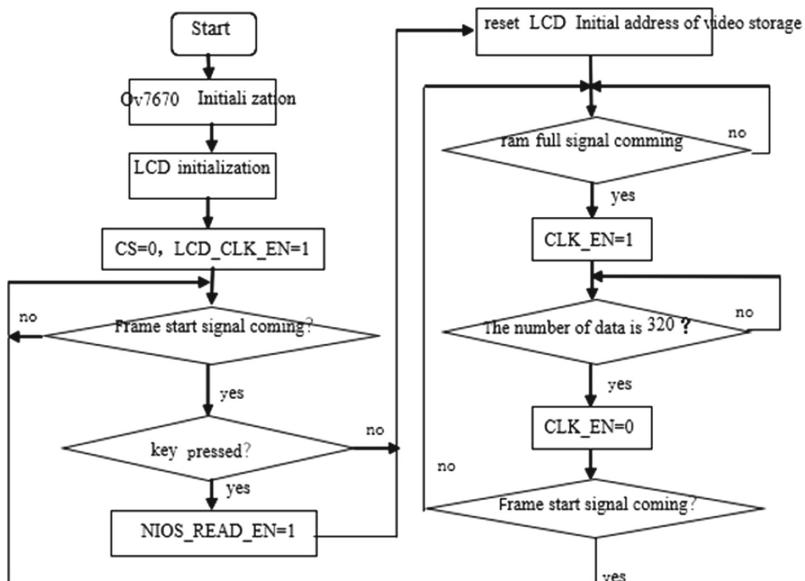


Fig. 6. The flow chart of system software

Use	C...	Module Name	Description	Clock	Base	End	Tags	IRQ
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	cpu	Nios II Processor	clk_0	0x00001000	0x00001fff		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	sdram	SDRAM Controller	clk_0	0x02000000	0x02ffffff		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	epcs_flash_controller	EPICS Serial Flash Controller	clk_0	0x00000000	0x000007ff		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	sysid	System ID Peripheral	clk_0	0x00001930	0x00001937		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	uart	UART (RS-232 Serial Port)	clk_0	0x00001800	0x0000181f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	sccb_control_clk_en	PIO (Parallel I/O)	clk_0	0x00001820	0x0000182f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	sccb_control_load	PIO (Parallel I/O)	clk_0	0x00001830	0x0000183f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	sccb_data	PIO (Parallel I/O)	clk_0	0x00001840	0x0000184f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	sccb_full	PIO (Parallel I/O)	clk_0	0x00001850	0x0000185f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	ov7670_vsync	PIO (Parallel I/O)	clk_0	0x00001860	0x0000186f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	pix_data	PIO (Parallel I/O)	clk_0	0x00001870	0x0000187f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	ram1_full	PIO (Parallel I/O)	clk_0	0x00001880	0x0000188f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	ram2_full	PIO (Parallel I/O)	clk_0	0x00001890	0x0000189f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	nios_read_clk_en	PIO (Parallel I/O)	clk_0	0x000018a0	0x000018af		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	data_clk_en	PIO (Parallel I/O)	clk_0	0x000018b0	0x000018bf		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	lcd_clk_en	PIO (Parallel I/O)	clk_0	0x000018c0	0x000018cf		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	nios_read_clk	PIO (Parallel I/O)	clk_0	0x000018d0	0x000018df		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	lcd_DC	PIO (Parallel I/O)	clk_0	0x000018e0	0x000018ef		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	lcd_CS	PIO (Parallel I/O)	clk_0	0x000018f0	0x000018ff		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	lcd_wr	PIO (Parallel I/O)	clk_0	0x00001900	0x0000190f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	lcd_rd	PIO (Parallel I/O)	clk_0	0x00001910	0x0000191f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	lcd_data	PIO (Parallel I/O)	clk_0	0x00001920	0x0000192f		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	key	PIO (Parallel I/O)	clk_0	0x00001930	0x0000193f		

Fig. 7. The structure of Nios II soft core

5 Conclusion

Based on the research and analysis of the development of placement machine and its vision system at home and abroad, Nios II system is applied to the vision system of Mounter, and a design scheme of vision system based on FPGA is proposed, including the selection of camera, the planning, design and implementation of hardware and software scheme of vision system, and the hardware implementation of necessary image processing algorithm in the process of image acquisition and processing in the logic chip is completed.

References

- Chen, K.: Research on operating system configuration method based on SOPC technology. Beijing University of technology (2010)
- Wang, H., Yang, X.: Design of high speed image acquisition system based on FPGA. In: Application of Single Chip Microcomputer and Embedded System, vol. 3, pp. 28–30 (2009)
- Xie, Y.: Research on the positioning algorithm of the machine vision system of the high precision automatic placement machine. Opt. Technol. 5, 449–451 (2008)
- Gong, J.: Image processing based on FPGA. Master's thesis of Wuhan University of science and technology, 4 (2007)



Air Pollutants NO₂ Concentration Prediction Based on LSTM Neural Network method

Jihan Li¹, Xiaoli Li^{1,2(✉)}, Jian Liu¹, and Kang Wang¹

¹ Faculty of Information Technology, Beijing University of Technology,
Beijing 100124, China
lxxiaolibjut@bjut.edu.cn

² Beijing Key Laboratory of Computational Intelligence and Intelligent System,
Beijing Advanced Innovation Center for Future Internet Technology,
Engineering Research Center of Digital Community, Ministry of Education,
Beijing 100124, China

Abstract. In recent years, the China's economy has developed rapidly. The people's living standard has been improved. The number of cars has been increasing, and the pollutant NO has been produced continuously, which leads to the formation of NO₂. These harmful particles have an impact on human health. Thus, the effective and accurate NO₂ concentration prediction model plays an effective role in people's health and prevention. For this problem, this paper presents a prediction model based on the long short-term memory (LSTM) method to predict NO₂ concentration. Firstly, the PM₁₀, SO₂, NO₂, CO, O₃, temperature in a campus monitoring point in Beijing is collected as the research object in this paper. Then, the LSTM prediction model and BP (back propagation) neural network prediction model are established respectively. Finally, the accuracy of the two prediction models for the prediction of NO₂ concentration is compared. The results show that the prediction model based on LSTM method is superior to BP neural network model, and the prediction accuracy is more accurate.

Keywords: NO₂ · Long short-term memory · BP neural network · Prediction

1 Introduction

With the rapid development of urban integration in China, the city has developed rapidly. But the environment has been damaged with the continuous expansion of industry. A large number of harmful substances are discharged into the atmosphere, among which carbon monoxide (CO), sulfur dioxide (SO₂) and other air pollutants also seriously affect the air quality and people's health. Especially the pollution is more serious in Beijing Tianjin Hebei region of China [1,2]. Therefore, these harmful substances of air pollution have been widely concerned. The

carbon dioxide is a toxic gas with a reddish brown and pungent smell at room temperature. The Long term exposure to high concentration of NO₂ will damage people's immune system and pollute the environment. In order to prevent the impact of air quality pollution on people, it is necessary to establish an effective and accurate model to predict the concentration of NO₂.

The artificial intelligence is widely used in recent years. A large number of prediction methods are widely used in the concentration prediction of air pollutants [3]. For example, the multiple linear regression method [4], the time series method [5] and the Kalman filter technology [6]. Due to the strong nonlinearity of the variables in the atmospheric environment, the linear method is easy to be affected by the changes of environmental variables and make the prediction inaccurate. Therefore, the artificial neural network method with strong nonlinear ability is proposed and widely used [7]. Aizhi Wang proposed a BP neural network method for air pollution prediction [8]. The prediction results show that the method is effective and has good nonlinear and generalization ability. But BP neural network has some problems, such as difficult to determine the network structure, slow training speed and low prediction accuracy. Saratha Sathasivam proposes a recursive neural network (RNN), which folds the hidden layer in time, and iteratively calculates the memory ability of the state characteristics of time series data [9]. The RNN has the problem of gradient vanishing, which makes it difficult for RNN to learn the state characteristics of long-term sequence data. In order to solve this problem, the LSTM method is proposed. For this method, the memory unit is added to the neural unit of RNN hidden layer to effectively solve the long-term dependence of information and the problem of gradient disappearance.

Thus, this paper takes a campus in Beijing as the research area. The hourly NO₂ historical concentration data is collected by the atmospheric monitor. The BP neural network model and LSTM neural network model are established to predict NO₂ concentration respectively, and the prediction accuracy of the two models is compared. The rest of this paper is introduced. The mathematical model is described for the second part. The third part describes the data source and data process. The fourth part gives the simulation analysis and comparison. Finally, the conclusion is obtained in this paper.

