

Yingmin Jia
Junping Du
Weicun Zhang *Editors*

Proceedings of 2019 Chinese Intelligent Systems Conference

Volume II

Lecture Notes in Electrical Engineering

Volume 593

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 Lab, Karlsruhe Institute for Technology, Karlsruhe, Baden-Württemberg, 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 Martin, 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 Lab, 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, Baden-Württemberg, 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

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

Aninda Bose, Senior Editor (aninda.bose@springer.com)

Japan

Takeyuki Yonezawa, Editorial Director (takeyuki.yonezawa@springer.com)

South Korea

Smith (Ahram) Chae, Editor (smith.chae@springer.com)

Southeast Asia

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

USA, Canada:

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

All other Countries:

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

Christoph Baumann, Executive Editor (christoph.baumann@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 · Junping Du · Weicun Zhang
Editors

Proceedings of 2019 Chinese Intelligent Systems Conference

Volume II



Springer

Editors

Yingmin Jia
Beihang University
Beijing, China

Junping Du
Beijing University of Posts
and Telecommunications
Beijing, China

Weicun Zhang
University of Science
and Technology Beijing
Beijing, China

ISSN 1876-1100 ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-32-9685-5 ISBN 978-981-32-9686-2 (eBook)

<https://doi.org/10.1007/978-981-32-9686-2>

© Springer Nature Singapore Pte Ltd. 2020

This work is subject to copyright. All rights are reserved 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

| | |
|---|----|
| Research on Vector Control of PMSM Based on Fuzzy PI | 1 |
| Xiangsheng Liu, Zhengxin Zhou, Qiuzhen Wu, Yuanyuan Yang, and Xukun Zhu | |
| Key Methods of Recognizing Container Number Automatically Using Video Stream in Intelligent Tally | 10 |
| Jiwei Wu, Jianming Guo, and Xueqi Li | |
| Stability of Single Species Dynamics Chained-Form System with Stage Structure via Finite-Time Control | 18 |
| Qiao Zhang, Chenglin Li, Suda Wang, and Yue Song | |
| On the Complexity of Encapsulated Agent Cell | 24 |
| Yunzhong Song, Fengzhi Dai, and Huimin Xiao | |
| Research on Novel Method of the Branch Series Resonance Recognize | 31 |
| Yongle Ai, Qunfeng Liu, and Shuai Li | |
| Simulation Research on Fault Diagnosis of CNC Machine Tools Based on Fuzzy Petri Net | 41 |
| Na Lu, Bing Xu, Zheyi Chen, Di Ren, Yuanyuan Yang, and Peng Bian | |
| Application of CMAC in Sensorless Control of PMSM | 50 |
| Qiangqiang Xu, Bulai Wang, and Yuanyuan Yang | |
| Uniform Demagnetization Fault Diagnosis in Permanent Magnet Synchronous Motors Based on the Back-EMF Gyration Radius | 58 |
| Yanjie Nie, Caixia Gao, and Ziyi Fu | |
| Dynamic Gesture Capture, Location and Tracking Based on MeanShift Algorithm | 66 |
| Shenjian Gong, Changdi Li, Junyi Hou, and Lei Yu | |

| | |
|--|-----|
| Design and Implementation of Powered Parafoil System Based on FreeRTOS | 75 |
| Lin Guo, Qinglin Sun, Zengqiang Chen, Sai Chen, and Hongchen Jia | |
| Research on the Control of Permanent Magnet Synchronous Motor Based on High Frequency Injection and Optimized Sliding Mode Controller | 83 |
| Peng Bian, Bulai Wang, Qiangqiang Xu, Yuanyuan Yang, and Na Lu | |
| Atmospheric Turbulence Rejection Control for a Geared Turbofan Engine | 91 |
| Yunxiao Liu, Zhongzhi Hu, and Jiqiang Wang | |
| Kalman-Filter-Group Based Aero-Engine Sensors Fault Diagnosis and Verification | 100 |
| Ying Liu, Jiqiang Wang, and Zhongzhi Hu | |
| Hypersonic Periodic Cruise Trajectory Optimization Based on Flexible Use of Pseudo-spectral Method | 107 |
| Haitao Gao, Mingwei Sun, and Zengqiang Chen | |
| Research on Global Stability Control Strategy of Under-Actuated Crane System Based on Equivalent Input Interference | 117 |
| Xiangsheng Liu, Zhengxin Zhou, Yunxia Jiang, Yuanyuan Yang, and Qingqing Chen | |
| Truck Size Measurement System Based on Computer Vision | 126 |
| Zhengguang Xu and Meiling Yang | |
| Dual Space Vector Control Study of Two Stage Matrix Converter Fed PMSM | 135 |
| Huida Duan and Xin Peng | |
| Visual Pose Estimation Based on the DenseNet Network | 143 |
| Lixin Lu and Weicun Zhang | |
| Kinematics Numerical Inverse Solution Algorithm for Arm Type AFP Machine Based on D-H Modeling | 151 |
| Kanzhong Yao and Yanjuan Zhu | |
| General Correntropy Based DOA Estimation for Coherent Source with Impulsive Noise | 159 |
| Sheng Guan, Jin Chen, Kaikai Li, and Yuqi Zhao | |
| Design and Research of 100KV DC Power Supply Based on SiC | 168 |
| Feng Ji and John Xu | |
| Design and Simulation Analysis of NA-PMSM Based on ANSYS | 175 |
| Yuyang Yuan, Bulai Wang, Jiale Zhong, and Xiangsheng Liu | |

| | |
|---|-----|
| A Novel Hybrid Model of Wind Speed Forecasting Based on EWT, BiLSTM, SVR Optimized by BOA in Inner Mongolia, China | 183 |
| Renqiang Ding, Wuneng Zhou, and Hangyang Cheng | |
| A Novel Bearing Intelligent Fault Diagnosis Method Based on Modified VMD and 1-D CNN | 192 |
| Xin Chen and Kaixiang Peng | |
| The Design and Implementation of Simple Corexy Structure Writing Robot | 201 |
| Rongchuan Yu, Yahang Qin, Jiansheng Peng, Tan Guo, and Xiaoping Tang | |
| The Research and Design of Intelligent Monitoring and Management Platform for River and Lake Water Environment | 214 |
| Jiabin Yu, Zhe Shen, Xiaoyi Wang, Jiping Xu, Li Wang, Qian Sun, Zhiyao Zhao, and Huiyan Zhang | |
| A Parametric Mesh Generation Method for Aero-Engine Compressor Blades | 223 |
| Min Jiang, Fengshun Lu, Long Qi, and Yufei Pang | |
| Centrifuge Fault Diagnosis Method Based on Random Forests Algorithm | 233 |
| Yanxi Li, Jing Bai, and Hezheng Li | |
| Medical Image Super-Resolution Based on the Generative Adversarial Network | 243 |
| Xu Huang, Qianyao Zhang, Guoli Wang, Xuemei Guo, and Zhonghua Li | |
| A Self-tuning Controller Design Method Based on LQG/LTR and Back Propagation Algorithm | 254 |
| Weimin Liu, Zhongzhi Hu, Jiqiang Wang, and Shengyi Liu | |
| Research on Tracking Guidance Method for On-orbit Service Autonomous Homing Based on Proportional Guidance and Trajectory Planning | 260 |
| Qi Song, Xinlong Chen, Qiang Chen, Yajing Ren, and Mengping Zhu | |
| SOC Estimation for Lithium Battery Based on Segmented Model UKF Filter | 270 |
| Xue Li, Sijia Cao, Zhishu Xu, Zongshuai Yue, and Qiaochu Guo | |
| Simplification of Gauss Spherical Point Cloud Based on K-Mean | 279 |
| Tao Ma, Xiaobin Li, and Tianyang Yu | |
| Trajectory Planning for Seven-DOF Robotic Arm Based on Seventh Degree Polynomial | 286 |
| Shuang Fang, Xianghua Ma, Jiarui Qu, Sijing Zhang, Na Lu, and Xin Zhao | |

| | |
|---|-----|
| A Hybrid Approach for Solving the Flexible Job-Shop Scheduling Problem with Considering Transporting Time and Deteriorating Effect | 295 |
| Zhuang Huang and Jianjun Yang | |
| Deep Learning and Binocular Stereovision to Achieve Fast Detection and Location of Target | 306 |
| Qingbin Wang, Yueqian Liang, Ziteng Wang, Wenyan Li, Zhiguo Jiang, and Yanjie Zhao | |
| Position Tracking and Ellipse Fitting of UAV Flying by Elliptical Trajectory | 314 |
| Qingbin Wang, Zhiqiang Jia, Yueqian Liang, Xiangfeng Wang, Zhiguo Jiang, and Yanjie Zhao | |
| A Novel Stability Analysis Method for Large-Scale Photovoltaic System | 324 |
| Xiaofan Gao, Shu Tian, Mengtian Zhang, and Yunlong Li | |
| Simulation Analysis of Nonlinear Characteristics in Complex System | 334 |
| Jiaomei Zhou, Fengshun Lu, Sumei Xiao, Long Qi, and Yufei Pang | |
| Scene Classification of Remote Sensing Images Based on ResNet | 343 |
| Lin Li and Da Li | |
| A Method of Ball Drop Point Prediction Based on Racket Pose | 352 |
| Fei Yan and Chaoli Wang | |
| Rigid-Flexible Dynamics Analysis of a Large Deformation Flexible Beam Based on Finite Volume Beams | 359 |
| Wei Han, Long Qi, and Zhongyun Xiao | |
| An Image Mosaic Algorithm Based on Improved ORB | 370 |
| Xianjin Du and Lin Li | |
| Optimization Path Planning Algorithm Based on STL File Reconstruction for Automated Fiber Placement | 379 |
| Yanjuan Zhu and Kanzhong Yao | |
| Decoupling Design Based on Active Disturbance Rejection Control Approach for Near Space Vehicle Autopilot | 388 |
| Hui Li, Zhide Yin, Fengyi Li, and Hongyu Liu | |
| Design of Speed Regulation of Ping-Pong Ball Ejector Based on RBF Neural Network PID Control | 397 |
| Min Xu and Chaoli Wang | |

| | |
|--|-----|
| Structural Analyses of Chinese Passenger Airline Network from the Perspective of Multi-layer Networks | 407 |
| Hao Wang, Shiwen Sun, Li Wang, and Chengyi Xia | |
| Observer-Based Leader Following Consensus of Nonlinear Multi-agent Systems with a Leader of Bounded Input | 419 |
| Ziyi Wang and Lin Li | |
| Target Recognition Method of Street Lamp Based on Point Cloud Data | 429 |
| Pan Cheng, Xiaobin Li, and Tianyang Yu | |
| Research on Combination and Optimization of Path Planning Algorithm Based on Bridge Crane | 437 |
| Shuai Shao and Wei Lin | |
| Extrinsic Parameter Co-calibration of a Monocular Camera and a LiDAR Using Only a Chessboard | 446 |
| Shuguang Tu, Yafei Wang, Chenyang Hu, and Yuneng Xu | |
| Detection of Load Loss by Using a New Bolt Sensor | 456 |
| Lutao Yan, Qi Wang, Jinhai Wang, and Bin Zhang | |
| Ground Verification of In-Orbit Lunar Relay Communication Satellite | 461 |
| Wenma Jin | |
| Trajectory Planning of UAV in Unknown Dynamic Environment with Deep Reinforcement Learning | 470 |
| Jia Wang, Weihong Wang, and Qian Wu | |
| Search-Based Motion Planning for Aggressive MAV with Limited Sensors FOV | 481 |
| Yu Zheng, Bowen Li, Zihe Zhao, Xu Zeng, and Xuedong Zhao | |
| An Improvement of Gravitational Search Algorithm | 490 |
| Chaohang Zheng, Hua Li, and Lin Wang | |
| Specular Reflection Separation for a Single Image Using IR-SIP | 504 |
| Yongqi Guo, Jianlin Wang, Kepeng Qiu, Liqiang Zhao, Xuesong Fu, and Rutong Wang | |
| A Calibration Method for 6-UPS Stewart Platform | 513 |
| Xuedong Jing, Yisheng Fang, and Ziteng Wang | |
| Image of Burden Point Cloud Based on Kmeans-Bayesian Segmentation with Energy Estimation | 520 |
| Jiangxue Wu, Xianzhong Chen, Qingwen Hou, Yang Liu, and Yue Meng | |

| | |
|--|-----|
| An Efficient Loop Closure Detection Based on Convolutional Autoencoder Neural Network for Autonomous Robots | 529 |
| Jian Wang and Sen Zhang | |
| Quasi-periodic J_2 Invariant Relative Orbits for Formation Flying | 542 |
| Qiang Luo, Yaru Zheng, and Ming Xu | |
| Stereo Visual Inertial Odometry for Unmanned Aerial Vehicle Autonomous Flight | 550 |
| Quanpan Liu, Zhengjie Wang, and Huan Wang | |
| A Traffic Network Congestion Diffusion Model Based on Gene Regulation Mechanism | 563 |
| Xiangwei Wu, Ning Huang, and Kan Xu | |
| Search-Based Motion Planning for MAVs Decentralized Formation Control | 573 |
| Yu Zheng, Yunsong Ge, Longhong Wang, Xu Zeng, and Xuedong Zhao | |
| Researching the Monitoring Scheme Built for the Characteristic Parameters of Incineration Equipment | 582 |
| Fengguang Huang, Lin Tian, Daoqing Hu, Wei Wang, Bosen Li, Qi Lu, Xiaolin Xing, Xi Chen, Xiaoman Zhai, and Yansha Luo | |
| Finite-Time Consensus Control for the Second-Order Multi-agent Systems with Communication Delays | 589 |
| Yan Cui and Qinghua Li | |
| Social Network Emergency Incident Portrait Based on Attention Mechanism | 601 |
| Jiannan Chen, Junping Du, Lei Shi, Zhe Xue, and Feifei Kou | |
| Network Attack Detection Method Based on Convolutional Neural Network | 610 |
| Nan Ding, Yang Liu, Yimin Fan, and Dangyang Jie | |
| Study on Path Planning for Family Service Robot Based on Improved Genetic Algorithm | 621 |
| Xuedong Jing, Ya-nan Chen, and Yuquan Xue | |
| Study on Dose Distribution Prediction of Esophageal Cancer Patients Using U-Net Model | 632 |
| Jing Wang, Ronghu Mao, Jiwei Liu, and Jianfei Liu | |
| Primary Equipment Fault Diagnosis Strategy for Box Substation Based on PSO-BP Neural Network | 641 |
| Ao Hou, Fuzhong Wang, and Jili Ren | |

| | |
|---|-----|
| Defect-Oriented Test Case Prioritization | 651 |
| Jinfeng Li, Yan Zhang, and Jilong Bian | |
| A Positioning Method Using Teaching-Learning Based Optimization for GNSS | 660 |
| Jing Zhang, Jinheng Zhang, Zihang Ni, and Lan Cheng | |
| Opinion Leader Mining Algorithm Based on Double-Graph Model | 670 |
| Xunpu Yuan, Junping Du, Wanqiu Cui, Zhe Xue, and Feifei Kou | |
| Model Reference Adaptive Control Based on Transformed Scalar Error for the Three-Tank Level Control System | 679 |
| Pei Zhou and Xuehui Gao | |
| Research on Self-compensation Technology of Strapdown Inertial Navigation System | 689 |
| Jianqiang Zheng, Lin Gao, Qinghua Ma, and Shujun Yang | |
| Scalable Observer Bandwidth Based Active Disturbance Rejection Control of Brushless DC Motor | 695 |
| Wei Wei, Zhiyuan Zhang, and Min Zuo | |
| Tool Wear State Monitoring Based on Long-Term and Short-Term Memory Neural Network | 703 |
| Zhu Xiang and Xie Feng | |
| A New V/f Control Method for Permanent Magnet Linear Synchronous Motor Drives | 713 |
| Leilei Cui, Hongwei Zhang, and Mingren Wang | |
| Multi-branch Input Structure for Pyramid Scene Parsing Network | 724 |
| Chuanya Wang and Jin Chen | |
| Author Index | 733 |



Research on Vector Control of PMSM Based on Fuzzy PI

Xiangsheng Liu^(✉), Zhengxin Zhou, Qiuzhen Wu, Yuanyuan Yang,
and Xukun Zhu

Sanda University, Shanghai, China
lxssit@163.com

Abstract. The paper puts forward fuzzy PI control adopted in speed loop, while in the permanent magnet synchronous motor speed regulating system, PI control is usually used to get control of speed loop and showing weak control over power, robustness and anti-interference. Fuzzy control is the simulation of human intelligence to achieve the control effect required by the control system. At the same time, to build a model of reference self-adaptability according to the principle of non-position, to test and verify the idea by using the MATLAB Platform, the results of simulation show that the model of reference self-adaptability having advantage in fast response, small overshoot, strong robustness and anti-interference ability and showing better dynamic and static characteristics, comparing with the traditional PI model.