2 Mathematical Model

The LSTM method is a kind of recurrent neural network. The structure of LSTM is similar to BP neural network, and it is composed of input layer, hidden layer and output layer. The structure consists of three gates and a memory unit. The three gates are input gate, forgetting gate and output gate. The information can selectively affect the state of every moment in the recurrent neural network by three gates. Thus, the function of forgetting and memory can be realized, and the problem of long-term dependence can be solved. The LSTM solves the problem that the information interval is big or small and the length is different, and avoids the gradient disappearing. The LSTM gate structure is a sigmoid

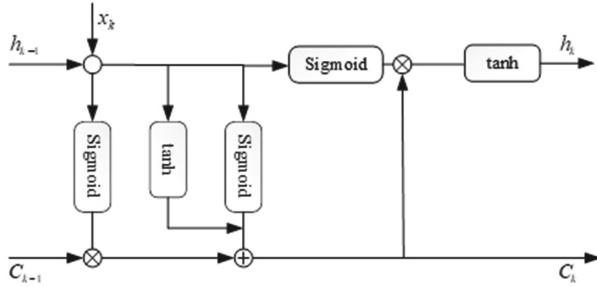


Fig. 1. The structure diagram of LSTM prediction model

neural network and bitwise multiplication. The structure diagram of LSTM is shown in Fig. 1.

The three gates functions are calculated as follows:

$$f_k = \sigma(W_f[h_{k-1}, x_k] + b_f) \quad (1)$$

$$i_k = \sigma(W_i[h_{k-1}, x_k] + b_i) \quad (2)$$

$$\tilde{C}_k = \tanh(W_c[h_{k-1}, x_k] + b_c) \quad (3)$$

$$C_k = f_k \cdot C_{k-1} + i_k \cdot \tilde{C}_k \quad (4)$$

$$O_k = \sigma(W_o \cdot [h_{k-1}, x_k] + b_o) \quad (5)$$

$$h_k = O_k \cdot \tanh(C_k) \quad (6)$$

where h_{k-1} and x_k is the previous output in forget gate, f_k is the output of h_{k-1} and x_k by sigmoid, C_{k-1} is the current cell. The σ is the sigmoid. The input date and the tanh function can obtain new candidate vector \tilde{C}_k . The W_f , W_i , W_c , W_o are four parameter matrices with dimensions of $[2n, n]$. b represents the corresponding paranoid.

LSTM calculates a vector $f = \text{sigmoid}(W_1 * x + W_2 * h)$ with dimension n according to the current input x_k and the output h_{k-1} at the previous moment by “forgetting gate”. Then, the state c_{k-1} of the last time is multiplied by the vector. The value range of sigmoid as activation function is $[0,1]$. When the value of f is close to 1, the information will be retained. When the value is 0, the information will be forgotten. Finally, the output of forgetting door is obtained. The output and current information state of the input gate at the above moment are processed by sigmoid function and tanh function. The output of the input gate is obtained, and updates the new status. Finally, the state update of memory unit is obtained by adding. In the output gate, the latest state C_k , the previous output h_{k-1} and the current input x_k are processed by the sigmoid function to get the output h_k at the time.



Fig. 2. Atmospheric monitoring equipment

3 Data Source and Processing

In this paper, the air pollution data and meteorological condition data is obtained from a monitoring point on a campus in Beijing. The air pollution degree of the area is evaluated by the data of the area points. The Air quality data of the monitoring point from May 2018 to October 2018 are selected as the original data set. The selected data set includes PM₁₀, NO₂, CO, SO₂, O₃ and temperature (T), which are measured hourly. Two prediction models are used to predict NO₂ concentration. The monitoring equipment consists of outdoor monitoring sensor probe, solar panel and distribution box, which is shown in Fig. 2.

In order to establish the corresponding model, the 900 groups data is selected as the training set, and the other 300 groups is the prediction set. For the dimensions of different variables, in the process of data processing, the collected sample data are normalized. It is given as follows:

$$x_g = \frac{x_{in} - x_{min}}{x_{max} - x_{min}} \quad (7)$$

where the value range after x_g normalization is [0,1], x_{in} is the collected sample data, x_{max} is the maximum value in the collected sample data set, and x_{min} is the minimum value in the collected data.

4 Simulation Analysis and Comparison

In this section, the comparison between the prediction accuracy of different prediction models is introduced. The two prediction model is established by the training data, and then the established prediction model is used to predict the

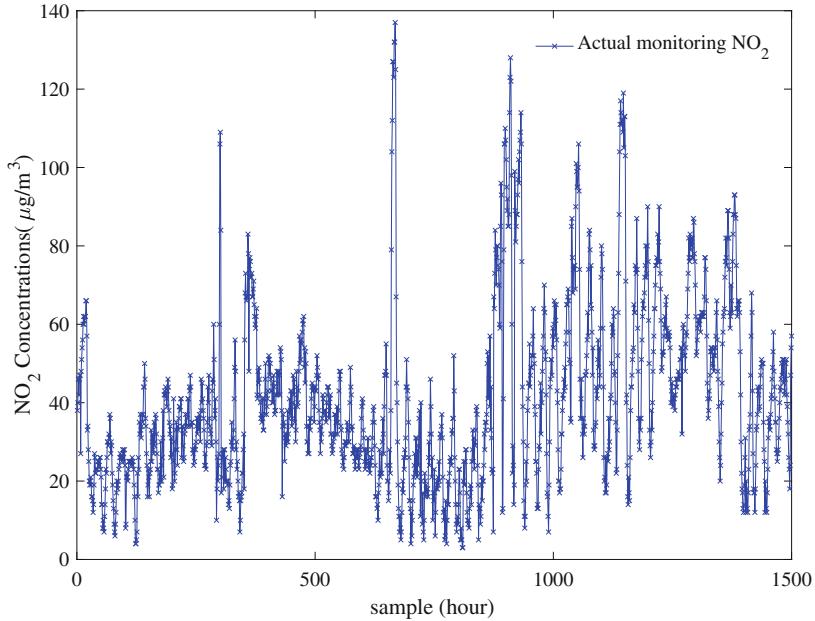


Fig. 3. The actual monitoring NO₂ concentration data

NO₂ concentration. This paper selects the data from May 2018 to October 2018 as the validation data to verify the prediction model. The Fig. 3 shows a time series diagram of the validation data. Firstly, we set the hidden layer as 9 for the LSTM prediction model, input layer is set as 6, output layer is set as 1. And the learning rate is set as 0.0004, the number of iterations is set as 1000.

In order to better compare the prediction accuracy between two prediction models, the different performance indexes are used to evaluate the accuracy of prediction NO₂. The indexes include mean absolute error (MAE), root mean square error (RMSE). The details are given as follows:

$$MAE = \frac{1}{n} \sum_{k=1}^n |h_k - \hat{h}_k| \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n [h_k - \hat{h}_k]^2} \quad (9)$$

where h_k is the actual measured NO₂ concentration value, the \hat{h}_k is the predicted NO₂ concentration. In order to further compare the prediction accuracy of BP and LSTM model for NO₂ concentration.

The Fig. 4 shows that LSTM prediction model and BP prediction model can predict NO₂ concentration well. The table 1 shows the error performance indicators of the two prediction models. From the error analysis, the prediction

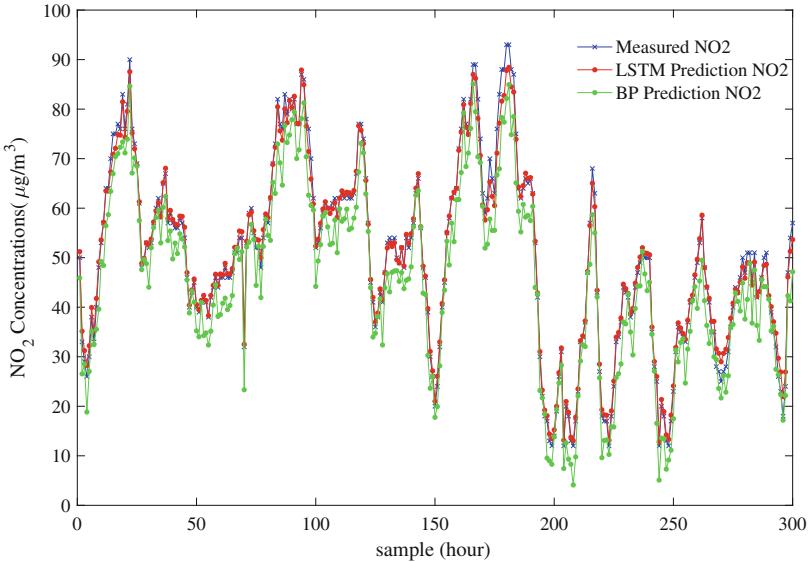


Fig. 4. The NO₂ concentration prediction results in the next hour

accuracy of LSTM prediction model is higher than that of BP prediction model. The MAE and RMSE of LSTM prediction model are 1.206 and 1.640 respectively. The MAE of BP prediction model is 4.936 and the RMSE is 5.859. The prediction accuracy of BP prediction model may not be obvious. In order to compare the accuracy of LSTM prediction model and BP prediction model, the NO₂ concentration is predicted in the next two hours.

As shown in Fig. 5, the prediction comparison of LSTM model and BP model for NO₂ concentration in the next two hours. It can be clearly seen from Fig. 5 that the prediction effect of LSTM is better than that of BP prediction model. When the actual measurement value is between 0–100, the BP model prediction result is obviously inconsistent with the actual measurement value. Which leads to the large deviation and obvious degradation of prediction ability. The prediction value of LSTM model is better than BP model. The MAE and RMSE of LSTM prediction model are 1.521 and 1.747 respectively. The MAE of BP prediction model is 6.238. The RMSE was 6.984. The table 1 shows the error performance indicators of the two prediction models. From the analysis of the error performance, the LSTM prediction model has better prediction performance for the longer time NO₂ concentration, and the LSTM model can predict the NO₂ concentration well. It is further proved that LSTM prediction model can control discarding or adding information through its own structural characteristics. The NO₂ concentration information can selectively pass through, so as to realize the memory function of state characteristics of time series process. However, the BP neural network lacks this ability, which leads to its insufficient prediction ability and reduced prediction accuracy.

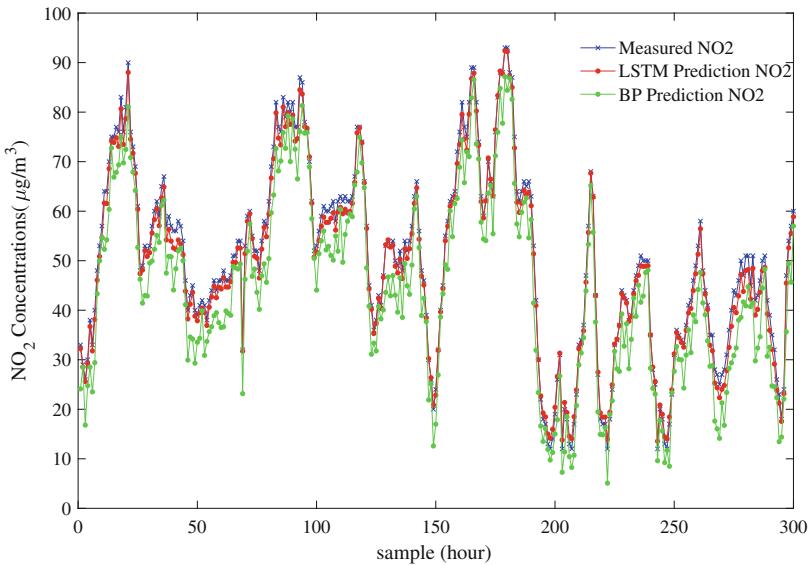


Fig. 5. The NO₂ concentration prediction results in the next two hours

Table 1. The error performance index of LSTM model and BP prediction model

Prediction model	Predicted hour	MAE	RMSE
LSTM	1	1.206	1.640
BP	1	4.936	5.859
LSTM	2	1.521	1.747
BP	2	6.238	6.984

5 Conclusion

In this paper, the BP prediction model and LSTM neural network prediction model are established by the obtained historical data and used to predict the NO₂ concentration. By comparing the prediction accuracy of two different models, the results shown that the proposed LSTM model in this paper is more accurate. The MSE of 1.206 and RMSE of 1.640. The MSE and RMSE of BP neural network prediction model are 4.936 and 5.859 respectively. The prediction accuracy of LSTM model is higher than BP neural network model. When the prediction time is extended, the MSE and RMSE of LSTM are 1.521 and 1.747 respectively. The MSE of BP model is 6.238, and the RMSE is 6.984. The reduction speed of The prediction ability for LSTM model is slower than BP model. Thus, the results show that the proposed prediction model in this paper is effective, and the model predicts NO₂ concentration more accurately and has long range prediction and strong adaptive ability.

Acknowledgments. This study is support by National Natural Science Foundation of China (61873006, 61673053), and National Key Research and Development Project (2018YFC1602704, 2018YFB1702704).

References

1. Zheng, S., Yi, H., Li, H.: The impacts of provincial energy and environmental policies on air pollution control in China. *Renew. Sust. Energ. Rev.* **49**, 386–394 (2015). <https://doi.org/10.1016/j.rser.2015.04.088>
2. Zhang, H., Wang, S., Hao, J.: Air pollution and control action in Beijing. *J. Clean Prod.* **112**, 1519–1527 (2016). <https://doi.org/10.1016/j.jclepro.2015.04.092>
3. Kavousi-Fard, A., Samet, H., Marzbani, F.: A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. *Expert Syst. Appl.* **41**, 6047–6056 (2014). <https://doi.org/10.1016/j.eswa.2014.03.053>
4. Donnelly, A., Misstear, B., Broderick, B.: Real time air quality forecasting using integrated parametric and non-parametric regression techniques. *Atmos. Environ.* **103**, 53–65 (2015). <https://doi.org/10.1016/j.atmosenv.2014.12.011>
5. Kumar, U., Jain, V.K.: ARIMA forecasting of ambient air pollutants (O_3 , NO, NO_2 and CO). *Atmos. Environ.* **24**, 751–760 (2010). <https://doi.org/10.1007/s00477-009-0361-8>
6. Ridder, K.D., Kumar, U., Lauwaet, D., et al.: Kalman filter-based air quality forecast adjustment. *Atmos. Environ.* **50**, 381–384 (2012). <https://doi.org/10.1016/j.atmosenv.2012.01.032>
7. Nayebi, M., Khalili, D., Amin, S.: Daily stream flow prediction capability of artificial neural networks as influenced by minimum air temperature data. *Biosyst. Eng.* **95**, 557–567 (2006). <https://doi.org/10.1016/j.biosystemseng.2006.08.012>
8. Bai, H.M., Shen, R.P., Shi, H.D.: Forecasting model of air pollution index based on BP neural network. *Environ. Sci. Technol.* **36**, 186–189 (2013). <https://doi.org/10.3969/j.issn.1003-6504.2013.03.039>
9. Sathasivam, S.: Logic learning in hopfield networks. *Mod. Appl. Sci.* **2**, 57–63 (2008). <https://doi.org/10.5539/mas.v2n3p57>



Fractional Integral Sliding Mode Control for Trajectory Tracking of Baxter Robot Manipulators

Xiru Wu^(✉) and Jingtong Jia

College of Electronic Engineering and Automation,
Guilin University of Electronic Technology, Guilin 541004, Guangxi, China
xiruwu520@163.com

Abstract. A fractional integral sliding mode control algorithm is proposed to solve the problems of poor tracking accuracy and poor robustness for multi-degree-of-freedom manipulators during trajectory tracking. This method combines the advantages of sliding mode control strategy with fractional integration, using exponential reaching law based on fractional integral sliding mode surface. In addition, the approximate estimation term of external disturbance is added to the system, which can achieve rapid convergence and has strong anti-interference ability. Moreover, the stability of the system could be guaranteed by the Lyapunov theory. Numerical simulations of the seven degree-of-freedom(7-DOF) Baxter robot manipulators shows promising results that validate the high-precision tracking performance and the better robustness of the proposed robot system with external disturbances.

Keywords: Multi-degree-of-freedom manipulator · Trajectory tracking · Sliding mode control · Exponential reaching law

1 Introduction

Multi-degree-of-freedom manipulators are widely used in important links of industrial production due to their high speed, dexterity, and precise repeatability. It is a nonlinear and strongly coupled time-varying system [1, 2], because of its limited tasks and joint space. There are a series of uncertain factors such as system modeling errors, joint friction, and unstable signal detection etc. The dynamic performance of the system is difficult to express with an accurate mathematical model. Therefore, the control of multi-degree-of-freedom manipulators has become a research hotspot of scholars at home and abroad in recent years [3, 4].

In view of the trajectory tracking problem of the robotic manipulators, there are many control methods proposed at present, including adaptive control [5],

PID control [6], sliding mode control [7], fuzzy theory and neural network control [8], etc. Sliding mode variable structure control makes the state of the system slide along the sliding mode surface by switching the control amount. The invariant characteristics of the system can reduce the influence of parameter perturbation and external interference, so it is widely used in robot control. Tairen Sun et al. [9] designed a sliding mode adaptive control based on neural network, which combines sliding mode technology and adaptive technology to ensure trajectory tracking of the robotic manipulators. Qiao Lei et al. [10] designed an adaptive second-order fast non-singular terminal sliding mode control scheme, which is used to track the trajectory of a fully driven autonomous underwater vehicle in the presence of dynamic uncertainty and time-varying external interference. Amir Salimi Lafmejani et al. [11] studied the trajectory tracking control of a 6-DOF pneumatic Gough-Stewart parallel robot, the position control of the pneumatic actuator is based on the dynamic model of the system based on back-stepping sliding mode controller. In this way, trajectory tracking control is performed on different trajectories by using the rotation sensor and the calculated position based on the joint space and the task space at the same time.