Keywords: Fuzzy control · Parameter estimation · PMSM · Model Reference Adaptive System

1 Introduction

The permanent magnet synchronous motor (PMSM) is widely used in the day life for its characteristics, such as small size, high efficiency, simple rotor structure, reliable operation, wide speed range and so on. PMSM has been widely used in life. In recent years, the vector control technology, improving dynamic performance and control-responding speed efficiently, is widely used in PMSM system. Then the control of PMSM is similar to DC motor in speed regulation performance. However, vector control has disadvantages such as mounting mechanical sensor to feed back the position information of the rotor's rotational speed and rotational angle. The installed mechanical sensor also brings many difficult problems. For example, under high temperature and hige pressure working conditions, the position information of the rotor speed and rotation angle provided by the rotor is not trustworthy; the cost of the motor control system is increased and the maintenance is difficult; The stability is reduced, so the research without position sensor control has become a hot spot [1].

The PI controllers are used in speed outer-loop in the traditional PMSM speed control system. The PI controllers are linear control system, if nonlinear characteristics or parameters of controlled object has changes, the PI controllers of linear constant parameters can be inconsistent with the performance at the time of design, and the robustness is often unsatisfactory. PMSM is a typical nonlinear system. The fuzzy

controllers have received more and more attention and has become a research hotspot at present, in order to further improve the rapidity, stability and robustness of the permanent magnet synchronous motor speed control system [2, 3]. The fuzzy controllers need not to establish a mathematical model of the controlled object, and has strong adaptability to the time lag, nonlinearity and time-varying of the controlled object. The fuzzy controllers are nonlinear control system, different from the conventional PI controller, which has better adaptability to the parameter variation or nonlinearity of the control object, and has stronger suppression ability against interference and noise.

At present, Sensorless research is mainly divided into low speed and medium speed [4]. In low speed, high frequency injection method is generally used, but this method is used for motors with salient pole effect, and the scope of use is limited. In the case of medium and high speed, the No-position control method includes flux linkage estimation method, model reference adaptation method [5], extended Kalman method, and sliding mode variable structure method. The extended Kalman method has a large amount of computation, and the sliding mode variable structure method has the problem of bucket vibration. In comparison, the model reference adaptation method has good stability [6–10], which is not affected by the system control strategy.

The simulation results based on MATLAB demonstrate that the Fuzzy PI controller has good control performance for the speed and torque control.

2 The Basic Principle of Model Reference Adaptive Algorithm (MRAS)

The main idea of MRAS is to use the equation containing the parameters to be estimated as an adjustable model, and the equation without unknown parameters as the reference model, and the two models have the same output in the physical sense. Taking the output value of the reference model as the ideal response, the parameters of the adjustable model are adjusted in real time through a suitable adaptive law. Let the output difference of the two models approach zero, and the output of the control object can track the reference model output [11]. The MRAS schematic is shown in Fig. 1.

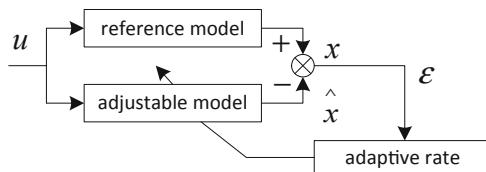


Fig. 1. MRAS schematic structure

In Fig. 1, reference model output is, adjustable model output is, The output of the controlled object is compared with the output x of the reference model, and the difference is sent to the adaptive mechanism for feedback adjustment. Adjusting the parameters of the adjustable model through feedback. Let x and consistent, that is, the value of tends to zero.

3 Mathematical Model of PMSM

The voltage equation of PMSM in the dq coordinate system is:

$$u_d = R_s i_d + L_s p i_d - \omega_r L_s i_q \quad (1)$$

$$u_q = R_s i_q + L_s p i_q - \omega_r L_s i_d + \omega_r \psi_f \quad (2)$$

In the formula: u_d, u_q, i_d, i_q is the voltage and current of the stator under the dq axis, R_s is the stator phase resistance, L_s is the stator per phase inductance, ω_r is speed, ψ_f is rotor permanent magnet flux linkage, p is differential operator, the simplification of (1) and (2) gives (3) and (4) as the reference mode.

$$p i_d + \frac{p \psi_f}{L_s} = -\frac{R_s}{L_s} (i_d + \frac{\psi_f}{L_s}) + \omega_r i_q + \frac{1}{L_s} (u_d + R_s \frac{\psi_f}{L_s}) \quad (3)$$

$$p i_d = -\omega_r (i_d + \frac{\psi_f}{L_s}) - \frac{R_s}{L_s} i_q + \frac{1}{L_s} u_d \quad (4)$$

$i'_d = i_d + \frac{\psi_f}{L_s}$, $i'q = iq$, $u'_d = u_d + R_s \frac{\psi_f}{L_s}$, $u'_q = u_q$, so the reference model is

$$p i'_d = -\frac{R_s}{L_s} i'_d + \omega_r i'_d + \frac{1}{L_s} u'_d \quad (5)$$

$$p i'_q = -\omega_r i'_d - \frac{R_s}{L_s} i'_q + \frac{1}{L_s} u'_q \quad (6)$$

Constructing model reference adaptive tunable model from Eqs. (5) and (6)

$$p \tilde{i}'_d = -\frac{R_s}{L_s} \tilde{i}'_d + \tilde{\omega}_r \tilde{i}'_q + \frac{1}{L_s} \tilde{u}'_d \quad (7)$$

$$p \tilde{i}'_q = -\tilde{\omega}_r \tilde{i}'_d - \frac{R_s}{L_s} \tilde{i}'_q + \frac{1}{L_s} \tilde{u}'_q \quad (8)$$

From the generalized error $\varepsilon = i'_s - \tilde{i}'_s$, we know that, Eqs. (5)–(7), (6)–(8) can obtain the model reference adaptive formula:

$$p \varepsilon_d = -\frac{R_s}{L_s} \varepsilon_d + \omega_r \varepsilon_q + (\omega_r - \tilde{\omega}_r) p \tilde{i}'_q \quad (9)$$

$$p \varepsilon_q = -\frac{R_s}{L_s} \varepsilon_q - \omega_r \varepsilon_d - (\omega_r - \tilde{\omega}_r) p \tilde{i}'_d \quad (10)$$

In the formula $\varepsilon_d = i'_d - \tilde{i}'_d$, $\varepsilon_q = i'_q - \tilde{i}'_q$.

The parameters in MRAS generally use the structure of proportional integral, which can be set by Eqs. (9), (10) and Popov ultra-stability theory:

$$\tilde{\omega}_r = \int_0^t k_i (\tilde{i}_d \tilde{i}'_q - \tilde{i}'_d \tilde{i}_q) d\tau + k_p (\tilde{i}'_d \tilde{i}'_q - \tilde{i}_d \tilde{i}'_q) + \tilde{\omega}_r(0) \quad (11)$$

From $\dot{i}'_d = i_d + \frac{\psi_r}{L_s}$, $\dot{i}'_q = i_q$ and (11), the speed is represented by i_d , i_q , \tilde{i}_d , \tilde{i}'_q .

4 Fuzzy PI Regulator Design

Fuzzy control is the fuzzy set theory, which expresses the experience of human experts in an algorithmic language that can be recognized by computers, thus simulating human intelligence and achieving the control effect required by the control system.

In Fig. 2, the fuzzy controller is shown by a dotted line. It includes an inference engine, a knowledge base, etc. The knowledge base includes a database and a rule base. The fuzzy control system is a closed-loop control system.

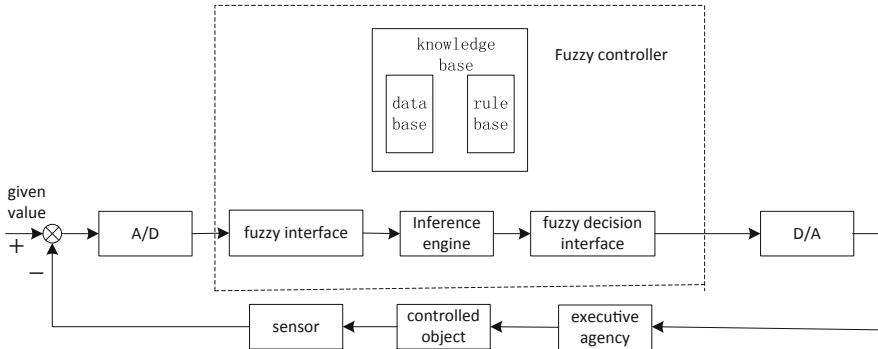


Fig. 2. Closed loop system block diagram of fuzzy control

Figure 3 shows that the input of the fuzzy controller is the derivative of the velocity deviation and the velocity deviation.

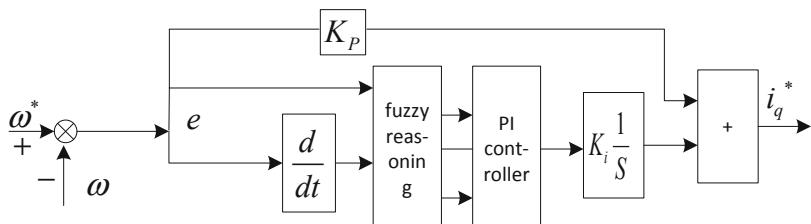


Fig. 3. Fuzzy control schematic

In Fig. 3, deviation e is the deviation between the given value ω^* and the feedback value u . The deviation e and the deviation change rate de/dt are differentially controlled by fuzzy inference to control the PI controller. The output of the PI controller is integrated and e passes through the KP. The gain is used as the d-axis current.

The fuzzy controller is designed by MATLAB fuzzy logic system toolbox. The fuzzy controller is designed, and Fig. 4 is the membership function of the system input deviation e and the deviation change rate de/dt . The fuzzy inference rules are shown in Table 1.

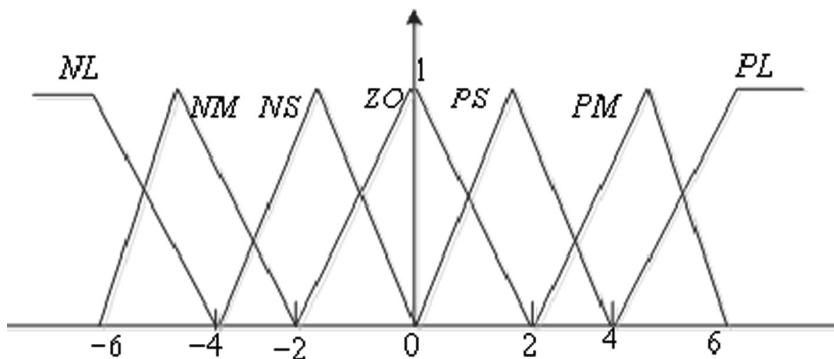


Fig. 4. Membership functions of input and output variables of the fuzzy controller

Table 1. Fuzzy controller RULES TABLE

| e de | NL | NM | NS | ZO | PS | PM | PL |
|---------|----|----|----|----|----|----|----|
| NL | NL | NL | NL | NM | NM | NS | ZO |
| NM | NL | NL | NM | NM | NS | ZO | PS |
| NS | NL | NM | NM | NS | ZO | PS | PM |
| ZO | NM | NM | NS | ZO | PS | PM | PM |
| PS | NM | NS | ZO | PS | PM | PM | PL |
| PM | NS | ZO | PS | PM | PM | PL | PL |
| PL | ZO | PS | PM | PM | PL | PL | PL |

5 System Simulation and Analysis

5.1 PMSM Sensorless Vector Control System

Figure 5 shows that the PMSM vector system is composed of a speed outer loop and a current inner loop. The value of the speed estimated by MRAS is compared with the set value, and the deviation value is sent to the fuzzy PI controller to obtain the cross-axis current. Let the d-axis current be equal to 0, compare it with the feedback value, and then output the dq axis voltage via PI. The rotation angle is estimated as the reference angle, and the coordinate rotation is performed to obtain the two-phase static voltage. The switching signal is generated by SVPWM modulation, and the permanent magnet synchronous motor is controlled by the inverter.

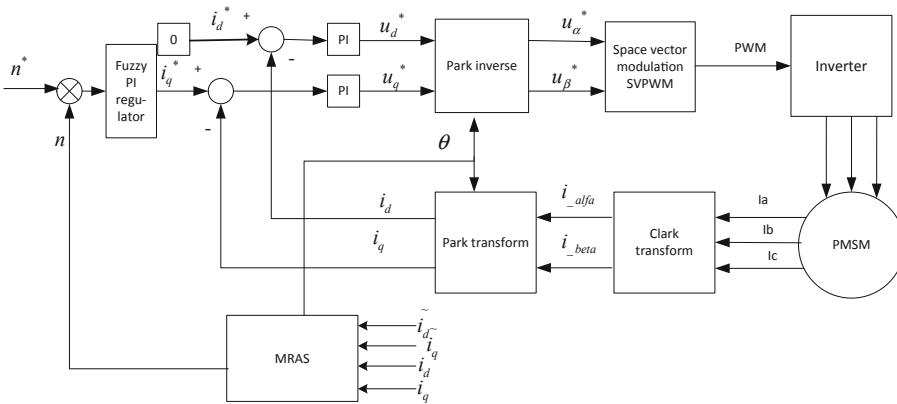
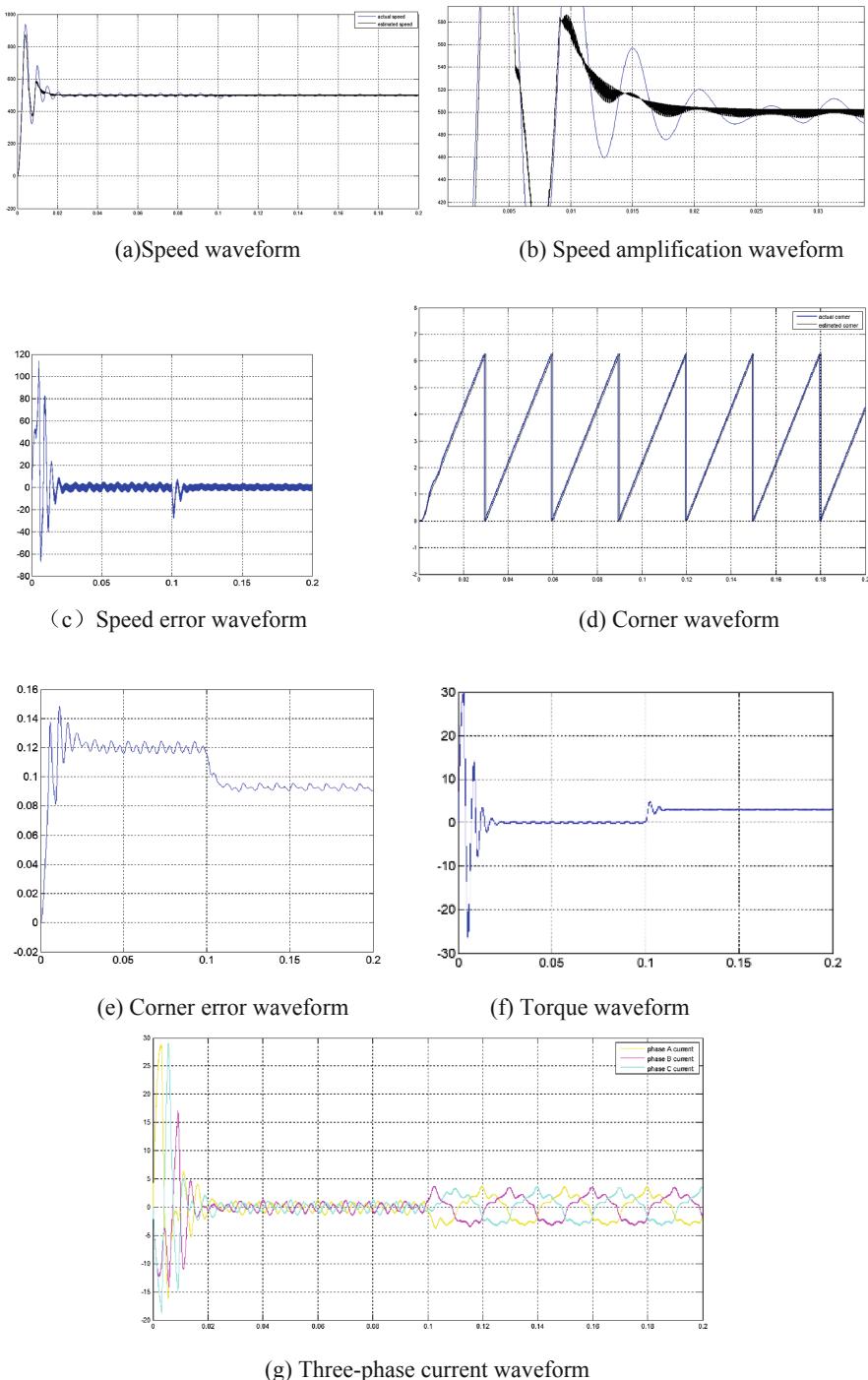


Fig. 5. Vector control block diagram of PMSM

5.2 Simulation Results and Analysis

The controller designed in this paper is simulated and analyzed. The main parameters of the motor are: the pole pair is 4, the stator resistance is 2.87Ω , the inductance L_d and L_q is 8.5 mH , and the flux ψ_f is 0.17 Wb . The load torque value at 0.1 s is 3 N m .