Fractional integral sliding mode combines the dual advantages of fractional and sliding mode control. It can further improve the control performance of the system on the basis of traditional sliding mode control, and has become an important research direction of modern nonlinear control. Ahmet Dumlu et al. [12] used a fractional-order adaptive integral sliding mode control method to perform trajectory tracking control on a six-degree-of-freedom robot. The method is composed of fractional-order adaptive integral sliding mode control to achieve finite time convergence and better tracking performance. In addition, fractional-order sliding mode can also reduce chattering and discontinuity of control input. In this paper, based on the trajectory tracking of Baxter robot manipulator, using the memory and genetic characteristics of fractional integration operator, fractional integration operator is introduced into traditional sliding mode control. Besides, a fractional integration sliding mode control algorithm based on exponential reaching law is designed. The effectiveness of the proposed control method is verified by a simple manipulator system model robustness.

2 Dynamic Model of Manipulator

In this paper, Baxter robot is used to discuss the trajectory tracking of the robotic arm, the structure of the robot is shown in Fig. 1. It uses a computer as a remote control center and uses the network to communicate with the robot's main computer to achieve real-time control of the two robot manipulators. The two manipulators of the system have seven degrees of freedom. The control system compares the actual position and attitude information of the robotic manipulator with the target information, and calculates the error of the two information in the controller. Then outputs the corresponding joint control torque to adjust the angle of each joint of the manipulator in real time to achieve the goal of track tracking.

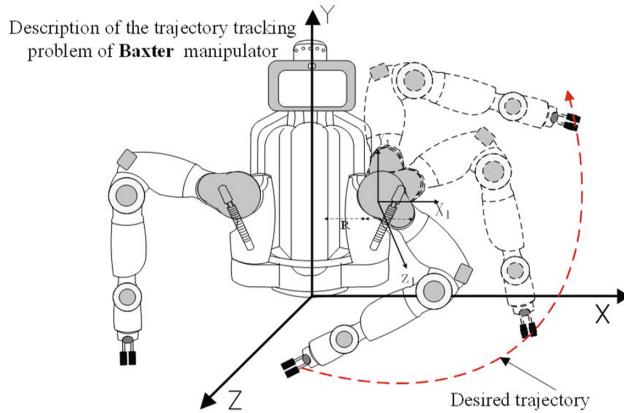


Fig. 1. The diagram of manipulator trajectory tracking

2.1 Kinetic Equation

As can be seen from Fig. 1, the Baxter robotic manipulator has 7-DOF, including 3 degrees of freedom of the shoulder joint, 2 degrees of freedom of the elbow joint and 2 degrees of freedom of the wrist joint. This paper analyzes its left arm, and establishes a D-H model of the robot's left arm according to the parameters of each part of the robot arm, as shown in Fig. 2. Where O is the origin of the base coordinate system, A is the shoulder joint, B is the elbow joint, C is the wrist joint, and D is the origin of the actuator end coordinate system. D-H parameters are shown in Table 1.

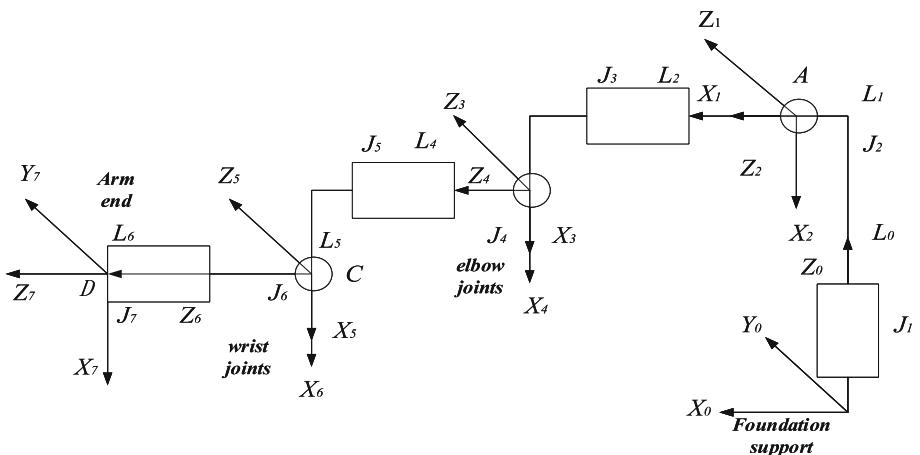


Fig. 2. Baxter robot left arm D-H coordinate system

Table 1. The D-H parameters of the left arm of the Baxter robot

Linkage i	Joint Angle θ_i	Length $d_i(m)$	Torsional Angle $\alpha_{i-1}(^\circ)$	Deviation $a_{i-1}(m)$	Angle Range $\theta_i(^\circ)$
1	θ_1	0	90	0	(−173.5, 173.5)
2	θ_2	0	90	0	(60, −123)
3	θ_3	0	−90	0.364	(51, −141)
4	θ_4	0.68	−90	0	(150, −3)
5	θ_5	0.68	90	0.374	(−175.25, 175.25)
6	θ_6	0	−90	0	(120, −90)
7	θ_7	0	0	0.15	(−175.25, 175.25)

The kinetic equation of the 7-DOF manipulator is expressed as:

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \tau - \tau_h - \tau_d \quad (1)$$

where, q, \dot{q}, \ddot{q} are joint position vector, velocity vector and acceleration vector respectively, $q, \dot{q}, \ddot{q} \in R^{7 \times 1}$; $M(q)$ is the inertia matrix, $M(q) \in R^{7 \times 7}$; $C(q, \dot{q})$ is the Coriolis force and centrifugal force matrix, $C(q, \dot{q}) \in R^{7 \times 7}$; $G(q)$ is the gravity vector, $G(q) \in R^{7 \times 1}$; τ_h is the operating torque, $\tau_h \in R^{7 \times 1}$, is the unknown disturbance torque, $\tau_d \in R^{7 \times 1}$; τ is the input torque to the joint, $G(q) \in R^{7 \times 1}$, (1) has the following characteristics:

$$M(q) = (M(q))^T, |M(q)| > 0$$

$$x^T(\dot{M}(q) - 2C(q, \dot{q}))x = 0, \forall x \in R^{7 \times 1}$$

In order to simplify the Eq. (1), define

$$N(q, \dot{q}) = C(q, \dot{q}) + G(q) + \tau_h, N(q, \dot{q}) \in R^{7 \times 1} \quad (2)$$

Thus, we can obtain

$$M(q)\ddot{q} + N(q, \dot{q}) = \tau - \tau_d \quad (2)$$

Here, the unknown external disturbance is regarded as the uncertainty of the model and the influence of factors such as friction is ignored.

3 Fractional Integral Sliding Mode Control

Figure 3 shows the block diagram of the Baxter robot's manipulator trajectory tracking system. After designing the controller and sliding mode surface, the trajectory error of robotic manipulator system will become smaller under the action of the designed control rate. Therefore, the error will gradually converge, which improves the system performance.

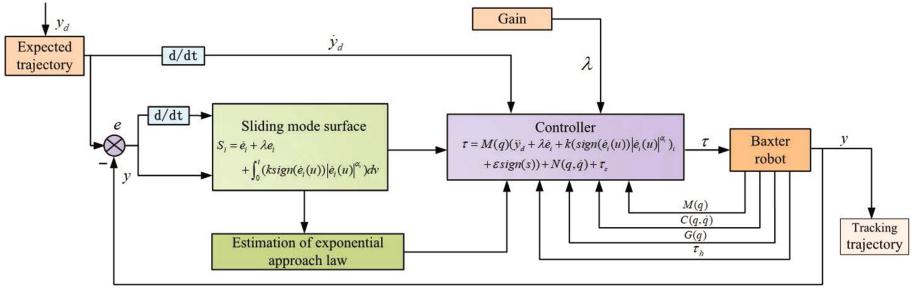


Fig. 3. The trajectory tracking control process

3.1 Controller Design

Let t_d be the desired joint vector given, which is second-order differentiable. Define tracking error $e = y_d - y$, $e = (e_1, e_2, \dots, e_7)^T$. We can design an integral sliding mode surface:

$$s_i = \dot{e}_i + \lambda e_i + \int_0^t (k_i \text{sign}(\dot{e}_i(u)) |\dot{e}_i(u)|^{\alpha_i}) dv \quad (3)$$

where λ is the control gain during real-time application, which is a diagonal normal number matrix of 7×7 , sliding surface parameter $k_i > 0$, $i = 1, 2, \dots, 7$. $0 < \alpha_i < 1$, $i = 1, 2, \dots, 7$. During the sliding state, $s_i(t) = 0$ ($i = 1, 2, \dots, 7$), we can obtain

$$\dot{e}_i + \lambda e_i + \int_0^t (k_i \text{sign}(\dot{e}_i(u)) |\dot{e}_i(u)|^{\alpha_i}) dv = 0 \quad (4)$$

Its derivative is

$$\ddot{e}_i + \lambda \dot{e}_i + k_i \text{sign}(\dot{e}_i(u)) |\dot{e}_i(u)|^{\alpha_i} = 0 \quad (5)$$

During the sliding consider $s_i = 0$, $\dot{s}_i = 0$, tracking error can achieve asymptotic convergence under non-zero initial conditions. Design the control input as:

$$\tau = M(q)(\ddot{y}_d + \lambda \dot{e}_i + k(\text{sign}(\dot{e}_i(u)) |\dot{e}_i(u)|^{\alpha_i})) + N(q, \dot{q}) + \tau_c \quad (6)$$

$$s = [s_1 \ s_2 \ \dots \ s_7]^T, k = \text{diag}[k_1, k_2, \dots, k_7]$$

$$(\text{sign}(\dot{e}_i) |\dot{e}_i|^{\alpha_i})_i = \left[\text{sign}(\dot{e}_1) |\dot{e}_1|^{\alpha_1} \ \dots \ \text{sign}(\dot{e}_7) |\dot{e}_7|^{\alpha_7} \right]^T$$

In Eq. (6), reaching law parameter $\varepsilon > 0$, control gain $\lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_7]$, $\lambda_i > 0$, τ_c is the estimate of τ_d , take the derivative of s with respect to time, combined Eqs. (2), (3), (5) and (6):

$$\dot{s} = -\varepsilon \text{sign}(s) - (M(q))^{-1}(\tau_c - \tau_d) \quad (7)$$

In order to ensure the rapid approach and weaken chatter, we should increase moderately and decrease moderately.