When the speed loop is traditional PI control, the simulation waveform is shown in Fig. 6:

**Fig. 6.** Traditional PI control simulation waveform

When the speed loop is fuzzy PI control, the simulation waveform is shown in Fig. 7:

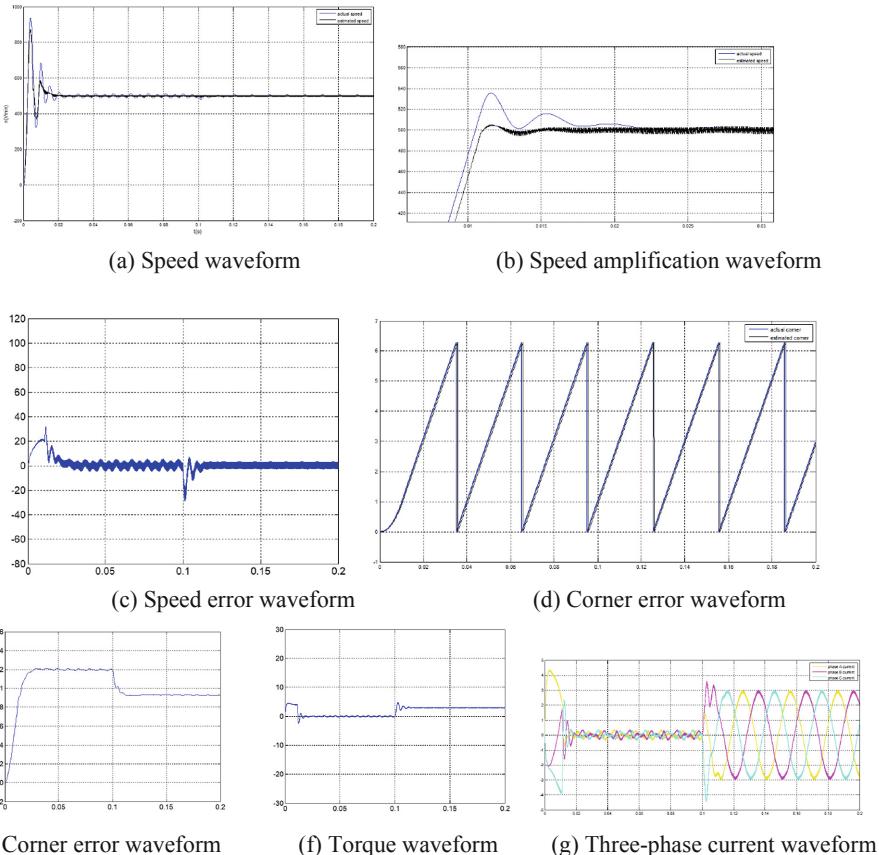


Fig. 7. Fuzzy PI control simulation waveform

By comparing the simulated waveforms of (a) and (b) in Figs. 6 and 7, it is concluded that when the speed loop is a traditional PI, the overshoot of the system is about 80%, and when the speed loop is a fuzzy PI, the overshoot is about 7%. Overshoot is significantly reduced. In the traditional PI, the speed is stable at 0.02 s, while the fuzzy PI control is almost stable at 0.0175 s; By comparing the simulated waveforms of (a) in Figs. 6 and 7, the error of the speed obtained by the fuzzy PI is smaller than that of the conventional PI; By comparing the simulated waveforms of (d) and (e) in Figs. 6 and 7, The corner error of fuzzy PI is more stable than traditional PI; By comparing the simulated waveforms of (f) in Figs. 6 and 7, The torque of fuzzy PI is smaller than that of traditional PI; By comparing the simulated waveforms of (g) in Figs. 6 and 7, The three-phase current waveform of the fuzzy PI is closer to the sine wave than the conventional PI.

6 Conclusions

In this paper, a fuzzy model reference adaptive control system is proposed. The speed regulator replaces the traditional PI with fuzzy PI. The fuzzy PI model reference adaptation is used to accelerate the response speed of the system. The simulation model was constructed by deriving the corner and speed formula. The simulation results show that the fuzzy PI model reference adaptive can track the rotation speed and rotation angle of the motor quickly and accurately. When the motor starts, the fuzzy PI speed regulator has small overshoot compared with the traditional PI regulator, small torque ripple, strong robustness, good anti-interference ability, high control accuracy, good dynamic performance, and fuzzy PI designed at the same time. At the same time, fuzzy PI model reference adaptation is easy to implement in hardware and software systems, and has certain practicability.

References

1. Wang F, Tian J (2016) The moment of inertia identification of permanent magnet synchronous motor based on improved model reference adaptive algorithm. *Motor Control Appl* 43(08):63–67+87
2. Qi L (2012) Robust control of T-S model and its application in speed control of PMLSM system. Shenyang University of Technology
3. Lv H, Ye Y (2012) Position sensorless control of low speed permanent magnet motor based on MRAS. *Mech Electr Eng* 29(3):307–310
4. Li C, Yin W, Feng X et al (2012) Stepless speed regulation system of brushless DC motor based on fuzzy adaptive PI control. *Mech Electr Eng* 29(1):49–52
5. Nait Seghir A, Henni T, Azira M (2013) Fuzzy and adaptive fuzzy PI controller based Vector control for permanent magnet synchronous motor. In: 2013 10th IEEE international conference on networking, sensing and control (ICNSC), Evry, pp 491–496
6. Omara AM, Sleptsov M, Diab AAZ (2018) Cascaded fuzzy logic based direct torque control of interior permanent magnet synchronous motor for variable speed electric drive systems. In: 2018 25th international workshop on electric drives: optimization in control of electric drives (IWED), Moscow, pp 1–6
7. Mohamed AS, Zaky MS, Ashraf SZ et al (2011) Comparative study of sensorless control methods of PMSM drives. *Innovative Syst Des Eng* 2(5):44–67
8. Aygün H, Gökdag M, Aktas M, Cernat M (2014) A novel sensorless field oriented controller for permanent magnet synchronous motors. In: 2014 IEEE 23rd international symposium on industrial electronics (ISIE), Istanbul, pp 715–720
9. Wang Q, Zhang C, Xing Z (2008) Variable structure MRAS speed identification for permanent magnet synchronous motor. *Proc CSEE* 28(9):71–75 (in Chinese)
10. Guo Y, Huang L, Qiu Y, et al (2000) Inertia identification and auto-tuning of induction motor using MRAS. In: Proceeding of the 3rd power electronics and motion control conference, Shanghai, China, August 2000, pp 1006–1011
11. Guo X (2013) Research on rotor speed identification of permanent magnet synchronous motor based on improved MRAS. Hunan University, Hunan



Key Methods of Recognizing Container Number Automatically Using Video Stream in Intelligent Tally

Jiwei Wu, Jianming Guo^(✉), and Xueqi Li

School of Automation, Wuhan University of Technology, Wuhan 430000, China
jmguo62@163.com

Abstract. An automatic identification algorithm of container number based on video stream is proposed in this paper. The algorithm is designed to use Selective Search and frame difference method to locate the front and rear sides of the container and the land and sea side of the container. And then the region of the number is located through noise filtering. Ultimately, the complete container number is identified by the template matching and convolutional neural network (CNN). The algorithm can realize the real-time positioning of the container surface from the four-side video stream of the container docking as well as the identification of the container number.

Keywords: Intelligent tally · Number identification · Video stream

1 Introduction

Intelligent tally achieves automatic collection of dock data, intelligent identification and processing of containers through computer vision and artificial intelligent technology, reducing the problem of harsh environment and high risk of manual on-site tally operation, and improving the efficiency of tally and the level of automation [1]. It is common to realize the identification of the container number by computer vision methods and artificial intelligent in seaports. However, most of them use the height information of the current spreader to trigger the camera's capturing image. The method of capturing the container by preset camera capture time is often indeterminate. If there is an error in the time, the early shooting, late shooting, missing shooting and false shooting will appear in the sample, which directly affect the number identification's result.

Considering the existed problems, our paper proposes a container positioning algorithm based on video frames to obtain the images of the four sides of the containers, so as to locate the whole container number. The use of video frames solves the problems of missing containers and defective container numbers due to errors in the preset capture time, and can directly filter most of the background area during the container acquisition. It is more conducive to locate and identify the number from the acquired container.

2 Container Positioning Based on Video Stream

In the docking tally video, the complete container surface will appear in a single frame image from the container's front and rear sides video streams. However, the sea and land sides surface images cannot be obtained completely from a single-frame picture due to the excessive sea surface area of containers. According to the characteristics of two types of surfaces, they are divided into two parts: the front and rear side surface positioning and the sea and land side surface positioning. In the front and rear side surface positioning algorithms, the intact box surface is as the discriminating basis of the whole container number. Differently, the intact container number is judged from the moment that the upper and lower edges of the box surface appear simultaneously in the sea and land side surface positioning algorithm.

2.1 Front and Rear Side Container Positioning

The front and rear side box surface positioning algorithm is divided into three parts: rough positioning, verification container surface result and precise positioning. The positioning results is shown in Fig. 1.

During the rough positioning, the Selective Search [2] algorithm is used to obtain the ROI first, the candidate regions satisfying the conditions are filtered and merged according to the various features. Then the cascade classifier is used for classification [3]. At last, the area with the highest credibility is screened out as the rough positioning result of the front and back side box surfaces, as is shown in Fig. 1(a). The rough positioning result of the container surface obtained in Fig. 1(a) is used to locate the four corners of the container surface by using the multi-scale detection algorithm. Then the result of the coarse positioning of the container surface is corrected by the corner piece



(a) Rough positioning of the container box



(b) Positioning of the corner



(c) Accurate positioning

Fig. 1. Positioning results from the front and rear side of the container

positioning result, which is shown in Fig. 1(b). In order to avoid missed judgments as much as possible, the container positioning results of consecutive frames are analyzed comprehensively, and the most reliable positioning result from the container image positioned in multiple frames is selected as the front and rear side of the container. The positioning result is shown in Fig. 1(c).

2.2 The Sea and Land Side Container Positioning

In a single-frame of the container's sea and land side video streams, it is often impossible to obtain a complete box image from a single frame picture, which brings great difficulties in selection of video frames and container number positioning.

It has been found that when the container enters the image range, the inter-frame gradient differential lateral projection waveform is more sensitive to the container surface. For the point $I_t(x, y)$ in the grayscale image at time t, the deviations value $D_t(x, y)$ is:

$$D_t(x, y) = \begin{cases} \tau |I_t(x, y) - I_{t-1}(x, y)|, & \text{if } |I_t(x, y) - I_{t-1}(x, y)| \leq \sigma \\ |I_t(x, y) - I_{t-1}(x, y)| & \text{others} \end{cases} \quad (1)$$

σ is a threshold value considering factors such as illumination, and τ is a coefficient constant smaller than 1. The sum of $D_t(x, y)$ in the horizontal direction is counted in order to get the gradient deviations lateral projection waveform P_t at time t:

$$P_t(y) = \sum_x D_t(x, y) \quad (2)$$

When the container does not appear, the waveform diagram is relatively stable, as shown in Fig. 2(a). When the lower edge of the container appears in the image, the waveform shows a distinct peak at the corresponding edge of the container surface, such as Fig. 2(b). When the container continues to descend but the upper edge of the container still does not appear in the image, peak-cluster appears near the waveform at the container's lower edge due to the rust, characters and other interference on the surface of the container, which is shown in Fig. 2(c). When the upper and lower edges of the container enter the monitoring image at the same time, the peak-cluster will appear near the corresponding waveform at the upper edge, which is shown in Fig. 2(f). As the container continues to be lowered, the number of peaks is converted from a double peak group to a single peak group until the waveform image tends to be stable, which is shown in Fig. 2(i).

From the analysis, the inter-frame difference vector can reflect the actual situation of the container surface. Whether the container number in the current frame is complete is predicted by discriminating whether the upper and lower edges of the box surface appear at the same time. In the sea and land side surface positioning algorithm, the cumulative value of the inter-frame gradient difference in the horizontal direction is used to form the feature vector of the current frame, and the feature vector is classified by SVM. The classification result includes two types, the current frame available and the current frame not available. Afterwards, the multi-frame discrimination result is

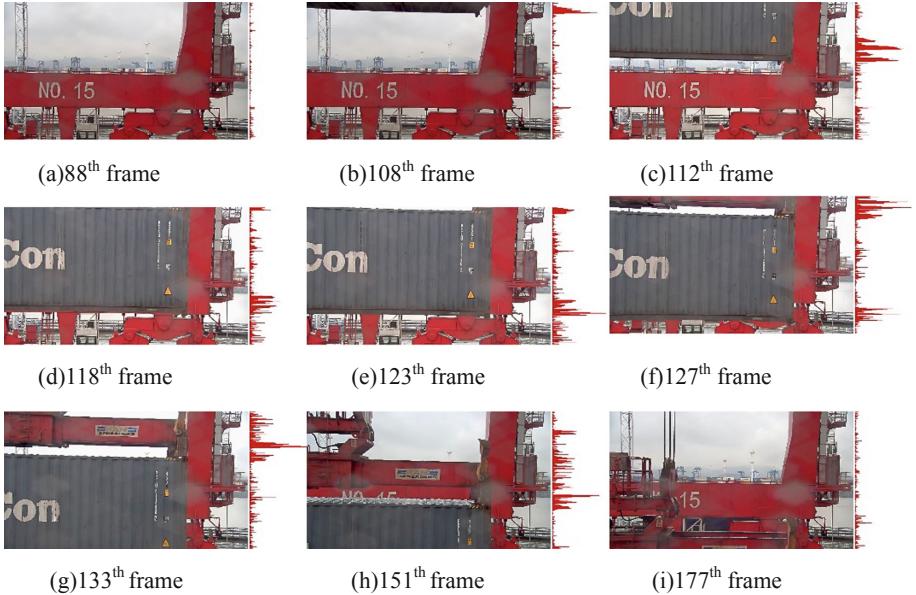


Fig. 2. Container sea and land sides frames and corresponding feature vector waveforms

used to correct and filter the classification results, and the appropriate frame is filtered out as a sample for identification of the container number in the next stage.

3 The Location and Identification of Container Number

According to the complexity of the docking tally scene and the diversity of the container number arrangement, the number positioning algorithm of containers is divided into two parts: container image preprocessing and container number positioning.

3.1 Noise Filtering in Container Images

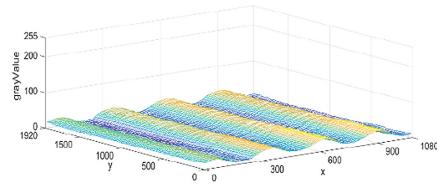
Two kinds of contaminated samples are often found. One is that the horizontal lateral stripes interfere caused by nighttime fill light superimposed on the sample image, and the other is that the contrast of the image is low due to changes in illumination. It is showed in Fig. 3(a) that the horizontal stripe noise which affects the location of the container number. The horizontal stripe noise will cause the local horizontal gray value to be greatly improved. The corresponding three-dimensional diagram is shown in Fig. 3(b-1). It can be assumed that the increased gradation value is a partial component of the fringe noise. The Fourier transform is used to change the image of the positioning area to the frequency domain, and the frequency value $\varphi(x)$ of the longitudinal direction of the image is calculated:

$$\varphi(x) = \frac{1}{h} \sum_y I(x, y) \quad (3)$$

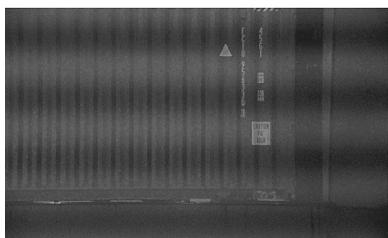
h is the width of the image. The frequency corresponding to the stripe noise can be obtained, and an adaptive algorithm is used to select an appropriate filter radius. The Gaussian low-pass filter can effectively filter the stripe noise. A horizontal stripe noise adaptive filtering algorithm based on noise gray estimation is proposed in this paper to filter out the horizontal stripe noise. First, the horizontal projection feature vector of the 72-dimensional image is extracted and the SVM is used to discriminate whether the sample contains horizontal streak noise. The three-dimensional map corresponding to the noise estimation image and the filtered image is shown in Fig. 3(b-2) and (b-3). The adaptive filtering algorithm can filter out the horizontal stripe noise on the basis of ensuring the character details of the container character area, and overcome the influence of the noise intensity change on the filtering effect. It can also enhance the contrast of the character area image.



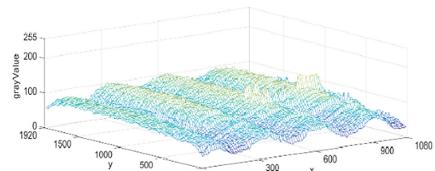
(a-1) pre image



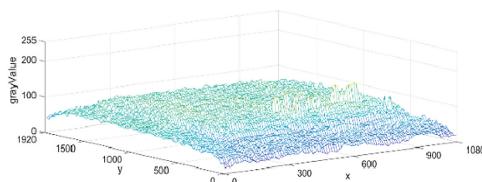
(b-1)3D diagram before filtering



(a-2)filtered image



(b-2)Estimated 3D diagram of noise



(b-3) 3D diagram corresponding to the filtered image

Fig. 3. Comparison of adaptive filtering effects

3.2 Number Positioning of Containers

In this paper, the edge feature and MSER [4, 5] feature of the image are used in the container number localization algorithm. In the actual application process, the type of the extracted feature image is selected according to the container surface color of the current image and the orientation information of the image acquisition. Table 1 shows the average time taken by the algorithm to locate a single sample, and the time is controlled within 0.9 s, which meets the real-time requirements of practical application.