3.2 Stability Analysis

Consider a positive definite Lyapunov function

$$V = \frac{1}{2} s^T s \quad (8)$$

Derivation of (8) about time, combine Eq. (8) we can obtain

$$\begin{aligned} \dot{V} &= s^T \dot{s} = s^T (-\varepsilon sign(s) - (M(q))^{-1}(\tau_c - \tau_d)) \\ &= -\varepsilon s^T sign(s) - s^T (M(q))^{-1}(\tau_c - \tau_d) \leq -\varepsilon \|s\|_1 \leq 0 \end{aligned} \quad (9)$$

$$\|s\|_1 = \sum_{i=1}^7 |s_i|$$

Equation (9) shows that it is bounded, when $t \rightarrow \infty$, s_i asymptotically converges to zero. From the Eqs. (4) and (5), it can be seen that the tracking error of the system can quickly and steadily converge to the equilibrium point.

3.3 Simulation Results and Analysis

Each arm of Baxter robot has seven redundant degrees of freedom. In this paper, a single arm model of Baxter robot is established using MATLAB robot toolbox, as shown in Fig. 4. The proposed algorithm is applied to the redundant manipulator with seven degrees of freedom, set the desired trajectory of each joint to $y_d = \sin(\pi t)$, then there is $\dot{y}_d = \pi \cos(\pi t)$, $\ddot{y}_d = -\pi^2 \sin(\pi t)$, the initial state of the joint is all set to 0 rad. In Eq. (3), the sliding mode surface parameter is set as $k_1 = 100$, $k_2 = 120$, $k_3 = 120$, $k_4 = 100$, $k_5 = 100$, $k_6 = 100$, $k_7 = 100$, $\alpha_i = 0.9$, $i = 1, 2, \dots, 7$. Reaching law parameter $\varepsilon = 0.5$, $\lambda_i = 30$, $i = 1, 2, \dots, 7$. For other parameters, please refer to the D-H parameter table given above. Then use MATLAB to simulate the trajectory tracking error of each joint, the corresponding simulation results are shown in Fig. 5, 6, 7, 8, 9, 10 and 11. Comparing of the proposed algorithm with traditional sliding mode control method, the error generated by the fractional order integral sliding mode control method convergence rapidly.

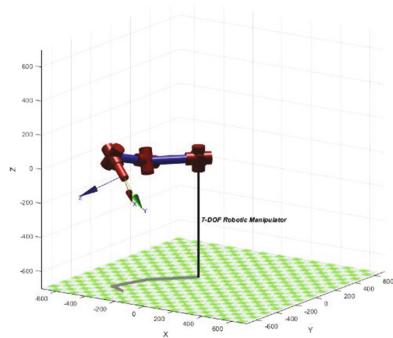


Fig. 4. Robot MATLAB model

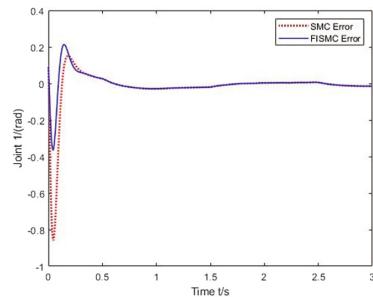


Fig. 5. The trajectory tracking error of joint 1

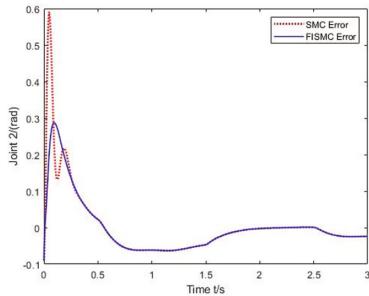


Fig. 6. The trajectory tracking error of joint 2

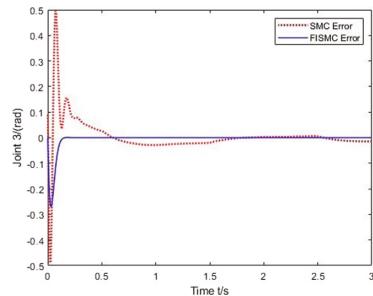


Fig. 7. The trajectory tracking error of joint 3

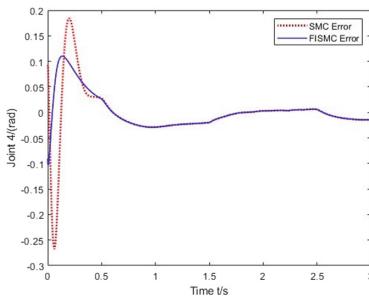


Fig. 8. The trajectory tracking error of joint 4

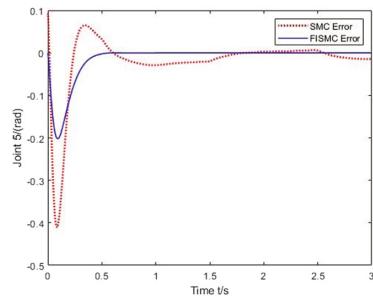


Fig. 9. The trajectory tracking error of joint 5

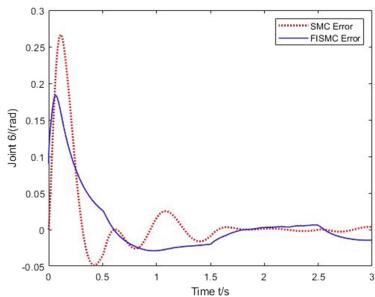


Fig. 10. The trajectory tracking error of joint 6

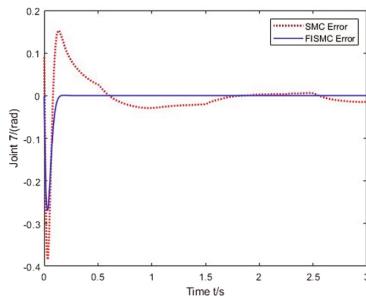


Fig. 11. The trajectory tracking error of joint 7

4 Conclusions

A fractional integral sliding mode control method is proposed. The approximate estimation term of external disturbance is added to the manipulator model, so that the system can realize rapid convergence and has strong anti-interference ability. The trajectory tracking of the 7-DOF manipulator is simulated. Result shows the control system not only achieve high-precision trajectory tracking performance, but also has strong anti-interference ability, thus verifying the effectiveness of the controller.

References

1. Ouyang, P.R., Pano, V.: Position domain synchronization control of multi-degrees of freedom robotic manipulator. *Magn. Reson. Semicond. Nanostruct. J. Dyn. Syst. Measur. Control* **136**(2), 173–186(2013). <https://doi.org/10.1115/1.4025755>
2. Wang, H.: Adaptive control of robot manipulators with uncertain kinematics and dynamics. *IEEE Trans. Autom. Control* **62**(2), 948–954 (2017). <https://doi.org/10.1109/TAC.2016.2575827>
3. Rani, K., Kumar, N.: Intelligent controller for hybrid force and position control of robot manipulators using RBF neural network. *Int. J. Dyn. Control* **7**(2), 767–775 (2019). <https://doi.org/10.1007/S40435-018-0487-Y>
4. Zirkohi, M.M.: Direct adaptive function approximation techniques based control of robot manipulators. *J. Dyn. Syst. Measur. Control-Trans. ASME* **140**(1), 98–111 (2018). <https://doi.org/10.1115/1.4037269>
5. Seo, D., Akella, M.R.: Non-certainty equivalent adaptive control for robot manipulator systems. *Syst. Control Lett.* **58**(4), 304–308 (2009). <https://doi.org/10.1016/J.SYSCONLE.2008.11.008>
6. Xu, J., Qiao, L.: Robust adaptive PID control of robot manipulator with bounded disturbances. *Math. Prob. Eng.* **2013**(13), 1–13 (2013). <https://doi.org/10.1155/2013/535437>
7. Tayebihaghghi, S., Piltan, F., Kim, J.M.: Control of an uncertain robot manipulator using an observation-based modified fuzzy sliding mode controller. *Int. J. Intell. Syst. Appl.* **10**(3), 41–49 (2018). <https://doi.org/10.5815/IJISA.2018.03.05>

8. Wu, X., Wang, Y., Dang, X.: Robust adaptive sliding-mode control of condenser-cleaning mobile manipulator using fuzzy wavelet neural network. *Fuzzy Sets Syst.* **235**(16), 62–82 (2014). <https://doi.org/10.1016/J.FSS.2013.07.009>
9. Sun, T., Pei, H., Pan, Y.: Neural network-based sliding mode adaptive control for robot manipulators. *Neurocomputing* **74**(14–15), 2377–2384 (2011). <https://doi.org/10.1016/J.NEUCOM.2011.03.015>
10. Qiao, L., Zhang, W.: Adaptive non-singular integral terminal sliding mode tracking control for autonomous underwater vehicles. *Control Theory Appl. IET* **11**(8), 1293–1306 (2017). <https://doi.org/10.1049/IET-CTA.2017.0016>
11. Lafmejani, A.S., Masouleh, M.T., Kalhor, A.: Trajectory tracking control of a pneumatically actuated 6-DOF Gough-Stewart parallel robot using backstepping-sliding mode controller and geometry-based quasi forward kinematic method. *Robot. Comput.-Integr. Manuf.* **54**, 96–114 (2018). <https://doi.org/10.1016/J.RCIM.2018.06.001>
12. Dumlu, A.: Design of a fractional-order adaptive integral sliding mode controller for the trajectory tracking control of robot manipulators. *Proc. Inst. Mech. Eng.* **232**(9), 1212–1229 (2018). <https://doi.org/10.1177/0959651818778218>



Fast Detection of Passion Fruit with Multi-class Based on YOLOv3

Xiru Wu^(✉) and Rongchai Tang

College of Electronic Engineering and Automation,
Guilin University of Electronic Technology, Guilin 541004, China
Xiru.WU520@163.com

Abstract. In order to quickly identify and locate passion fruits with different maturity in natural environment, a multi-class passion fruit fast detection algorithm based on YOLOv3 is proposed. Firstly, the data is preprocessed and made into VOC2007 data set format. Secondly, the Densenet network is added to YOLOv3 feature extraction module for enhancing the feature propagation of convolutional layers. Finally, the multi-scale prediction is reduced to single-scale prediction in YOLOv3, and medium- sized object detectors are retained among large, medium, and small object detectors. The experimental results show that the method can effectively detect passion fruit with different maturity in natural environment.

Keywords: YOLOv3 · Densenet network · Passion fruits · Multi-class

1 Introduction

In recent years, target detection plays an important role in the field of computer vision [1–4]. With the development of intelligent agriculture, target detection technology is widely used in the field of fruit picking and classification. In order to realize the automatic picking of passion fruit, it is necessary to solve the classification, recognition and positioning of passion fruits with different maturity. There are many inevitable interference factors in passion fruit detection, such as illumination changes, fruit overlapping, branches and leavesis occlusion. The main research content of this paper is to detect passion fruit with different maturity in natural environment. With the development of deep learning [5–8], a large number of researchers use neural networks to learn the target features in the pictures. A pre-training model is obtained by training the neural network, which can be used to identify and locate the target. In the classification and detection of fruits, different researchers fully extract the features of fruits through neural networks to achieve higher learning efficiency and improve recognition accuracy. Such as, Zhang et al. [9] proposed a fruit classification method based on a 13-layer deep convolutional neural network to classify 18 different fruits. Yu et al. [10] proposed strawberry detection based on Mask-RCNN, which uses Resnet50 as

© The Editor(s) (if applicable) and The Author(s), under exclusive license

to Springer Nature Singapore Pte Ltd. 2021

Y. Jia et al. (Eds.): CISC 2020, LNEE 706, pp. 818–825, 2021.

https://doi.org/10.1007/978-981-15-8458-9_87

the backbone network and combines with the feature pyramid to extract strawberry features. Aiming at the problem of locating the picking point of the picking robot, Zhuang et al. [11] proposed a computer-vision-based method to locate the acceptable picking point for litchi group by using the iterative retenex algorithm which can improve the accuracy of the litchi picking point when the picking robot picks the litchi. With the real-time requirements of neural networks, the series networks of YOLO [12–14] are proposed to improve the real-time performance of target detection. In an unstructured growth environment, Xue et al. [15] use YOLOv2 as the basic network and add Densenet network to the Darknet-53 network to detect immature mangoes. In view of the problems of fruit overlapping and branches and leaves occlusion, Tian et al. [16] proposed a real-time detection method for apple with different growth stages in real orchards based on YOLOv3 network. Considering the complexity of the background interference, this paper uses the YOLOv3 network as the main algorithm to detect the passion fruit with different maturity. The multi-scale prediction is reduced to single-scale prediction, and then Densenet network is added to the convolution layer for enhancing the feature extraction ability of the network to realizes the fast detection of passion fruit with multi-class in natural environment.

2 Methodology

2.1 Densenet Network

Densenet network learns the idea of Resnet network, in which any two layers are interconnected to each other and the input of each layer comes from the output of all the previous layers in order to improve its feature extraction ability. With the increasing of convolution layers, the transmission of the feature map will weaken gradually and disappears in the feature extraction module of YOLOv3. In order to solve the problem, DenseNet network was used to enhance feature extraction and propagation. The DenseNet network structure is shown in Fig. 1. The output of the m th layer is a set of nonlinear transformations for the inputs of all previous layers.

$$X_m = H_m([X_0, X_1, \dots, X_{m-1}]) \quad (1)$$

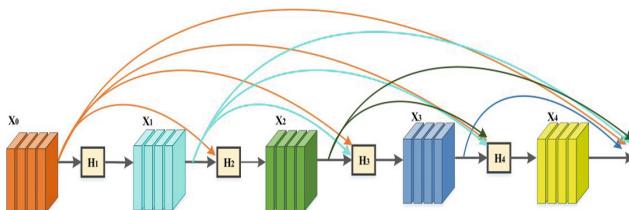


Fig. 1. The structure diagram of DenseNet

where $[X_0, X_1, \dots, X_{m-1}]$ represents the splicing of feature maps of layers X_0, X_1, \dots, X_{m-1} and H_m is the function for splicing features. It can be seen that the dense connection of each module can effectively use the shallow and deep features, making the network efficient and reducing the network complexity and calculation.

2.2 YOLOv3-Densenet Network

YOLOv3 uses Darknet53 as the feature extraction network. The final output of Darknet53 will reduce the size of the original image by 1/32, making it difficult to detect the small target in the final feature map. In response to the above problems, the Densenet network for enhancing feature propagation is added to the YOLOv3 feature extraction module. According to the detection of the same target of passion fruit, the multi-scale prediction is reduced to single scale prediction. Figure 2 is the structural diagram of the YOLOv3-Densenet network. The Densenet network is added to 13 × 13 and 26 × 26 scales to enhance the feature

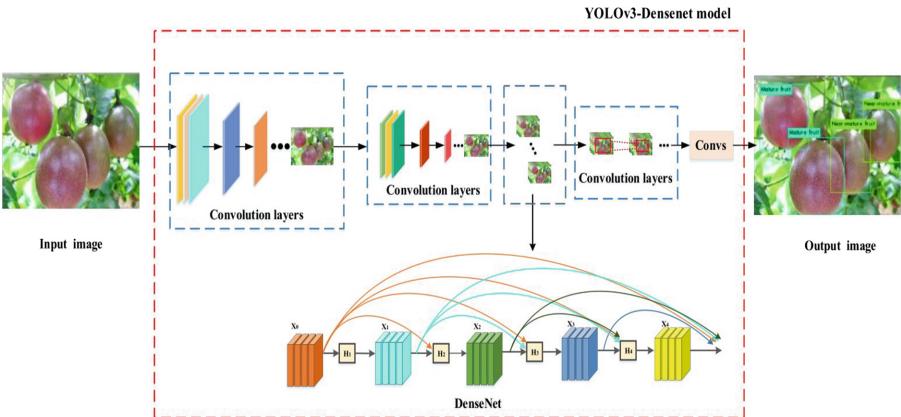


Fig. 2. The structure diagram of YOLOv3-Densenet

Table 1. The process of K-mean clustering

K-mean classification algorithm

Step1: Randomly selecting k samples $\mu_1, \mu_2, \dots, \mu_k$ as clustering centers and initializing k clustering centers

Step2: Every dataset sample x_1, x_2, \dots, x_i is distributed to the nearest cluster center

Step3: Recalculating the clustering center J of every dataset sample.

$$J = \sum_{k=1}^k \sum_{i=1}^n \|x_i - u_k\|^2$$

Step4: When the clustering center of every dataset sample no longer changes or reaches the set number of iterations, the classification result will be output, otherwise step 1 will be repeated

extraction capability of the network, and then the final feature map is predicted on 13×13 scale detector through the 3×3 and 1×1 convolutional layers. In the classification process, YOLOv3 uses K-mean cluster analysis algorithm. Firstly, K samples are randomly selected as the initial cluster centers, then calculate the distance between each sample and each cluster center. Finally, the distance between the sample and its cluster center is minimized by iterative method. The process of clustering algorithm is shown in Table 1. During the picking process, due to the picking robots detect the passion fruit at different distances, the feature information of different maturity passion fruit can be extracted effectively by adding Densenet network.