Table 1. Average time taken by each algorithm to locate

| Location algorithm | Algorithm based on edge image | Algorithm based on MSER image | Using the combined strategy |
|--------------------|-------------------------------|-------------------------------|-----------------------------|
| Average time (s) | 0.6135 | 0.7632 | 0.8983 |

3.3 Number Identification of Containers

In this paper, two character recognition algorithms are used to identify the container number characters, and the recognition results of the two algorithms are combined by the credibility of the two algorithms and the check rule of the box number.

- (1) Container number character recognition based on template matching [6]. In this paper, character multi-feature extraction method is adopted in character extraction based on template matching in character recognition.
- (2) Container number character recognition based on convolutional neural network (CNN) [7]. In this paper, the convolutional neural network LeNet-5 [8] is used to identify the container number characters based on the convolutional neural network (Fig. 4).

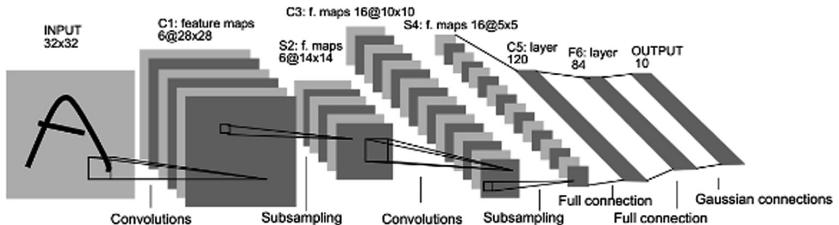


Fig. 4. Architecture of Lenet-5, a convolutional neural network for character recognition

- (3) Container number character combination verification [9]. In this paper, images of 4 sides of a box are acquired in each container, and each side has 5 sets of parameters for positioning and segmentation. After the template matching algorithm and CNN algorithm are used respectively, the container number will correspond up to 40 recognition results. Even if each recognition result is fallacious, it is possible to obtain the correct number recognition result by combining them.

Table 2. Container number accuracy

| Accuracy | Template matching | CNN | Combined strategy |
|------------------|-------------------|--------|-------------------|
| Container number | 0.9724 | 0.9893 | 0.9934 |
| Container type | 0.9802 | 0.9880 | 0.9955 |
| Combined | 0.9661 | 0.9868 | 0.9930 |

As shown in Table 2, According to the analysis, the comprehensive correct rate of the container number identification under the combination strategy of this paper is 99.3%, which is obviously higher than the recognition accuracy of the template matching algorithm and the CNN algorithm.

4 Experiment

The container number identification algorithm in the process of docking the container based on video stream in this paper is mainly divided into three parts: container positioning, container number character location and container number identification. According to the basic process of container identification, 2,428 sets of 3 different bridge cranes in a domestic seaport are selected for testing. A total of 10,568 container samples are obtained through the container positioning algorithm. The experimental test part of this paper is divided into two parts: daytime sample test and nighttime sample test. The test results are shown in Table 3.

Table 3. Accumulated correct rate of each part of the algorithm

| Category | Daytime | | Nighttime sample | | All | |
|-----------------------|-----------|-------------|------------------|-------------|-----------|-------------|
| | Each part | Accumulated | Each part | Accumulated | Each part | Accumulated |
| Container positioning | 0.9895 | 0.9895 | 0.981 | 0.981 | 0.9860 | 0.9860 |
| Number positioning | 0.9908 | 0.9804 | 0.959 | 0.942 | 0.9783 | 0.9646 |
| Number identification | 0.9949 | 0.9754 | 0.994 | 0.936 | 0.9944 | 0.9592 |

It can be seen from the analysis in Table 3 that the accuracy of the container box surface positioning algorithm is 98.6%, which is obviously better than obtaining the container surface by camera capturing. The correct rate of the container number localization algorithm and the container number identification algorithm reaches 97.8% and 99.4%, respectively, which verified the effectiveness of the localization and recognition algorithm.

5 Conclusion

The experimental results show that our algorithm based on video stream is robust to the positioning of the container, the positioning and identification of the container number in different environments. It is an effective and efficient identification algorithm of the

container number. However, the accuracy at night is significantly lower than that of the daytime. The main reason is that the quality of the nighttime samples is easily disturbed by factors such as illumination. This is the future study of the algorithm in this paper that needs to be improved.

References

1. Huang S, Weng M, Shi Y, et al (2018) Container box number identification based on computer vision. *Port Handling* 25(1):1–4
2. Uijlings JR, van de Sande KE, Gevers T, Smeulders AW (2013) Selective search for object recognition. *Int J Comput Vision (IJCV)* 104(2):154–171
3. Cheng WC, Jhan DM (2011) A cascade classifier using Adaboost algorithm and support vector machine for pedestrian detection 42(1):1430–1435
4. Wu W, Liu Z, Chen M et al (2012) An automated vision system for container-code recognition. *Expert Syst Appl* 39(3):2842–2855
5. Weng M, Liu Q, Guo J (2017) MSER and connected domain analysis based algorithm for container code locating process. In: International conference on industrial informatics - computing technology, intelligent technology, industrial information integration. IEEE Computer Society, pp 83–86
6. Wen Y, Lu Y, Yan J, Zhou Z, Von Deneen KM, Shi P (2011) An algorithm for license plate recognition applied to intelligent transportation system. *IEEE Trans Intell Transp Syst* 12:830–845
7. Mei L, Guo J, Liu Q, et al (2017) A novel framework for container code-character recognition based on deep learning and template matching. In: International conference on industrial informatics - computing technology, intelligent technology, industrial information integration. IEEE, pp 78–82
8. LeCun Y, Bottou L, Bengio Y, et al (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86:2278–2324
9. Yoon Y, Ban KD, Yoon H et al (2016) Automatic container code recognition from multiple views. *ETRI J* 38(4):767–775



Stability of Single Species Dynamics Chained-Form System with Stage Structure via Finite-Time Control

Qiao Zhang^(✉), Chenglin Li, Suda Wang, and Yue Song

School of Hohai University, Changzhou 213022, China
996266928@qq.com

Abstract. This paper presents a finite-time control scheme to stabilize the single species dynamics model with stage structure. To begin with, the single species dynamics model with stage structure is linearized by the Jacobian matrix with the equilibrium. The main advantage of the paper is the finite-time controller with disturbance proposed based on the stability control theory, transforming the dynamics model into a chained-form system via rigorous convergence and stability analysis. Finally, we demonstrate the performance of the proposed finite-time control approach according to the further simulation results.

Keywords: Species dynamics model · Chained-form system · Finite-time control

1 Introduction

As is known to us, individuals in almost all species live several different stages through life history, for the simplicity of our model, we only consider the stage structure, one is juvenile (immature) stage and the other is adult (mature) stage. Actually, survival, reproduction, and development in a population and other common rates always depend on stage structure [1], while individuals in two stage perform similarly in characteristics. Apart from that, strong interaction relationship exists in juvenile population and the adult population, which has a great influence on the persistence and extinction of the species. Furthermore, Zhang has done a series of studies around the dynamics of species with stage structure in [2].

By the economic theory of a common-property resource [3], there is a phenomenon of bio-economic equilibrium when the population of immature and mature both approach to zero. Recently, finite-time control is widely used in various fields, especially in target tracking. According to Wang et al. [4], a finite-time nonlinear controller satisfying the distance constraints were proposed for vehicles to track a moving target with desired distance and bearing angles. In terms of tracking issues with chained-form systems, Chen et al. [5–7] has conducted a series of studies, finite-time tracking control for extended nonholonomic chained-form systems with external disturbance and parametric uncertainty has been considered, which also applies to the proposed system.

As for the stability analysis, a kind of stability analysis based on the full nonlinear dynamics stability analysis has been put forward by Sun in [8], which gives the paper a

lot instructive help. For the sake of investigating the dynamics of such species model, we introduce a finite-time controller with disturbance to guarantee perfect stabilization, at where disturbance can be policy factors et al.

In this paper, we intend to consider the stage structure of species with a finite-time control scheme, which is proved possible to stabilize the single species dynamics model with stage structure. The main work of this article is organized as the following points:

- (1) Firstly, the single species dynamics model with stage structure is linearized by the Jacobian matrix with the equilibrium.
- (2) Secondly, the finite-time controller with disturbance proposed based on the stability control theory, transforming the dynamics model into a chained-form system via rigorous convergence and stability analysis.

2 System Description

Considering the stage structure of species, a model describing the dynamics form proposed by Zhang can be expressed as follows:

$$\begin{cases} \dot{x}_1 = \alpha x_2 - \gamma_1 x_1 - \beta x_1 - \eta x_1^2 \\ \dot{x}_2 = \beta x_1 - \gamma_2 x_2 \end{cases} \quad (1)$$

where x_1 and x_2 represent population density of immature species and mature species, respectively; $\alpha, \beta, \eta, \gamma_1, \gamma_2$ are positive constants, α denotes the intrinsic growth rate, β denotes transition rate of the immature population, γ_1, γ_2 represent the death rate of the immature and mature population, respectively; and the growth of the immature population is restricted by population density $-\eta x_1^2$.

In order to discuss the stabilization of system (1), at first, in order to convert the dynamics model into a linear system, we adopt the Jacobian, which here is described as:

$$\begin{bmatrix} \frac{\partial \dot{x}_1}{\partial x_1} & \frac{\partial \dot{x}_1}{\partial x_2} \\ \frac{\partial \dot{x}_2}{\partial x_1} & \frac{\partial \dot{x}_2}{\partial x_2} \end{bmatrix}$$

the Jacobian matrix of the equilibrium $(0, 0)$ is

$$\begin{bmatrix} -\gamma_1 - \beta & \alpha \\ \beta & -\gamma_2 \end{bmatrix}$$

on the basis of manipulations from linearization operation, the characteristic equation can be proved as the following description:

$$\begin{cases} \dot{x}_1 = \alpha x_2 - (\gamma_1 + \beta)x_1 \\ \dot{x}_2 = \beta x_1 - \gamma_2 x_2 \end{cases} \quad (2)$$

Then, for the convenience of description, we make the following transformation:

$$\begin{aligned} X_1 &= x_1 \\ X_2 &= \dot{x}_1 = (-\gamma_1 - \beta)x_1 + \alpha x_2 \end{aligned}$$

in order to convert the linear systems into a chained-form system, we take the derivative of the above equations:

$$\begin{aligned} \dot{X}_1 &= X_2 \\ \dot{X}_2 &= (-\gamma_1 - \beta)X_2 + \alpha \dot{X}_2 \end{aligned}$$

Here, a practical input u and external disturbance d are introduced into the differential equations to convert the linear systems into a chained-form system:

$$\dot{X}_2 = (-\gamma_1 - \beta)X_2 + \alpha \dot{X}_2 + u + d \quad (3)$$

and next, we design a controller τ to stabilize the chained-form system

$$\tau = (-\gamma_1 - \beta)X_2 + \alpha \dot{X}_2 + u$$

therefore, the final equations can be represented as:

$$\begin{cases} \dot{X}_1 = X_2 \\ \dot{X}_2 = \tau + d \end{cases} \quad (4)$$

3 Stability Analysis

In order to stabilize system (4), take the following discontinuous controller:

$$\tau = \tau_f + \tau_e \quad (5)$$

meanwhile, it is proved that

$$\begin{cases} \tau_f = \bar{v}_2 \\ \dot{\tau}_e + p\tau_e = -(o_q + o_p + \varepsilon)sgn(\varphi) \end{cases} \quad (6)$$

where

$$\tau_e(0) = 0, \varphi = \dot{X}_2 - \bar{v}_2$$

positive constants $\varepsilon, p, o_p, o_q > 0$, the change rate of the existing external disturbance is bounded satisfying $d \leq o_q$ and

$$\begin{cases} \bar{v}_1 = -l_1 w_1^{\frac{r_1+o}{r_1\sigma_0}}, \bar{v}_2 = -l_2 w_2^{\frac{r_2+o}{r_2\sigma_1}} \\ \bar{w}_1 = X_1^{1+o}, \bar{w}_2 = X_2^{\sigma_1} - v_1^{\sigma_1} \end{cases} \quad (7)$$

where $\sigma, l, r_i > 0$ ($i = 1, 2$) are real numbers.

Based on (5), it is obvious to denote a variable equation as

$$\varphi = \dot{X}_2 - \bar{v}_2 = \tau_e + d \quad (8)$$

In order to give reasonable stability analysis, we introduce Lyapunov function [9], utilize its time derivative to prove the stabilization.

$$\begin{aligned} \dot{V}_1 &= \varphi \dot{\varphi} = \varphi(\dot{\tau}_e + \dot{d}) \\ &= -(o_q + o_p + \varepsilon)|\varphi| + \dot{d}\varphi - p\tau_e\varphi \\ &\leq -(o_q + o_p + \varepsilon)|\varphi| + o_q|\varphi| - p\tau_e\varphi \end{aligned} \quad (9)$$

through simplification, the following equation can obtain

$$\dot{V}_1 \leq -(o_q + \varepsilon)|\varphi| - p\tau_e\varphi \quad (10)$$

according to the initial condition $\tau_e(0) = 0$, the following posture can be proved with the Eq. (8):

$$o_p \geq p \max(\tau_e) \geq p|\tau_e| \quad (11)$$

considering the above certification, the following equation can obtain:

$$\begin{aligned} \dot{V}_1 &\leq -(o_q + \varepsilon)|\varphi| + p|\tau_e||\varphi| \\ &\leq -\varepsilon|\varphi| - (o_p - p|\tau_e|)|\varphi| \\ &\leq -\varepsilon|\varphi| = -\sqrt{2\varepsilon}|V_1|^{\frac{1}{2}} \leq 0 \end{aligned}$$

Therefore, subsystem (4) is proved capable of stabilizing to the desirable equilibrium in a finite time $t \geq T_1$ ($T_1 < +\infty$) satisfying $X_1 = 0, X_2 = 0$.

4 Simulations and Test Results

In this section, we demonstrate the performance of the proposed finite-time control approach, then prove the effectiveness of the above method by MATLAB simulation.

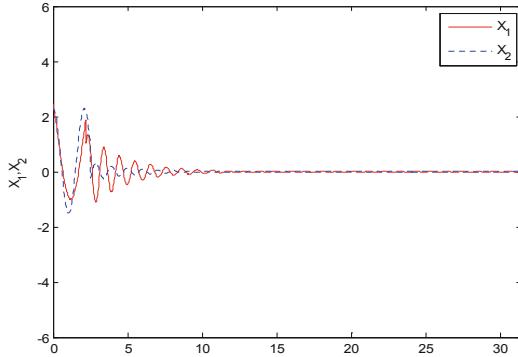


Fig. 1. The stabilization process of X_1 , X_2 .

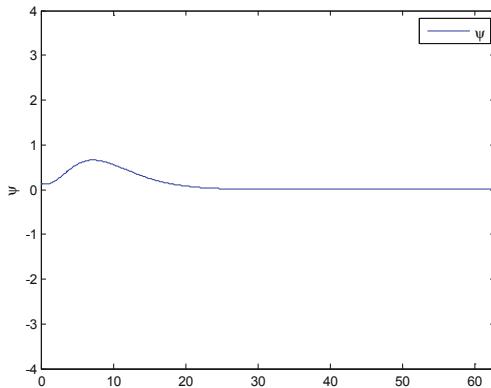


Fig. 2. The stabilization process of φ .

As shown in Fig. 1, the system can converge to zero in a finite time. To begin with, X_1 , X_2 in the system are fluctuating with the amplitude no more than 2. In addition, it is concluded that X_2 will stabilize to 0 faster within 8 s, and then X_1 begins to approach and later stabilize at 0. After around 12 s, X_1 and X_2 are both shown to stabilize to 0, and then the stabilization of the system is achieved.

Figure 2 shows that variable φ in Lyapunov function is bound to converge to zero in a finite time. Therefore, the stability of the system is better proved.

5 Conclusion

In the paper, a finite-time control scheme is considered to stabilize the single species dynamics model with stage structure. In order to obtain chained-form system for clear proof, the single species dynamics model with stage structure is linearized by the Jacobian matrix with the equilibrium. Moreover, we propose a finite-time controller

with disturbance according to the stability control theory, transforming the dynamics model into a chained-form system via rigorous convergence and stability analysis. And finally, the effectiveness of the finite-time control approach is verified by the simulation results, the stabilization of the system is achieved.

References

1. Li L, Qingling Z, Zhu B (2015) Fuzzy stochastic optimal guaranteed cost control of bio-economic singular markovian jump systems. *IEEE Trans Cybern* 45(11):2512–2521
2. Zhang X, Chen L, Avidan U (2000) The stage-structured predator-prey model and optimal harvesting policy. *Math Biosci* 168(2):201–210
3. Gordon H (1954) The economic theory of a common property resource: the fishery. *J Polit Econ* 62(2):124–142
4. Wang Y, Wang D, Ng BC (2017) Finite time moving target tracking using nonholonomic vehicles with distance and bearing angle constraints. In: American control conference. IEEE, pp 2962–2967
5. Chen H, Ding S, Chen X et al (2014) Global finite-time stabilization for nonholonomic mobile robots based on visual servoing. *Int J Adv Rob Syst* 11(10):180
6. Chen H, Zhang B, Zhao T, et al (2016) Finite-time tracking control for extended nonholonomic chained-form systems with parametric uncertainty and external disturbance. *J Vibr Control*. <https://doi.org/10.1177/1077546316633568>
7. de Wit CC, Khennouf H (1995) Quasi-continuous stabilizing controllers for nonholonomic systems: design and robustness consideration. In: Proceedings of the 3rd ECC, pp 2630–2635
8. Sun N, Wu Y, Fang Y, Chen H (2017) Nonlinear antiswing control for crane systems with double-pendulum swing effects and uncertain parameters: design and experiments. *IEEE Trans Autom Sci Eng*. <https://doi.org/10.1109/tase.2017.2723539>
9. Bhat SP, Bernstein DS (1998) Continuous finite-time stabilization of the translational and rotational double integrators. *IEEE Trans Autom Control* 43(5):678–682



On the Complexity of Encapsulated Agent Cell

Yunzhong Song^{1(✉)}, Fengzhi Dai², and Huimin Xiao³

¹ School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454003, China
songhpu@126.com

² School of Electronic Information and Automation,
Tianjin University of Science and Technology, Tianjin 300222, China

³ School of Computer and Information Engineering,
Henan University of Economics and Law, Zhengzhou 450046, China

Abstract. In order to amplify the importance of the minimal cell of the agent systems, the encapsulated agent cell is put forward. And on account of that, the encapsulated cell, which is composed by one of the first order agents and one of the second order agents, is touched upon here. Where three different situations are reported, they are neutral style, positively coupled style and negatively coupled style. Results shown that the neutrally coupled style is a little more complex for its reservation of all the inputs and the outputs of the original systems, whereas the positively coupled and negatively coupled styles are comparatively simple, and the latter two has no distinct difference in control performance.

Keywords: Encapsulated agent · Neutrally coupled style · Positively coupled style · Negatively coupled style

1 Introduction

Networked agent systems have attracted a lot of researchers and engineering scientists since its inception. Focused area started from consensus, and thereafter to the topic of rendezvous and some important points [1–4]. Agent systems can be in a very wide sense, they can be vehicles [5–7], they can be animals and they can be groups of fishes and anything else.

The traditional ways to deal with the networked systems are always being kicked off from single agent locally, and globally ended at the whole systems, tools like linear and nonlinear dynamics, graph theory and Lyapunov theory are often being used to verify the usage of the main ideas. Local agents are too simple to be single and that is not the real thing in the real world. For example, in order to have the huge building, bricks are not poured in the buildings directly today, they are always mixed the other bricks to be in a compact cell, and it is the cell, not the brick itself to be filled in the buildings. The distinguished points lie in that the flexible required characteristics can be met according to the different composite ways of the building materials. Inspired by this point, this paper will focus on the composite agent cell, which is composed by one of the first order agents and one of the second order agents. The control performance

under H infinity, together with the magnitude of the feedback gain of the controller as well as the complexity of the systems, will be explored in this paper.

The paper is arranged as follows, system description will be introduced in Sect. 2, and controller design and simulation results will be provided in Sect. 4, conclusions will be available in the final part.

2 System Description

The encapsulated agent cell is composed of one single first order agent and one single second order agent, and the encapsulated agent cell composite style can be three different ways. And they are positive, neutral and negative, respectively.

2.1 Encapsulated Agent Cell from Positive Composite

Encapsulated agent cell from positive composite, refers to positive feedback between the second order integral agent and the first order integral agent. It can be demonstrated as,

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}x + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}u, \\ y &= [0 \ 0 \ 1]x.\end{aligned}\quad (1)$$

This system can be viewed as the composite of the two reduced order system as follows:

For the second order agent,

$$\begin{aligned}\dot{p} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}p + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u_p, \\ q &= [1 \ 0]p.\end{aligned}\quad (2)$$

For the first order agent,

$$\begin{aligned}\dot{r} &= u_r, \\ s &= r.\end{aligned}\quad (3)$$

And assume

$$\begin{aligned}u_p &= k(r - p_1), \\ u_r &= k(p_1 - r).\end{aligned}\quad (4)$$

And let the overall augmented system output comes from the first order agent output, the overall augmented system input comes from the second order agent, then the augmented system can be recast as,

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & 1 & 0 \\ k & 0 & -k \\ -k & 0 & k \end{bmatrix}x + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}u, \\ y &= [0 \ 0 \ 1]x.\end{aligned}\quad (5)$$

And further assume that the feedback coefficient be equal to positive scalar 1 will turns out to be the composite System 1.

2.2 Encapsulated Agent Cell from Negative Composite

Encapsulated agent cell from negative composite, refers to negative feedback between the second order integral agent and the first order integral agent. It can be demonstrated as,

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}x + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}u, \\ y &= [0 \ 0 \ 1]x.\end{aligned}\quad (6)$$

It is not so difficult to find how to transfer System 1 to System 6, and the process is omitted here in case of simplicity.

2.3 Encapsulated Agent Cell from Neutral Composite

Encapsulated agent cell from neutral composite, refers to neutral connection between the second order integral agent and the first order integral agent. It can be demonstrated as,

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}x + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \\ y &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}x.\end{aligned}\quad (7)$$

Obviously, the overall augmented system output comes from the first order agent and the second order agent, all of them, and the overall augmented system input comes from the second order agent and the first order agent, all of them. Obviously, the input of the augmented system should be one vector, which is 2 times 1, column vector. At the same time, the output of the augmented system should be one vector, which is 1 times 2, row vector. And this composite is the decoupling composite. And that is neutral in feedback.

3 Controller Design and Simulation Results

Consider the composite agent cell, which is packed hybrid order, can be described by,

$$\begin{aligned}\dot{x} &= Ax + B_u u + B_\omega \omega, \\ z &= Cx.\end{aligned}\tag{8}$$

where $x \in R^n$ is the state of the agent cell, $u \in R^p$ is the control input, $\omega \in L_2^m[0, \infty)$ is the exogenous input that denotes disturbances, and $z \in R^m$ is the output that denotes performance.

3.1 Controller Design

The controller is defined as the full state feedback form,

$$u = Kx.\tag{9}$$

Lemma 1 (Bounded Real Lemma, Jia [8]). Let $\gamma > 0$ and $G(s) = C(sI - A)^{-1}B_u$. Then, the following two statements are equivalent:

- (1) A is stable and $\|G(s)\|_\infty < \gamma$.
- (2) There exists an $X > 0$ such that

$$AX + XA^T + XC^T CX + \gamma^2 B_u B_u^T < 0.$$

In the following, a necessary and sufficient condition is gotten for the existence of a controller for (9) having an unbounded H_∞ performance region.

Lemma 2 (Jia [8]; Li, Duan and Chen [9]). For a positive scalar $\gamma > 0$, a controller (9) can enable the system (8) has an unbounded H_∞ performance region if and only if the following linear matrix inequality (LMI) about $P > 0$ and a scalar $\tau > 0$ satisfy,

$$M(P, \tau) = \begin{bmatrix} AP + PA^T - \tau B_u B_u^T & 0 & PC^T \\ * & -\gamma^2 I & 0 \\ * & * & -I \end{bmatrix} < 0.\tag{10}$$

Corollary 3. The H_∞ performance of System (8) under Controller (9) is determined by the following optimization problem,

$$\begin{array}{ll}\text{minimize} & \gamma \\ \text{subject to} & \text{LMI(10)}.\end{array}\tag{11}$$

Algorithm 1 (Loberg [10]; Sturm [11]). For any $\bar{\gamma} > \gamma$, where γ is determined via (11), the controller (9) solving the H_∞ control problem of System (8), can be configured as follows,

- (1) Solve LMI (10) for a feasible solution: P, τ ;
- (2) Choose the feedback gain matrix $K = -\frac{1}{2}B_u^T P^{-1}$.

3.2 Simulation Results

According to the controller design method, we listed the resulted controller for three different encapsulated agent cells, respectively.

For the positive coupled agent cell,

$$K = -\frac{1}{2}B_u^T P^{-1} = [-10.0437 \ -4.6225 \ 12.8317];$$

For the negative coupled agent cell,

$$K = -\frac{1}{2}B_u^T P^{-1} = [-1.6343 \ -2.3364 \ -0.5618];$$

and for the neutral coupled agent cell,

$$K = -\frac{1}{2}B_u^T P^{-1} = \begin{bmatrix} -1.6662 & -2.4822 & 0 \\ 0 & 0 & -1.5763 \end{bmatrix}.$$

Remark 1. Here we selected $\gamma = 1$, and system parameters of encapsulated agent cell as the same as Sect. 2;

In case of saving the space, we only provide the results for the neutral coupled agent cell, and the result is listed as Fig. 1.

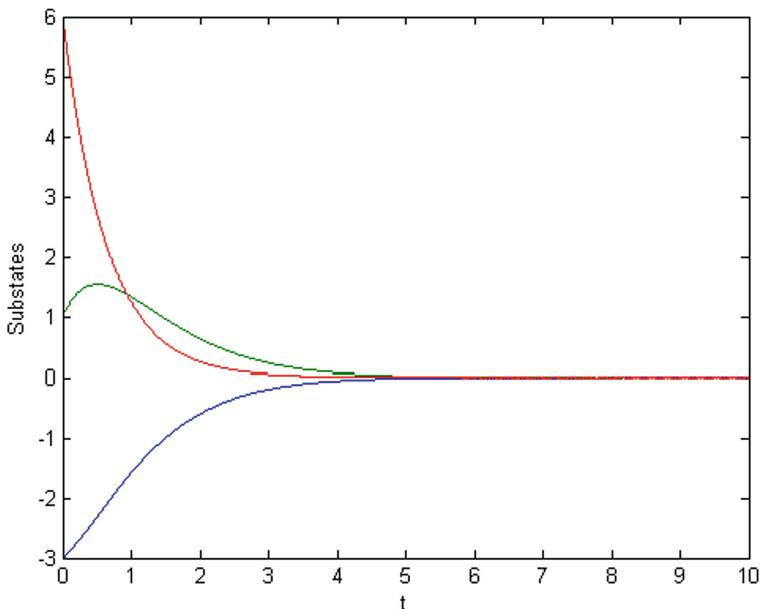


Fig. 1. Control effects of neutral coupled agent cell

For the simulation, the disturbance signals are chosen as.

$$\omega(t) = \begin{bmatrix} \sin 300t \\ \sin 300t \end{bmatrix}, \text{ and disturbance input matrix } B_\omega = \begin{bmatrix} 0 & 0 \\ 0.1 & 0 \\ 0 & 0.2 \end{bmatrix}.$$

Remark 2. The encapsulated agent cell reported here are totally controllable and observable. So the assumption of the states feedback is feasible. And in this case, we do not want to discern the output signals from the performance signals.

Remark 3. We do not have the sharp difference in mathematical modeling and controller design for both positive and negative coupled agent cell, and this is easily understood that when we looked the controller and the cell uniformly, the difference in coupled of them is lessened via feedback control.

Remark 4. The neutral coupled agent cell is a little more complex, since this is the most loosely coupled format between the first order and the second sub-agent.

Remark 5. From the process of H_∞ controller design, we do not see any extra cost compared with the single agent, and this is so good for us to explore the more plentiful composite agent cells in the near future.

4 Conclusions

Three encapsulated agent cell have been provided to enrich the corpus of the networked agent systems research. The key point lies in that with the encapsulated agent cell, we can have the networked agent cell systems have distinguished characteristics as locally heterogeneous, and globally homogeneous; and this heterogeneous and homogeneous collaboration can be promising in the near coming research of multi-agent systems.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grants 61340041 and 61374079, and by the Natural Science Foundation of Henan Province, China under Grant 182300410112. And the first author want to express his thanks to Professor Yingmin Jia from Beihang University for the kindly offer of the monograph titled as *Robust H_∞ Control*, and this book is very helpful in doing research of this paper.

References

1. Olfati-Saber R, Murray RM (2004) Consensus problems in networks of agents with switching topology and time delays. *IEEE Trans Autom Control* 49(9):1520–1533
2. Song YZ (2016) Consensus of agents with mixed linear discrete dynamics. *Int J Control Autom Syst* 14(4):1139–1143
3. Song YZ, Zhao W (2014) Multi-agent system rendezvous via refined social system and individual roles. *WSEAS Trans Syst Control* 9:526–532
4. Song YZ, Fu ZY, Wang FZ (2016) Socialized multi-agent system rendezvous via networks of networks. *J Rob Netw Artif Life* 3:136–139
5. Jia YM (2000) Robust control with decoupling performance for steering and traction of 4WS vehicles under velocity varying motion. *IEEE Trans Control Syst Technol* 8(9):554–569

6. Jia YM (2003) 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* 48:1413–1416
7. Jia YM (2009) General solutions to diagonal model matching control of multi-output delay systems and its applications in adaptive scheme. *Prog Nat Sci* 19:79–90
8. Jia YM (2007) Robust H_∞ Control. Science Press, Beijing
9. Li ZK, Duan ZS, Chen GR (2011) On H_∞ and H_2 performance regions of multi-agent systems. *Automatica* 47:797–8036
10. Loberg J (2004) YALMIP: a toolbox for modelling and optimization in MATLAB. In: Proceedings of 2004 IEEE international symposium on computer aided control system design, pp 284–289
11. Sturm JF (1999) Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optim Methods Softw* 11(1):625–653



Research on Novel Method of the Branch Series Resonance Recognize

Yongle Ai, Qunfeng Liu^(✉), and Shuai Li

School of Electrical Engineering, Henan Polytechnic University,
Jiaozuo 454000, China
lqf7333@163.com

Abstract. For the identification of series resonance in power system branches, the identification method of branch series resonance based on port equivalent method is proposed. First of all, a typical three-bus test system is used to analyze the series resonance of the branch. It is found that the branch series resonance can be identified by the port equivalent admittance formed between the bus node and the reference node. Then, by establishing the node port association matrix and using the node impedance matrix, the multi-port Thevenin equivalent impedance matrix is obtained, then the equivalent admittance matrix of the port is formed, and the load branch of the node is eliminated, as well as the effect of the node load branch on the series resonance of the identification branch is eliminated. The algorithm effectively utilizes the node admittance matrix commonly used in power system. Finally, the feasibility and correctness of the three-bus system and IEEE9 network system are proved by simulation analysis. The validity of this novel formulation is confirmed through theoretical analysis and case studies.

Keywords: Branch series resonance · Node admittance matrix · Node impedance matrix · Equivalent port

1 Introduction

The harmonics mainly come from power electronic equipment such as electric arc furnaces and converters. The access of these nonlinear loads injects large amount of harmonics into the grid. The existence of harmonics is inevitable because there is nothing to replace the position of nonlinear loads in the fields of metallurgy and chemical engineering. The impedance of the system is inductive in normal operation of the system. There are often several series-parallel resonance points, which are affected by the presence of reactive compensation capacitors and high-voltage transmission line capacitors. The operation of the system and the safety of the electrical equipment will be compromised, when the power system resonates the excitation source frequency is equal to the resonance point frequency [1, 2]. How to efficiently and comprehensively identify the existence of series-parallel resonance points is very meaningful for power system harmonic resonance analysis.

It's no doubt that the analytical and spectrum analysis methods can accurately identify and determine the series-parallel resonance point among scholars worldwide

[3]. However, the limitations of traditional analytical methods have become increasingly prominent for the current complex power networks. Thence, scholars at home and abroad have conducted more in-depth research. In [4, 5], Resonance Mode Analysis (RMA) is proposed. By decomposing the eigenvalues of the admittance matrix of the network node, the impedance, voltage and current of the corresponding modes are obtained, and the coupling between the nodes is obtained. Decoupling can accurately identify parallel resonances and analyze related resonance information, but this method cannot identify and analyze series resonances. Literature [6, 7] proposed a loop modal analysis method for analyzing series resonance, which is different from the series resonance point identified by spectrum analysis method, but this method can only identify the series resonance point of the loop, and cannot identify the branch in the branch. Series resonance point. In [8], a series resonance analysis method combining modal method and virtual branch method is proposed. Based on the idea of node admittance matrix analysis of parallel resonance, the series resonance is analyzed by loop impedance matrix. By analyzing the characteristics of voltage harmonic source and current harmonic source, under the premise of ensuring the network structure is unchanged, a virtual branch is added at each node, and the series resonance in the loop is identified by means of eigenvalue decomposition. However, it is difficult to establish a loop impedance matrix in a complex network, and adding a virtual branch multiplies the scale of the impedance matrix. The literature [9, 10] proposed the branch method to analyze the series resonance. This method of resonance point identification is the same as the RMA, and only the parallel resonance can be identified. In [11], an improved virtual branch method is proposed to analyze series resonance. This method uses the node admittance matrix to analyze the series resonance. It is not necessary to reestablish the loop impedance matrix, but it is still necessary to establish the admittance matrix and the node relationship matrix of the branch. The method is based on a single branch to investigate the series resonance, but whether the parallel resonance or the series resonance is caused by the interaction of components on several branches, the single branch cannot identify the resonance properties in more detail. Literature [12] proposed a method of using capacitors to scan and identify parallel resonance, which provides a new solution to the resonance problem of reactive compensation capacitors.

This paper proposes a branch circuit resonance identification method based on the port equivalent method. It is found that the branch series resonance can be identified by the port equivalent admittance formed by the bus node and the reference node, through the analysis of the three-bus test system. A multi-port Thevenin equivalent impedance matrix is obtained by establishing a node-port association matrix in order that form a port equivalent admittance matrix. The series resonance of the driving point is caused by the series resonance, so that the series resonance point of the branch becomes the frequency point corresponding to the peak value of the frequency-port equivalent admittance curve. The algorithm effectively utilizes the node admittance matrix commonly used in power systems. Finally, the feasibility and correctness of the proposed method are verified by case study.