3 Experimental Results and Analysis

3.1 Data

The experimental data are collected from a farmer's planting base in Longzao Village, Guilin city, and a total of 1000 pictures of passion fruit are collected. The passion fruit is divided into four different maturity, namely young fruit, near-young fruit, near-mature fruit and mature fruit as shown in Table 2.

Table 2. Type and quantity description of passion fruit

Class description	Colour description	Total number of images
Young fruit	Full Green	400
Near-young fruit	20–50% Green	200
Near-mature fruit	50–80% Red	200
Mature fruit	Full Red	200

3.2 Experimental Environment

The experiment is based on Pytorch framework, and the specific hardware and software configuration is shown in Table 3. The hyperparameter settings of the training model are shown in Table 4. During model training, batch size is set to 16, i.e., 16 passion fruit pictures are randomly selected for each network training. The YOLOv3-Densenet network randomly selects 16 pieces of marked data for training is shown in Fig. 3.

3.3 Evaluating Indicator

In this paper, the recognition accuracy and target position of passion fruit detection with different maturity are particularly important. If the accuracy of the

Table 3. Configuration of Software and hardware

Name	type
CPU	Intel I76700
RAM	32G
System	Windows 10
Framework	Pytorch
Graphics card	GeForce RTX2080Ti@11G

Table 4. Hyperparameter of the network

Paiametei	Nsmeiical valse
Batch size	16
Ssbdivisions	8
Max_batches	20000
Momentsm	0.9
Learning rate	0.001

detection and the target location is incorrect, the picking robot will make mistakes in selecting and grabbing the target. Therefore, this paper selects detection accuracy as the main evaluation index. The definition of precision is as follows:

$$P = \frac{TP}{TP + FP} \quad (2)$$

where TP is the number of positive samples which are detected correctly, and FP is the number of positive samples which are detected incorrectly. In the training of YOLOv3 network model, the loss function is defined as:

$$\text{Loss} = E_{coord} + E_{confidence} + E_{class} \quad (3)$$

The E_{coord} is the coordinate prediction error, the $E_{confidence}$ is the confidence prediction error, and the E_{class} is classification prediction error.

Table 5. Detection accuracy of passion fruits with different maturities

Class	mAP%	Detection speed/s
Young fruit	0.92	36
Near-young fruit	0.88	-
Near-mature fruit	0.82	-
Mature fruit	0.95	-

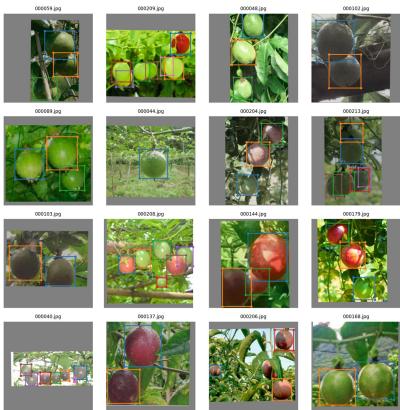


Fig. 3. Training set samples

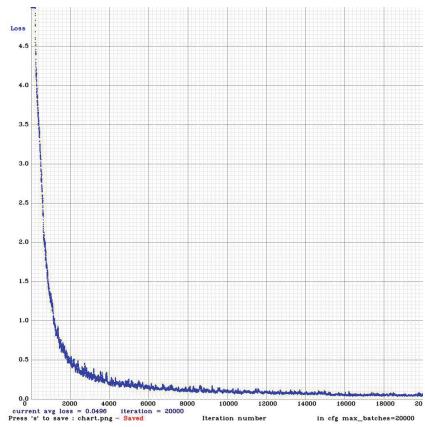


Fig. 4. Training loss

3.4 Experimental Results

In order to illustrate the detection effect, YOLOv3-Densenet is used to detect passion fruit with different maturity in natural environment. The total training times is 20000, and the loss curve is shown in Fig. 4. When the max_batches is small, the loss curve drops fast, which indicates that the network model begins to learn the image features. With the increase of training times, the loss curve decreased slowly. While the max_batches is equal to 18000, the loss curve is basically unchanged, and the final loss value is stable at about 0.05. Table 5 shows the experimental results of the same test data using different algorithms. The Fig. 5 shows the detection of passion fruit with different maturity using YOLOv3-Densenet network in natural environment. The detection accuracy of



Fig. 5. Passion fruit detection with different maturity in natural environment

passion fruit with different maturity is shown in Table 6. It can be seen that YOLOv3-Densenet can accurately identify passion fruits of different maturity and give the calibration box of each target location.

Table 6. Experimental results of different algorithms

Network	mAP%	F1%	FPS
Faster-RCNN	83.6	82.4	12
YOLO-tiny	82.3	81.6	56
YOLO-V3	81.5	83.4	48
Ours	86.7	87.3	36

4 Conclusions

This paper changes the multi-scale prediction and increases the feature extraction network of YOLOv3 in order to propose a fast detection method for passion fruit with different maturity based on the YOLOv3-DenseNet network. The experimental results shows that the network can effectively detect passion fruits with different maturity in natural environments. The later work will further adjust the model, expand the data set, and improve the detection accuracy.

References

1. Zhang, Y., Fan, Y., Xu, M., Li, W., Zhang, G., Liu, L., Yu, D.: An improved low rank and sparse matrix decomposition based anomaly target detection algorithm for hyperspectral imagery. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **1** (2020). <https://doi.org/10.1109/JSTARS.2020.2994340>
2. Mai, X., Zhang, H., Jia, X., Meng, M.Q.-H.: Faster R-CNN with classifier fusion for automatic detection of small fruits. *IEEE Trans. Autom. Sci. Eng.* **1–15** (2020). <https://doi.org/10.1109/TASE.2020.2964289>
3. Zhang, M., Pang, K., Gao, C., Xin, M.: Multi-scale aerial target detection based on densely connected inception ResNet. *IEEE Access* **8**, 84867–84878 (2020). <https://doi.org/10.1109/ACCESS.2020.2992647>
4. Wang, D., Li, C., Song, H., Xiong, H., Liu, C., He, D.: Deep learning approach for apple edge detection to remotely monitor apple growth in orchards. *IEEE Access* **8**, 26911–26925 (2020). <https://doi.org/10.1109/ACCESS.2020.2971524>
5. Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F.-Y.: Traffic flow prediction with big data: a deep learning approach. *IEEE Trans. Intell. Transp. Syst.* **16**(2), 865–873 (2015). <https://doi.org/10.1109/TITS.2014.2345663>
6. Xin, Y., Kong, L., Liu, Z., Chen, Y., Li, Y., Zhu, H., Gao, M., Hou, H., Wang, C.: Machine learning and deep learning methods for cybersecurity. *IEEE Access* **6**, 35365–35381 (2018). <https://doi.org/10.1109/ACCESS.2018.2836950>

7. Liu, L., Ouyang, W., Wang, X., Fieguth, P.W., Chen, J., Liu, X., Pietikäinen, M.: Deep learning for generic object detection: a survey. *Int. J. Comput. Vis.* **128**(2), 261–318 (2020). <https://doi.org/10.1007/S11263-019-01247-4>
8. Han, L., Huang, Y.-S.: Short-term traffic flow prediction of road network based on deep learning. *IET Intel. Transp. Syst.* **14**(6), 495–503 (2020). <https://doi.org/10.1049/IET-ITS.2019.0133>
9. Zhang, Y.-D., Dong, Z., Chen, X., Jia, W., Du, S.: Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. *Multimed. Tools Appl.* **78**(3), 3613–3632 (2019). <https://doi.org/10.1007/S11042-017-5243-3>
10. Yu, Y., Zhang, K., Yang, L., Zhang, D.: Fruit detection for strawberry harvesting robot in non-structural environment based on mask-RCNN. *Comput. Electron. Agric.* **163**, 104846 (2019). <https://doi.org/10.1016/J.COMPAG.2019.06.001>
11. Zhuang, J., Hou, C., Tang, Y., He, Y., Guo, Q., Zhong, Z., Luo, S.: Computer vision-based localisation of picking points for automatic litchi harvesting applications towards natural scenarios. *Biosyst. Eng.* **187**, 1–20 (2019). <https://doi.org/10.1016/J.BIOSYSTEMSENG.2019.08.016>
12. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788. IEEE Press, New York (2016). <https://doi.org/10.1109/CVPR.2016.91>
13. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6517–6525. IEEE Press, New York (2017). <https://doi.org/10.1109/CVPR.2017.690>
14. Redmon, J., Farhadi, A.: YOLOv3: an incremental improvement[EB/OL]. (08 April 2018), 5 February 2020. <https://arxiv.org/abs/1804.02767>
15. Xue, Y., Huang, N., Tu, S., Mao, L., Chen, P.: Immature mango detection based on improved YOLOv2. *Trans. Chin. Soc. Agric. Eng.* **34**(7), 173–179 (2018). <https://doi.org/10.11975/j.issn.1002-6819.2018.07.022>
16. Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., Liang, Z.: Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Comput. Electron. Agric.* **157**, 417–426 (2019). <https://doi.org/10.1016/J.COMPAG.2019.01.012>