2 Resonance Analysis of Branch Series

The three-bus test system is selected shown in Fig. 1, where the impedance unit is Ω and the admittance unit is S. Analytical methods are used to solve system series and parallel resonance by the platform of MATLAB.

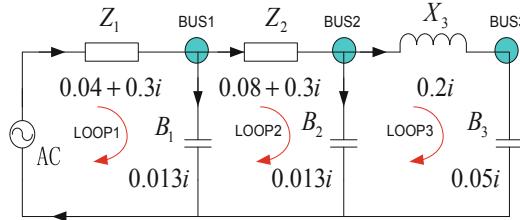


Fig. 1. Three-bus test system

2.1 Parallel Resonance

In Fig. 1, the equivalent impedance equations for the three-bus test system are established as shown in (1)–(3) respectively.

$$Z_{eqBUS1} = X_{C1}(h) // Z_1(h) // \{Z_2(h) + X_{C2}(h) // [X_3(h) + X_{C3}(h)]\} \quad (1)$$

The harmonic order calculated for resonant frequency are: $h = 4.58, 18.96, 29.54$.

$$Z_{eqBUS2} = X_{C2}(h) // \{Z_2(h) + Z_1(h) // X_{C1}(h)\} // \{X_3(h) + X_{C3}(h)\} \quad (2)$$

The harmonic order calculated for resonant frequency are: $h = 4.58, 18.96, 29.54$.

$$Z_{eqBUS3} = X_{C3}(h) // \{X_3(h) + X_{C2}(h) // [Z_2(h) + X_{C1}(h) // Z_1(h)]\} \quad (3)$$

The harmonic order calculated for resonant frequency are: $h = 4.58, 18.96, 29.54$.

Where: “+” represents the branches connected in series, and “//” represents branches connected in parallel. Z_{eqBUS1} , Z_{eqBUS2} , Z_{eqBUS3} are the equivalent impedance of BUS1, BUS2, BUS3 respectively; $X_{C1}(h)$, $X_{C2}(h)$ and $X_{C3}(h)$ are the capacitance impedances of the B_1 , B_2 and B_3 of branches respectively; $Z_1(h)$, $Z_2(h)$, $X_3(h)$ are the impedance of the branches respectively; h is the harmonic order.

2.2 Series Resonance

There are two types series resonance which are loop series resonance and branch series resonance.

- (1) Loop series resonance: In Fig. 1, three loop equivalent admittances $Y_{eqLOOP1}$, $Y_{eqLOOP2}$ and $Y_{eqLOOP3}$ are established as the basic loop based on “mesh” as shown

in Eqs. (4)–(6). The frequencies of series resonance points can be calculated through the Eqs. (4)–(6) respectively.

$$Y_{eqLOOP1} = 1/(Z_1(h) + X_{C1}(h) // \{Z_2(h) + X_{C2}(h) // [X_3(h) + X_{C3}(h)]\}) \quad (4)$$

The harmonic order calculated for resonant frequency are: $h = 4.58, 18.96, 29.54$ respectively

$$Y_{eqLOOP2} = 1/(Z_2(h) + X_{C1}(h) // Z_2(h) + X_{C2}(h) // \{X_2(h) + X_{C3}(h)\}) \quad (5)$$

The harmonic order calculated for resonant frequency are: $h = 4.58, 18.96, 29.54$

$$Y_{eqLOOP3} = 1/(X_3(h) + X_{C3}(h) + X_{C2}(h) // \{Z_2(h) + X_{C1}(h) // Z_1(h)\}) \quad (6)$$

The harmonic order calculated for resonant frequency are: $h = 4.58, 18.96, 29.54$

- (2) Branch series resonance. In Fig. 1, assume “ Z_2 ” and “ X_3 ” as harmonic source, for their left circuits, four branches equivalent admittances Y_{eq1} , Y_{eq2} , Y_{eq3} , Y_{eq4} are established as shown in the Eqs. (7)–(10). The frequencies of series resonance points can be calculated through the Eqs. (7)–(10) respectively.

$$Y_{eq1} = 1/(Z_2(h) + Z_1(h) // X_{C1}(h)) \quad (7)$$

The harmonic order calculated for resonant frequency is: $h = 22.64$.

$$Y_{eq2} = 1/(Z_2(h) + X_{C2}(h) // \{X_2(h) + X_{C3}(h)\}) \quad (8)$$

The harmonic order calculated for resonant frequency are: $h = 6.04, 26.54$.

$$Y_{eq3} = 1/(X_3(h) + X_{C2}(h) // \{Z_2(h) + X_{C1}(h) // Z_1(h)\}) \quad (9)$$

The harmonic order calculated for resonant frequency are: $h = 17.68, 29$.

$$Y_{eq4} = 1/(X_3(h) + X_{C3}(h)) \quad (10)$$

The harmonic order calculated for resonant frequency is: $h = 10.00$.

It can be known from the above equations (1) The types of resonance in the system cannot be simply divided into series resonance and parallel resonance. The series resonance needs to be further divided into loop series resonance and branch series resonance; (2) The resonance point from the loop series by the system “mesh” is same as the resonance point from the “node” point which means the parallel sub-circuits.

With MATLAB platform the system frequency-impedance diagram is obtained by impedance spectrum scanning simulation as shown in Fig. 2. The parallel resonance points can be determined by checking the peak impedance of the research points and oppositely the series resonance point can be determined by checking the valley impedance of the research points. From Fig. 2, it can be found that in the three-bus test

system there are three parallel resonance points and series resonance points which are 6.04 p.u., 10.00 p.u., 17.68 p.u., 22.64 p.u., 26.54 p.u., 29 p.u.

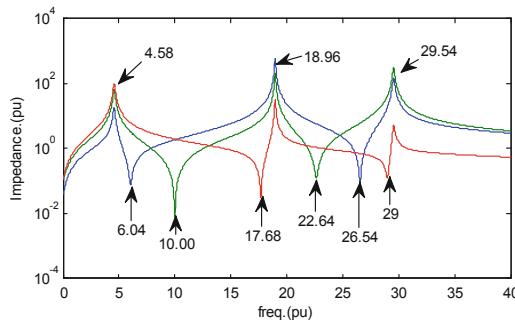


Fig. 2. Frequency impedance curve

2.3 Summary

It is found that the key of the series resonance identification is to select the branch researched. To complete the series resonance identification it will need three steps. First step, a sub-circuit researched which is an equivalent port formed between the node and the reference node is selected. Second step, the equivalent Thevenin impedance is calculated. Third step, the equivalent port admittance is obtained by converting the equivalent Thevenin impedance calculated. Taking the branch “ Z_2 ” shown in Fig. 1 as an example, the corresponding branch series resonant port is shown in Fig. 3.

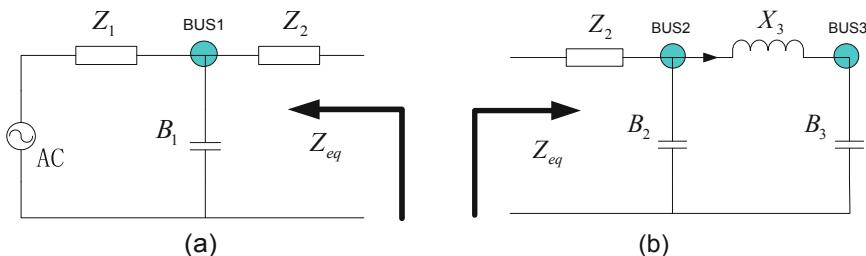


Fig. 3. Branch series port equivalent diagram

3 Port Equivalent Method

For a power network with N nodes (excluding reference nodes), as shown in Fig. 4, if you are interested in m ports (port subscripts denoted by a, b, \dots, m), The nodes - port correlation matrix can be established. Then a multi-port Thevenin equivalent port impedance matrix of $m \times m$ -dimension formed is analyzed. The Norton equivalent

port admittance matrix can be obtained by inverting the Thevenin equivalent port impedance matrix.

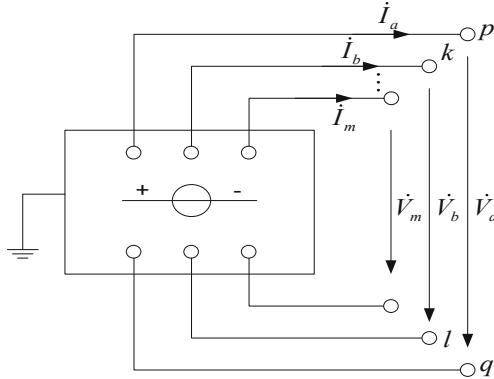


Fig. 4. Power network

In Fig. 4, \$p, q\$ are two nodes which are formed a port. The direction of the current is from first node to the second node which means \$p \rightarrow q\$. If both of the \$p, q\$ in a port is not reference node the correlation vector of the node-port is,

$$\mathbf{M}_a = [0 \ \cdots \ 1 \ \cdots \ -1 \ \cdots \ 0]^T \quad (11)$$

Where “1” and “−1” represents the positions of the \$p\$th and \$q\$th.

If the node of \$q\$ is a reference node, its corresponding \$N \times 1\$ dimensional node-port correlation vector is,

$$\mathbf{M}_a = [0 \ \cdots \ 1 \ \cdots \ 0 \ \cdots \ 0]^T \quad (12)$$

The node-port correlation vectors formed by the \$m\$ ports are put together in columns to form a \$N \times m\$ dimensional node-port correlation matrix \$\mathbf{M}\$.

$$\mathbf{M} = [\mathbf{M}_a \ \mathbf{M}_b \ \cdots \ \mathbf{M}_m] \quad (13)$$

Therefore, the formed \$m \times m\$-dimensional multiport Thevenin equivalent port impedance matrix is,

$$\mathbf{Z}_{eq} = \mathbf{M}^T \cdot \mathbf{Z} \cdot \mathbf{M} \quad (14)$$

Assume that the \$m\$ ports of the original network in Fig. 3. are opened, the equivalent electromotive voltage of the Thevenin equivalent port are,

$$\dot{\mathbf{V}}_{eq}^{(0)} = \mathbf{M}^T \dot{\mathbf{V}}^{(0)} = [\dot{\mathbf{V}}_a^{(0)} \ \dot{\mathbf{V}}_b^{(0)} \ \cdots \ \dot{\mathbf{V}}_m^{(0)}] \quad (15)$$

Then, Thevenin equivalent network of \$m\$ ports can be formed, as shown in Fig. 5.

It is known from Eq. (14) that the multi-port Norton equivalent port admittance matrix of the $m \times m$ -dimensional is

$$\mathbf{Y}_{eq} = \mathbf{Z}_{eq}^{-1} \quad (16)$$

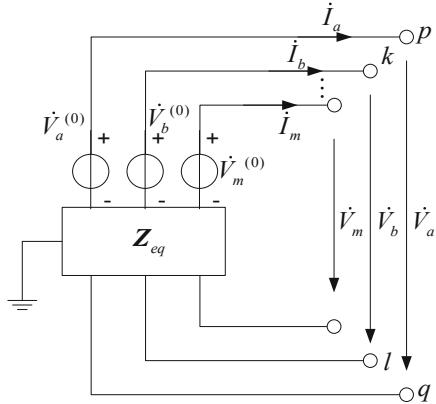


Fig. 5. m port's Thevenin equivalent network

For the three-bus system shown in Fig. 1, a port correlation matrix formed by its corresponding node and a reference node can be established.

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (17)$$

Also it can be known from (7)–(10) that when the port equivalent admittance is formed, the load branch of the corresponding node is not considered; therefore, in order to eliminate the influence of the node load branch on the identification branch series resonance, by using numerical processing means.

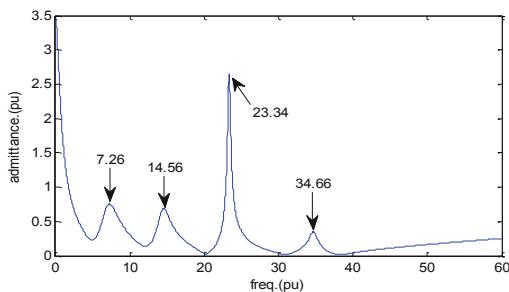
4 Case Studies

In order to further verify the feasibility and effectiveness of the proposed Thevenin-Norton equivalent port method to identify the branch series resonance, the IEEE9 node distribution network system is selected for simulation analysis. The parameters of each node load, line impedance, and capacitor configuration of the IEEE9 network system are shown in Table 1. For the IEEE9 node network system, the standard capacity is 100MVA, and a single capacity of the configured capacitor is 300kVar. At the same time, the load parameters of each node are modeled as series constant impedance load equivalent. The frequency series analysis method and the method proposed in this

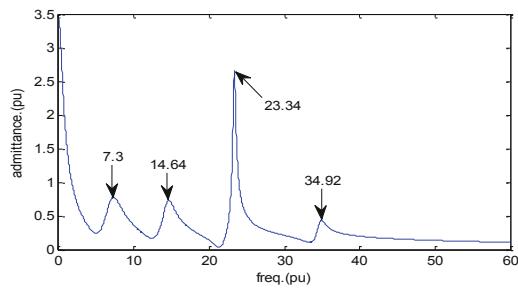
paper are used to identify the branch series resonance, and the branch series resonance point identification map is obtained as shown in Fig. 6.

Table 1. IEEE9 distribution network system parameter

| Branch number | First node | End node | R(p.u.) | X (p.u.) | End work (kW) | End reactive (kVar) | Capacitor number |
|---------------|------------|----------|---------|----------|---------------|---------------------|------------------|
| 1 | 0 | 1 | 0.0233 | 0.0780 | 1840 | 460 | 0 |
| 2 | 1 | 2 | 0.0026 | 0.1144 | 980 | 340 | 5 |
| 3 | 2 | 3 | 0.1411 | 0.2278 | 1790 | 446 | 0 |
| 4 | 3 | 4 | 0.1320 | 0.1150 | 1598 | 1840 | 2 |
| 5 | 4 | 5 | 0.3749 | 0.3266 | 1610 | 600 | 4 |
| 6 | 5 | 6 | 0.1711 | 0.1491 | 780 | 440 | 5 |
| 7 | 6 | 7 | 0.3885 | 0.2200 | 1150 | 60 | 0 |
| 8 | 7 | 8 | 0.9065 | 0.5134 | 980 | 130 | 5 |
| 9 | 8 | 9 | 1.0101 | 0.5721 | 1640 | 200 | 0 |



(a) Spectrum analysis



(b) Equivalent port method

Fig. 6. IEEE 9 system branch series resonance recognition diagram

5 Conclusion

The series-parallel resonance is divided into two types of the resonance problem, but the resonance in the power system cannot be simply divided into parallel resonance and series resonance. For series resonance, it can be divided into loop series resonance and branch series resonance. For parallel resonance, the modal method based on the node admittance matrix can be used to identify and analysis. For the series resonance of the branch, it cannot be simply regarded as a loop series resonance to identify the series frequency. The resonance can be identified using the port equivalent admittance formed by the node corresponding to the branch. Therefore, an equivalent port method is proposed to identify the branch series resonance using the three-bus test system, it is found that the branch series resonance can be identified by the port equivalent admittance formed between the bus node and the reference node. By establishing the node-port correlation matrix and using the node impedance matrix, the multi-port Thevenin equivalent correlation matrix is obtained. And then the port equivalent admittance matrix can be obtained. By using characteristic of the frequency-port equivalent admittance the branch series resonance frequency point can be determined and is equal to the peak point frequency of the frequency-port equivalent admittance curve.

References

1. Eghtedarpour N, Karimi MA, Tavakoli M (2014) Harmonic resonance in power systems-a documented case. In: IEEE international conference on harmonics and quality of power. IEEE, pp 857–861
2. Li HJ, Wang T, Tong XH et al (2016) Analysis and treatment for resonance phenomena of filter capacitor bank. Power Capacit React Power Compens 37(3):24–29
3. Dolara A, Gualdoni M, Leva S (2011) Impact of high-voltage primary supply lines in the 2 × 25 kV–50 Hz railway system on the equivalent impedance at pantograph terminals. IEEE Trans Power Deliv 27(1):164–175
4. Xu W, Huang Z, Cui Y et al (2005) Harmonic resonance mode analysis. IEEE Trans Power Delivery 20(2):1182–1190
5. Amornvipas C, Hofmann L (2010) Resonance analyses in transmission systems: experience in Germany. In: Power and energy society general meeting. IEEE, pp 1–8
6. Yang CX, Liu KP, Wang DX (2008) A loop modal analysis method for series harmonic resonance assessment. High Voltage Eng 34(11):2459–2462
7. Lei ZL, Ai X, Cui MY et al (2011) Simulation on series harmonic resonance of microgrid based on modal assessment method. J Zhejiang Univ (Eng Sci) 45(1):178–184
8. Zhou H, Wu YW, Lou SH et al (2007) Series resonance analysis based on modal analysis and dummy branch method. Proc CSEE 27(28):84–89
9. Hu HT, He ZY, Zhang M, et al (2012) Series resonance analysis in high-speed railway all-parallel AT traction power supply system. Proc CSEE 32(13):10+78–86
10. Wang B, Jiang XF, Gao SB et al (2012) Series resonance analysis based on branch-circuit method. J Southwest Jiaotong Univ 47(6):1021–1026

11. Wang B, Qiu ZC, Han XD et al (2015) Series resonance analysis based on improved dummy branch method for high-speed railway traction network. *Electric Power Autom Equip* 35(11):90–95
12. Heydt GT, Patil HU, Loehr J, et al (2016) Harmonic resonance assessment for transmission class shunt capacitors. In: Transmission and distribution conference and exposition. IEEE, pp 1–5



Simulation Research on Fault Diagnosis of CNC Machine Tools Based on Fuzzy Petri Net

Na Lu^(✉), Bing Xu, Zheyi Chen, Di Ren, Yuanyuan Yang,
and Peng Bian

School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai 201418, China
dargyluna@163.com

Abstract. In this paper, a Simulation Research on Fault Diagnosis of CNC Machine Tools Based on Fuzzy Petri Net is proposed. According to the fault information of the CNC machine tool electrical system, the Petri net and fuzzy reasoning are combined to establish the fuzzy Petri net model of the CNC machine tool electrical system. Describes the relationship between CNC machine faults. The fuzzy generation rule is represented by FPN. The fault diagnosis rule of Petri net is used for fault diagnosis reasoning. A fuzzy Petri net model combining reverse reasoning and forward excitation is proposed to analyze the causal relationship between abnormal behavior processes. Combining the feasibility of the occurrence relationship between faults and the frequency of faults, reverse reasoning and forward excitation can accurately and quickly find the root cause of the fault, and can find the fault cause more quickly than the traditional fault diagnosis method. The maintenance time is reduced. The fault diagnosis of the electrical system of CNC machine tools is taken as an example. The diagnostic model based on fuzzy Petri net is established. The correctness of the model and the effectiveness of the algorithm are verified by simulation analysis. Improve the usability of CNC machine tools.

Keywords: Fault diagnosis · CNC machine tools · Fuzzy petri nets

1 Introduction

Manufacturing is the foundation of the national economy, and its degree of development basically reflects the social productivity and technological level of a country [1]. CNC machine tools are an important part of the manufacturing industry. The intelligent technology of CNC machine tools determines the degree of development of the manufacturing industry [2]. In the traditional manufacturing industry, the occurrence of CNC machine tool failures is judged by manual experience. Work efficiency can't keep up with the production schedule, causing great losses to manufacturers. The advancement of science and technology has also promoted the upgrading of numerical control equipment [3, 4]. CNC equipment has complex technology, high operating conditions, and relatively high average operating cost. Highly integrated equipment often leads to downtime. If the fault cannot be corrected in time, it will lead to

long-term shutdown and extended delivery cycle. This will bring serious economic losses to the company. Based on the above phenomenon, this paper based on fuzzy Petri Net (Fuzzy Petri Net) numerical control machine tool fault simulation research, can accurately and quickly find the root cause of the fault, reduce the repair time, and thus enable manufacturers to maintain efficient production efficiency [5, 6].

Petri net is a method of modeling using mesh graphical representation system [6, 7]. It is a graphical modeling tool with strict mathematical definition. It has strong distribution description ability and is a dynamic system modeling for asynchronous, discrete and concurrent events. A powerful tool for performance analysis has been widely used in many fields [7]. The use of Petri nets in the fault diagnosis system can well describe the process of fault occurrence and know where the fault occurred. However, due to uncertainty or ambiguity, it is difficult to meet the requirements of fault diagnosis by using the common Petri net modeling method. Therefore, this paper combines fuzzy technology with Petri net, defines fuzzy Petri net model, combined with the characteristics of fault diagnosis, and takes the fault diagnosis of CNC machine tool as an example, and studies the Petri net model in the field of fault diagnosis [8]. At the same time, using the reverse reasoning and forward excitation of the FPN model, the causal relationship between abnormal behavior processes is analyzed, and the root cause of the fault is found accurately and quickly [9]. Compared with the traditional fault diagnosis method, the fault cause can be found more quickly, the repair time is reduced, and the availability of the numerical control machine tool is improved [10].

2 Definition and Inference Rules of Fuzzy Petri Nets

2.1 Definition of Fuzzy Petri Net

The FPN structure is defined as a ten-element array: $FPN = (P, T, D, I, O, f, \alpha, \beta, \gamma, \mu)$ [6], where $P = \{p_1, p_2, \dots, p_n\}$ is the non-empty finite set of the Place; $T = \{t_1, t_2, \dots, t_m\}$ is the non-empty finite set of Transition; $P \cap T = \emptyset$; $D = \{d_1, d_2, \dots, d_n\}$ is the non-empty finite set $|P| = |D|$ of the proposition; $I : P \times T \rightarrow \{0, 1\}$ is the input function, if $I(p, t) = 1$, It means that there is a relationship between t and p . In this case, p is the input Place of t , The set of all the Places of t is represented as represent as $\bullet t$ or $I(t)$; if $I(p, t) = 0$, there is no connection between p and t ; $O : T \times P \rightarrow \{0, 1\}$ is Input function, if $O(p, t) = 1$, it means that there is a relationship between t and p . At this time, p is the output Place of t . The set of all Places of t is represented as $t \bullet$ or $O(t)$; if $O(p, t) = 0$, there is no connection between t and p ; $\alpha : P \rightarrow [0, 1]$ is a correlation function of p , it represents a mapping to a real number between $[0, 1]$; $\beta : P \rightarrow D$ is a correlation function of p , which represent a bidirectional mapping of p to a set of propositions; μ is a transition credibility function, i.e. $\mu(t_i) = \mu_i$, $\mu_i \in [0, 1]$, indicating that the credibility of the transition t_i is μ_i ; γ is a transition frequency function, that is, $\gamma(t_i) = \gamma_i$, $\gamma_i \in [0, 1]$, indicating that the frequency of the transition t_i excitation is γ_i ; f is the transition reverse excitation function, $f(t_i) = \mu_i \times \gamma_i$, $f(t_i) \in [0, 1]$. Let A be a set of directed arcs. If $p_j \in I(t_i)$, there is a directed arc aji , from p_j to t_i , $aji \in A$; If $p_k \in O(t_i)$, there is a directed arc aik , from t_i to p_k , $aik \in A$. If $f(t_i) = \mu_i$, $\mu_i \in [0, 1]$,

then t_i is said to be associated with real μ_i ; If $\beta(pi) = di$, $di \in D$, then pi is associated with proposition di .

2.2 Accessibility of FPN

Let ta be a transition, pi , pj and pk are three places. If $pi \in I(ta)$ and $pk \in O(ta)$, then pk is said to be immediately reachable from pi ; if pj is immediately reachable from the place pk , then pj is reachable from pi . All sets of places that are immediately reachable from the place pi are called immediate reachable sets of pi , denoted as $IRS(pi)$; All sets of Places that are reachable from pi are called reachable sets of pi , denoted as $RS(pi)$. If the place pi and pk are the input places of the transition ta , that is $pi \in I(ta)$ and $pk \in O(ta)$, then pi and pk are said to be adjacent places of ta .

2.3 Fuzzy Inference Rule for Fault Diagnosis FPN

The ambiguity of FPN place and change. In the fault diagnosis FPN, a system consists of two parts: the place and the transition. The behavior and status of the fault represented by the place. The place is divided into the following three categories:

- (1) Only the place with $T \rightarrow P$ without $P \rightarrow T$ is called the fault place, and the corresponding proposition indicates the phenomenon of failure;
- (2) The proposition corresponding to the fault location place, which is both $T \rightarrow P$ and $P \rightarrow T$, indicates the location of the fault;
- (3) Only $P \rightarrow T$, the place without $T \rightarrow P$ is called the fault cause place, and the corresponding proposition indicates the root cause of the fault. The true degree of the fault corresponding to the original proposition is called the credibility of the place, has a certain degree of ambiguity, and takes values according to expert experience.

3 Application of FPN in Fault Diagnosis of CNC Machine Tools

3.1 Fuzzy Petri Net CNC Machine Tool Fault Diagnosis Model

Taking the fault part of CNC machine tool as the fault diagnosis example of Fuzzy Petri Net, the failure diagnosis of CNC machine tool in [5, 10] literature is used as the fault basis, the fault diagnosis model of Fuzzy Petri Net machine tool is shown in Fig. 1.

Transition set $T = (t_1, t_2, \dots, t_{36})$. Where t_1 is the servo system failure if the communication cable is broken or the connector is loose; t_2 is the servo system failure if the drive module fails; t_3 is if the control module is faulty, the servo system is faulty; t_4 is the servo system fault if the parameter is set incorrectly; t_5 is the control circuit fault if the relay or contactor fails; t_6 is If the air switch trips or blows, the control circuit is faulty; t_7 is if the power supply outputs low voltage or no voltage, then the control circuit is faulty; t_8 is if the scale is contaminated or damaged, the system fault is

detected; t_9 is the system fault if the readhead is faulty or contaminated; t_{10} is the actuator fault if the motor is braked; t_{11} If the power cable is faulty, the servo motor is faulty; if t_{12} is the fault of the inverter motor, the servo motor is faulty; t_{13} is if the transmission is worn or loose, the transmission system malfunction; t_{14} is the transmission system failure if the bearing is worn; t_{15} is the hydraulic station failure if the solenoid valve is faulty; t_{16} is if the flow and pressure switch are faulty, hydraulic station failure; t_{17} is the hydraulic pump ion failure if the oil pump motor fails; t_{18} is if the oil pump output is no oil or the oil quantity is too small, the hydraulic station is faulty; t_{19} is the hydraulic station failure if the coupling is damaged; t_{20} is the hydraulic system failure if the filter is clogged; t_{21} is if the pipeline is broken or blocked, hydraulic system failure; t_{22} is if the static pressure chamber floats abnormally, then the motion system fault; t_{23} is the contact surface wear, then the motion system fault; t_{24} strip If the pressure plate is loose or too tight, the motion system is faulty; t_{25} is if the servo system is faulty, then the control system is faulty; t_{26} is if the control circuit is faulty, then the control system is faulty; t_{27} is as Servo motor failure, actuator failure; t_{28} is the mechanical system failure if the transmission system fails; t_{29} is the mechanical system failure if the hydraulic station fails; t_{30} is the hydraulic system failure if the motion system fails; t_{31} is the electrical system fault if the control system fails; t_{32} is if the detection system fails, then the electrical System-like fault; t_{33} is the electrical system fault if the actuator fails; t_{34} is the mechanical system fault if the hydraulic system fails; t_{35} is the electrical system fault, the CNC machine fault; t_{36} is the mechanical machine fault if the mechanical system fails.

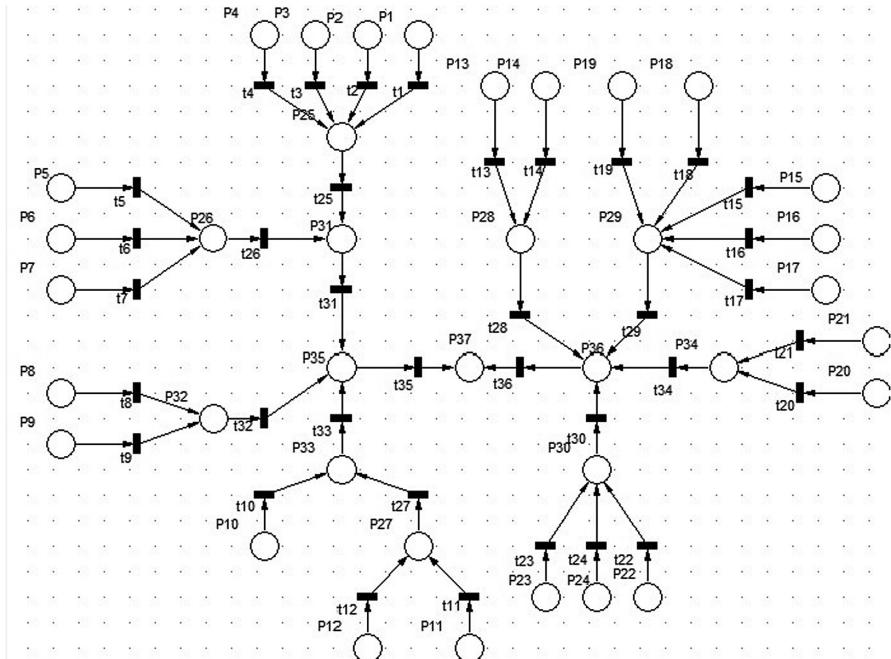


Fig. 1. Fault diagnosis model of CNC machine tools based on FPN

4 Analysis and Simulation of Electrical System Faults Based on FPN CNC Machine Tools

Take the electrical system fault of CNC machine tools as an example. The fuzzy Petri net model for fault diagnosis of CNC machine tools with hydraulic system fault is shown in Fig. 2. From Fig. 2, the immediate reachable set and reachable set of the propositions obtained by each place of electrical system fault diagnosis are shown in Table 1.

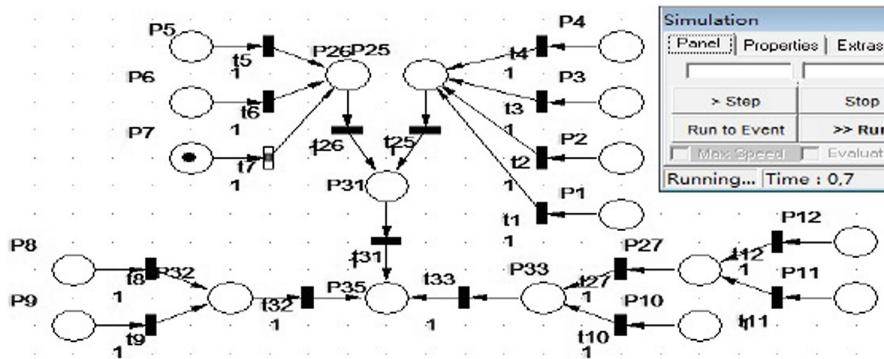


Fig. 2. FPN-based electrical control system failure cause place

Table 1. $IRS(pi)$ and $RS(pi)$

| $Place(pi)$ | $IRS(pi)$ | $RS(pi)$ |
|-------------|-----------|---------------------|
| $p1$ | { $p25$ } | { $p25, p31, p35$ } |
| $p2$ | { $p25$ } | { $p25, p31, p35$ } |
| $p3$ | { $p25$ } | { $p25, p31, p35$ } |
| $p4$ | { $p25$ } | { $p25, p31, p35$ } |
| $p5$ | { $p26$ } | { $p26, p31, p35$ } |
| $p6$ | { $p26$ } | { $p26, p31, p35$ } |
| $p7$ | { $p26$ } | { $p26, p31, p35$ } |
| $p8$ | { $p32$ } | { $p32, p35$ } |
| $p9$ | { $p32$ } | { $p32, p35$ } |
| $p10$ | { $p33$ } | { $p33, p35$ } |
| $p11$ | { $p27$ } | { $p27, p33, p35$ } |
| $p12$ | { $p27$ } | { $p27, p33, p35$ } |
| $p25$ | { $p31$ } | { $p31, p35$ } |
| $p26$ | { $p31$ } | { $p31, p35$ } |
| $p27$ | { $p33$ } | { $p33, p35$ } |
| $p31$ | { $p35$ } | { $p35$ } |
| $p32$ | { $p35$ } | { $p35$ } |
| $p33$ | { $p35$ } | { $p35$ } |
| $p35$ | Φ | Φ |

According to [6] fuzzy value range to determine the credibility of the transition and the excitation frequency of the transition, and according to the fuzzy inference rules, the calculated reverse excitation value is shown in Table 2.

Table 2. Credibility, excitation frequency, and reverse excitation value of transition

| μ_i | Numeric | γ_i | Numeric | $f(t_i)$ | Numeric |
|------------|---------|---------------|---------|-------------|---------|
| μ_1 | 0.97 | γ_1 | 0.84 | $f(t_1)$ | 0.81 |
| μ_2 | 0.93 | γ_2 | 0.95 | $f(t_2)$ | 0.88 |
| μ_3 | 0.95 | γ_3 | 0.91 | $f(t_3)$ | 0.86 |
| μ_4 | 0.96 | γ_4 | 0.97 | $f(t_4)$ | 0.93 |
| μ_5 | 0.90 | γ_5 | 0.97 | $f(t_5)$ | 0.87 |
| μ_6 | 0.92 | γ_6 | 0.87 | $f(t_6)$ | 0.80 |
| μ_7 | 0.98 | γ_7 | 0.96 | $f(t_7)$ | 0.94 |
| μ_{25} | 0.99 | γ_{25} | 0.41 | $f(t_{25})$ | 0.40 |
| μ_{26} | 0.93 | γ_{26} | 0.76 | $f(t_{26})$ | 0.70 |
| μ_{31} | 0.94 | γ_{31} | 0.89 | $f(t_{31})$ | 0.83 |

4.1 FNP-Based Electrical Control System Fault Diagnosis Reasoning Process

Reverse Reasoning

According to the failure phenomenon, the electrical system is reversely inferred. $\max[f(t_{25}), f(t_{26})] = f(t_{26})$, the transition t_{26} is preferentially reversely excited, and the input place p_{26} is the fault location place, and the corresponding proposition is the control circuit fault (d_{26}). With p_{26} as the goal, the reverse reasoning is continued. $\max[f(t_5), f(t_6), f(t_7)] = f(t_7)$, the transition t_7 is preferentially reversely excited, and the input place p_7 is the fault cause place. The corresponding proposition is that the power supply outputs low voltage or no voltage (d_7), and the reverse reasoning ends.