Author Index

B

Bai, Chenguang, 75, 171
Bai, Zongmei, 424

C

Cao, Tao, 284
Chai, Senchun, 193
Chai, Yi, 784
Chen, Chaoyang, 307
Chen, Guangzeng, 75, 171
Chen, Qun, 630
Chen, Zengqiang, 1
Cheng, Hanlin, 460
Cheng, Pengchao, 66
Cheng, Pengjie, 773
Cheng, Yun, 699, 709
Chi, Ronghu, 506
Cho, Koanhee, 571, 594
Cui, Kunxu, 533
Cui, Lingguo, 193
Cui, Yan, 496
Cui, Yuechen, 562

D

Dai, Fengzhi, 734, 742
Dai, Jiao, 257
Deng, Yang, 674
Dong, Xuemei, 674
Dou, Lingfei, 336
Du, Junping, 11, 20, 29, 38
Duan, Xuting, 316
Duo, Jingyun, 203

F

Fang, Rongchao, 180
Fang, Ting, 611
Fang, Xiaoyu, 717
Fu, Jian, 524, 533, 542
Fu, Yongling, 533
Fu, Ziyi, 442

G

Gao, Dai, 326
Gao, Jian, 479
Gong, Linjuan, 487
Gong, Lisen, 57
Guo, Cheng, 353
Guo, Huanyu, 222
Guo, Tuanhui, 533
Guo, Yaxin, 85
Guo, Yongqi, 212, 222

H

Han, Shichang, 649
Han, Sumin, 442
Han, Xudong, 524, 542
Hanajima, Naohiko, 434
Hao, Fei, 621
Hou, Guolian, 487, 630
Hou, Huidong, 649
Hou, Xujie, 793
Hu, Bo, 487, 630
Hu, Dongfei, 516
Hu, Haidong, 750
Hu, Qiyang, 750
Hu, Yan, 773

Hu, Zhongyi, 667
 Huang, Congzhi, 487, 630
 Huang, Fengguang, 452
 Huang, Miqi, 316
 Huang, Ting, 487
 Huang, Wei, 552
 Huang, Yingbo, 649

J

Jia, Jingtong, 809
 Jiang, Chengpeng, 284
 Jiang, HengBo, 180
 Jiang, Xinyu, 621
 Jin, Xinguang, 135

K

Kou, Feifei, 11, 20, 48

L

Lei, Tongtong, 275
 Li, Bingjie, 452
 Li, Hao, 717
 Li, Haoran, 404
 Li, Hongrui, 630
 Li, Jianshen, 773
 Li, Jihan, 801
 Li, Jing, 725
 Li, Jinxuan, 107
 Li, Mingxing, 275
 Li, Na, 773
 Li, Peng, 552
 Li, Rui, 404
 Li, Shurong, 657
 Li, Tong, 395
 Li, Wei, 524
 Li, Wenbo, 750
 Li, Xiaoli, 801
 Li, Yawen, 38, 48, 66, 107
 Liu, Ce, 266
 Liu, Chong, 29
 Liu, Guoqiang, 85, 97
 Liu, Jiajia, 424
 Liu, Jian, 801
 Liu, Jianfei, 460
 Liu, Jianhui, 516
 Liu, Jin, 725
 Liu, Jinzhuo, 717
 Liu, Jiwei, 460
 Liu, Lixia, 143
 Liu, Ouming, 424
 Liu, Qunpo, 434, 602
 Liu, Shangzheng, 602
 Liu, Tao, 415, 424
 Liu, Xukuai, 584

Liu, Yachao, 479
 Liu, Yeqiong, 674
 Long, Yirong, 307
 Lou, Yunjiang, 75, 161, 171
 Lu, Xiaoyu, 180
 Luo, Shiyu, 552
 Luo, Yujie, 180
 Lv, Jianghai, 38
 Lyu, Jianting, 326

M

Ma, Hongji, 562
 Ma, Yaofei, 374
 Ma, Yue, 571, 594
 Mao, Ronghu, 460
 Meng, Lingbo, 584
 Miao, Zhonghua, 143
 Min, Yong, 630

N

Na, Jing, 649
 Ni, Haoyuan, 97

O

OuYang, WanQing, 180

P

Pan, Lei, 793
 Pan, Suying, 152
 Pang, Zhen, 470, 506
 Peng, Cheng, 117
 Peng, Haotian, 232

Q

Qin, Huayang, 1
 Qiu, Kepeng, 212, 222
 Qiu, Zifeng, 773
 Qu, Jianfeng, 717

R

Ran, Jiangtao, 75, 171
 Ren, Xuemei, 641, 682, 691, 699, 709

S

Shao, Wei, 336, 470, 506
 Shen, Hong, 384
 Shen, Liuyang, 442
 Shi, Feng, 452
 Shi, Lei, 38
 Shi, Wuxi, 57
 Shi, Yingjing, 404
 Song, Jixin, 542
 Song, Pingting, 212
 Song, Yang, 117

Song, Yunzhong, 734, 742
Su, Huilin, 487
Sun, Jia, 126
Sun, Lianghui, 326
Sun, Mingjian, 584
Sun, Mingwei, 1
Sun, Pengjian, 460
Sun, Qinglin, 1
Sun, Yujiao, 249

T

Tang, Rongchai, 818
Tian, Daxin, 316
Tian, Lin, 452
Tong, Jiahui, 725

W

Wang, Baigeng, 657
Wang, Bing, 487
Wang, Boning, 336
Wang, Dayi, 750
Wang, Dejin, 742
Wang, Hao, 384
Wang, Haoyuan, 682
Wang, Jianlin, 212, 222
Wang, Jiqiang, 611
Wang, Kang, 801
Wang, Mengke, 434
Wang, Qin, 384
Wang, Shuangxin, 773
Wang, Xin, 326
Wang, Yucan, 85
Wei, Shanbi, 784
Wu, Xiru, 809, 818
Wu, Yalei, 602

X

Xiang, Lan, 152
Xiao, Lei, 667
Xiao, Wendong, 284
Xie, Wenjing, 552
Xie, Yonghui, 793
Xing, Bin, 552
Xu, Ling, 240
Xu, Minghai, 667
Xue, Qi, 496
Xue, Zhe, 48, 66, 107

Y

Yan, Zhanyi, 193
Yang, Hongyong, 249
Yang, Liu, 193
Yang, Yiwen, 232
Yang, Yize, 249

Yao, Wenlong, 336, 470, 506
Ye, Zhiyong, 152
Yi, Shenglun, 691
Yin, Di, 742
Yin, Hongpeng, 717
Yin, Yan-chao, 353
Ying, Chengye, 143
Yu, Fei, 126
Yu, Guizhen, 85, 97
Yu, Liming, 524, 542
Yu, Meiyen, 249
Yuan, Xunpu, 20, 48
Yuan, Yasheng, 734

Z

Zhai, Junyong, 257, 266, 298
Zhan, Tianming, 725
Zhang, Baihai, 193
Zhang, Bin, 346
Zhang, Huimin, 571, 594
Zhang, Jian, 180
Zhang, Jianhua, 630
Zhang, Kebei, 750
Zhang, Lanyu, 479
Zhang, Lixin, 415
Zhang, Lufeng, 709
Zhang, Mingang, 516
Zhang, Ningning, 516
Zhang, Qi, 135, 395
Zhang, Qiang, 11
Zhang, Tianping, 384
Zhang, Wensen, 524
Zhao, Hanxue, 336
Zhao, Jichao, 734, 742
Zhao, Junxin, 161
Zhao, Liping, 374
Zhao, Liqiang, 212
Zhao, Long, 203
Zhao, Sansan, 404
Zhao, Yafei, 363
Zhao, Yanjie, 363
Zhao, Yixin, 552
Zhao, Zhen, 161, 682
Zhong, Xi, 415
Zhong, Zhijie, 298
Zhou, Bin, 85, 97
Zhou, Jianxiong, 784
Zhou, Jin, 143, 152
Zhou, Xinjie, 222
Zhu, Qinghui, 602
Zhuang, Keyu, 470
Zou, Jie, 374
Zou, Yiping, 641