According to the reverse reasoning, the fault diagnosis reason of the fault phenomenon electrical control system fault is: $\beta(p_7) = d_7 \rightarrow \beta(p_4) = d_4 \rightarrow \beta(p_2) = d_2 \rightarrow \beta(p_5) = d_5 \rightarrow \beta(p_3) = d_3 \rightarrow \beta(p_{31}) = d_{31} \rightarrow \beta(p_1) = d_1 \rightarrow \beta(p_6) = d_6 \rightarrow \beta(p_{26}) = d_{26} \rightarrow \beta(p_{25}) = d_{25}$.

Positive Excitation

Reverse Reasoning The reason for the failure of the electrical control system is that the power supply outputs low voltage or no voltage. The maintenance personnel check the power supply output low voltage or no voltage and provide maintenance. According to the actual situation, the reliability of the fault cause place is given. Set the transition excitation threshold $\lambda = 0.76$. If the “power supply output low voltage or no voltage”, “very real”, give the place p_7 credibility $\alpha(p_7) = 0.97 > \lambda$, and then transition t_7 excitation, $\alpha(p_{26}) = \alpha(p_7) \times \mu(t_7) = 0.95 > \lambda$, transition t_{26} excitation, and finally

reach the fault phenomenon place p_{31} , and the cause of the fault is found. If the proposition “power supply output low voltage or no voltage”, “not too real”, give the place the feasibility $\alpha(p7) = 0.58 < \lambda$, the transition $t7$ can not be excited, indicating that the power supply output low voltage or no voltage is not the root cause of the electrical control system failure. Then check the fault cause parameter setting error ($d4$), and then perform fault diagnosis in turn until the cause of the fault is found.

4.2 FPN Visual Object N++ Simulation Analysis

According to the above diagnostic reasoning, the simulation results in Visual Object N++, the simulation results of the voltage control system failure are as follows:

In Fig. 2, the place $p7$ contains a token, indicating that the probable power supply voltage output low voltage or no voltage ($d7$) is established. When $\alpha(p7) = 0.97 > \lambda$, $t7$ is excited. At this time, the token moves from the place $p7$ to the place $p26$, as shown in Fig. 3. (a) shows that at this time, the place $p26$ contains a token, indicating that the proposition control circuit is faulty ($d26$). When $\alpha(p26) = \alpha(p7) \times \mu(t7) = 0.95 > \lambda$, $t26$ is excited. At this time, the token in the place $p26$ goes to the place $p35$, as shown in Fig. 3 (b), the last token arrives at the place $p35$, indicating that the electrical control system has failed. Throughout the simulation process, the fault cause place is the starting point through the fault location place, and finally reaches the electrical control system fault phenomenon place, as shown in Fig. 4.

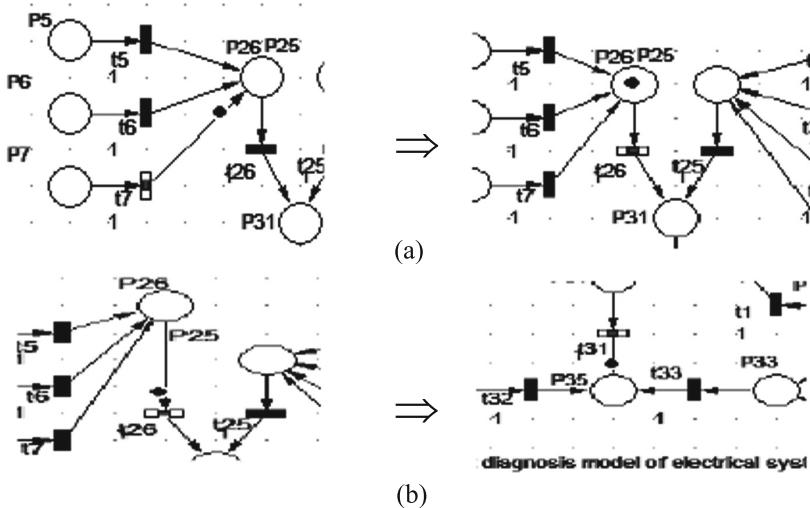


Fig. 3. FPN-based electrical control system fault diagnosis process

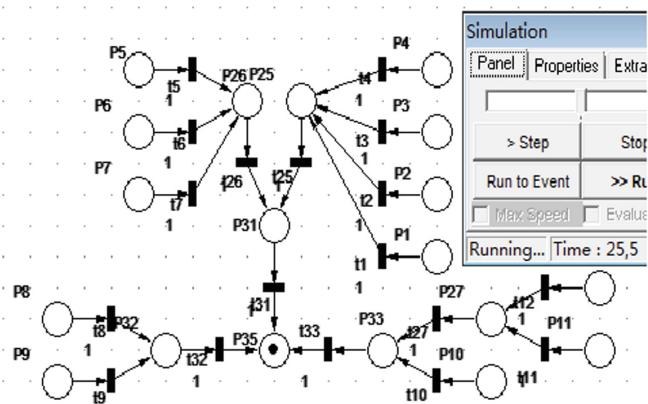


Fig. 4. FPN-based electrical control system fault phenomenon place

Based on the traditional FPN fault diagnosis, this paper introduces the concept of transition reverse excitation, which has a greater improvement on the rapidity of fault diagnosis, and the possibility of timely troubleshooting and improving the productivity of CNC machine tools. Based on the improved fuzzy Petri net, there are both quantitative analysis and qualitative analysis. The combination of computational reasoning and simulation, logical coherence and clarity, has laid a good foundation for the realization of computer software. At the same time, FPN can also be based on actual The situation is necessary to increase or decrease the necessary system failure of the CNC machine tool, so that the fault information of the CNC machine tool is continuously improved, and it is suitable for various CNC machine tools.

References

1. Yan CY (2001) Petri net principle and application. Electronic Industry Press, Beijing
2. Yu X (2015) Research on fault diagnosis expert system of air compressor based on web. Shandong Chemical Industry
3. Zhao SS, Lin XX, et al (2018) Title fault diagnosis of transition network based on fusion of time sequence and hierarchical transitional WFPN. In: Proceedings of 2018 Chinese intelligent systems conference. Lecture Notes in Electrical Engineering, vol 529, no 1, pp 143–151
4. Tong X, Xie H, Sun M (2017) Power system fault diagnosis model based on layered considering temporal constraint checking. Autom Electr Power Syst 37(2):63–68
5. Lin HY (2017) Design of fault diagnosis expert system for numerical control machine. Xihua University, Sichuan
6. Fan QM (2016) A fault diagnosis method for distillation column. Shanghai Institute of Technology, Shanghai
7. Shen GX, Wang ZQ, Zhang YZ, et al (2012) Fault diagnosis of CNC machine tool spindle based on fuzzy petri net. Manuf Technol Mach Tools 20(3):124–127

8. Xiang YS, Liu W, Yue XB, Qin N (2009) Research on the auto fault diagnosis simulation based on fuzzy petri nets. *Comput Eng Sci* 31(3):86–88
9. Liu QH (2019) Design of fault analysis and diagnosis system for CNC machine tools. *China's New Technology and New Products*
10. Wang XF (2018) Electrical fault diagnosis and maintenance of CNC machine tools. *Hubei Agric Mech* 12:17



Application of CMAC in Sensorless Control of PMSM

Qiangqiang Xu, Bulai Wang^(✉), and Yuanyuan Yang

School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai, China
wruoyi@163.com

Abstract. This study proposes a sensorless control method for permanent magnet synchronous motor (PMSM) based on a cerebellum model articulation controller (CMAC). Using high-frequency injection method to identify the magnetic field and rotational speed of the rotor. The cerebellum model neural network and PID composite controller are used to replace PI controller in speed loop. CMAC has a generalization ability from the beginning of input, suitable for a nonlinear mapping relationship. The gradient-type learning algorithm is used to train the CMAC in order to implement PID parameters online self-tuning. The simulation results show that the addition of CMAC algorithm can make the system have fast response, small overshoot and strong robustness.

Keywords: Permanent magnet synchronous motor · High-frequency injection method · Cerebellar model neural network

1 Introduction

The speed-sensorless drive of the motor has always been a research hotspot, and it has obvious advantages of low cost, easy maintenance and small volume compared with the motor control system with speed sensor. The traditional PMSM speed-sensorless vector control system usually adopts double closed loop PI control, and the algorithm is simple and widely used in industrial control field. However, PMSM is a multivariable, nonlinear, and strongly coupled control object. When the motor parameters change or are affected by external disturbances, it is difficult for the PID controller to obtain good performance for an entire speed range. The cerebellar model neural network has local generalization ability, fast response, simple learning algorithm, good nonlinear approximation ability, and can overcome the shortcomings of conventional PI controllers when controlling nonlinear systems.

At present, the application of the motor control field to the cerebellar model neural network is less, and related research mostly focuses motor control with speed sensor. In [2], CMAC-PID is applied to the BLDC motor control system. The experimental results prove that the hybrid CMAC-PID algorithm can improve dynamic performance of the control system, reduce overshoot, and has strong anti-interference ability. In [3], the PID controller based on CMAC neural network replaces the traditional PMSM speed outer loop PI controller to improve system performance. Literature [4] applied CMAC neural network to high-precision servo system, which improved the system's

high-precision dynamic tracking performance. In [5], CMAC-PID is applied to the vector control system of three-phase asynchronous motor. The gradient learning algorithm is used to adjust the weight of the cerebellar model neural network to realize real-time adaptive identification of motor speed. Literature [6] introduced BP neural network into the speed PI controller of permanent magnet synchronous linear motor system, which improved the system performance.

2 Mathematical Model

There are many uncertainties and nonlinear terms in the electromagnetic relationship of the motor in practical applications. The mathematical model of PMSM is simplified, and PMSM is in the d-q coordinate system [7]. The model is as follows:

$$\begin{aligned} u_d &= R_s i_d - p\omega_r L_q i_q + \frac{di_d}{dt} L_d \\ u_q &= R_s i_q + p\omega_r L_d i_d + p\omega_r \psi_f + \frac{di_q}{dt} L_q \\ \frac{d\omega_r}{dt} &= \frac{1}{J} (T_e - T_L - B\omega_r) \\ T_e &= 1.5p[\psi_f i_q + (L_d - L_q)i_d i_q] \end{aligned} \quad (1)$$

From the above formula, there is a mutual constraint between the motor back EMF, phase current, rotor speed and load torque. These quantities are often time-varying in the actual control system, and interference is often introduced in the system. PI controllers are not suitable for solving this complex nonlinear problem. The neural network control effect does not depend on the exact model. The cerebellar model neural network and PID intelligent control strategy of the permanent magnet synchronous motor proposed in this paper can take into account the advantages of simple and rapid PID regulation and the nonlinear approximation ability of neural networks.

3 Hybrid CMAC-PID Controller

3.1 A Subsection Sample

CMAC is an adaptive control model that imitates brain connections. After Miller successfully applied CMAC neural network to robot control, cerebellar model neural network has developed rapidly. After nearly 40 years of research and exploration, CMAC has a wide range of applications in mechanical control, intelligent identification, signal processing and so on.

The working principle of CMAC is shown in Fig. 1. The input state of the system is taken as a pointer, and the input space is divided into a number of partitions, each of which corresponds to the address of the memory. These blocks are learned by the algorithm, and the corresponding information is obtained, and then there is a memory address of the adjacent block. Usually the number of storage units is much less than the maximum possible number of partitions in the input space. Therefore, this mapping

relationship is many-to-one, that is, multiple partitions are mapped at the same memory address.

The architecture of CMAC is shown in Fig. 1, it consists of four parts.

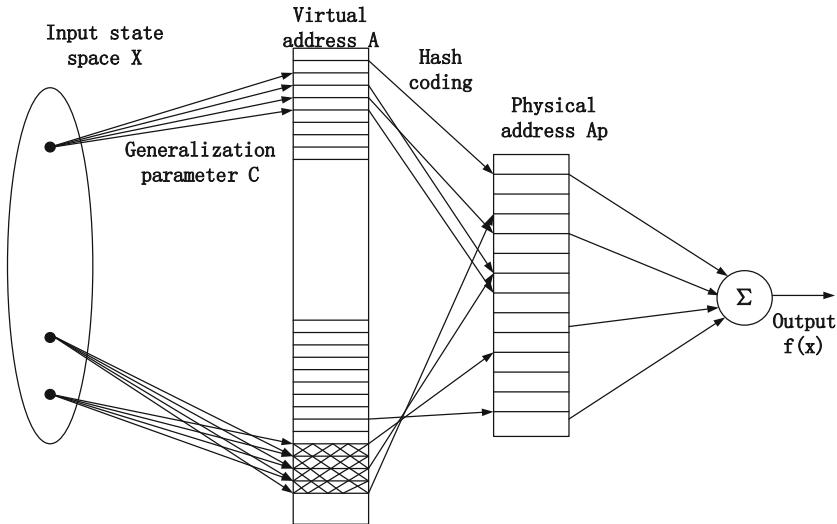


Fig. 1. The architecture of the CMAC.

- (1) Input vector X quantization. The signal of the X is measured by the sensor. The number of the perceptron is called the quantization level. For any input, the perceptron of the corresponding quantized value can be found. Near the input quantized values, there are C perceptrons that are simultaneously excited to output non-zero values. C is called a generalization parameter, which characterizes the size of the area within the network that affects the output of the network.
- (2) X to A address mapping. According to the principle that the closer vectors in the input space are similar in the output space, the input is obtained. The virtual address of the A memory corresponding to the sample. If the input vector contains m components, each component N quantization levels, then A will require storage spaces.
- (3) A to Ap spurious mapping. Not all neurons are activated, the neurons excited in A are sparse, and the spurious technique is used to compress A into a relatively smaller real space Ap , which is the storage area of the actual deposit right. The probability of address collisions in the spurious process can be reduced by appropriately selecting the C and Ap sizes.
- (4) Ap to f network output. The C memory cells in the Ap are randomly arranged, and the input is mapped to the weighted sum output stored in the C cells of the actual memory.

Therefore, when studying the CMAC model, we will study CMAC mapping system, physical structure of the front end and learning algorithm of the back end. It can be

seen that the network has a clear network connection and strong local generalization. Capabilities, simple LMS learning algorithms, and faster convergence speeds.

3.2 The Design of CMAC-PID Controller

The schematic of CMAC-PID controller is given in Fig. 2.

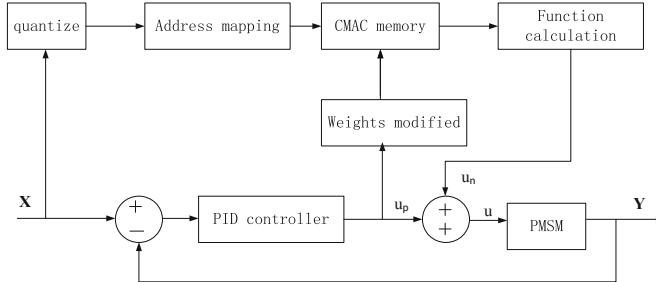


Fig. 2. Schematic of CMAC-PID controller.

The input vector of CMAC-PID intelligent controller is represented by a point $X^p = (x_1^p, x_2^p, \dots, x_n^p)^T$ in the input state space X , X can be quantized into q levels, and a possible input value is stored in the virtual memory address A , and the number of cells in the A storage area is m , $m \geq q^n$. Each virtual memory address in A corresponds to a point in X . A with q^n memory locations is mapped into a smaller physical address space by hash encoding. There are C random addresses for each input sample in A_p . The output $u_p(k)$ of the PID controller is obtained through the Jacobian iteration to obtain the network connection right, which is memorized by the CMAC, and the stored weight is calculated by the CMAC function to obtain the output $u_n(k)$ of the CMAC.

$u_n(k)$ is calculated every cycle, compared with the total control input $u(k)$, and the weight is corrected for learning. By continuously learning, the difference between $u(k)$ and $u_n(k)$ is close to 0. Finally $u(k)$ is equivalent to being generated by the CMAC.

Controller algorithm is as follows:

$$u_n(k) = \sum_{i=l}^c w_i a_i \quad (2)$$

$$u(k) = u_n(k) + u_p(k) \quad (3)$$

a_i indicates neuron status. If the neuron i is active, a_i is 1; If it is in a suppressed state, a_i is 0. The weight adjustment goal is to make $E(k)$ close to 0 by δ rule. The objective function is:

$$E(k) = \frac{1}{2} (u(k) - u_n(k))^2 \cdot \frac{a_i}{c} \quad (4)$$

$$\Delta w(k) = \eta \frac{u(k) - u_n(k)}{c} a_i = \eta \frac{u_p(k)}{c} a_i \quad (5)$$

$$w(k) = w(k-1) + \Delta w(k) + \alpha(w(k) - w(k-1)) \quad (6)$$

Where η is the network learning rate, α is the inertia amount; $E(k)$ is the CMAC network output error at time k , $w(k)$ is the weight at time k , and $w(k-1)$ is the weight at time $k-1$ value, $\Delta w(k)$ is the weight adjustment at time k .

4 Hybrid CMAC-PID Controller

To check the feasibility of CMAC-PID algorithm, simulations have been done in the MATLAB. System simulation block diagram shown in Fig. 3.

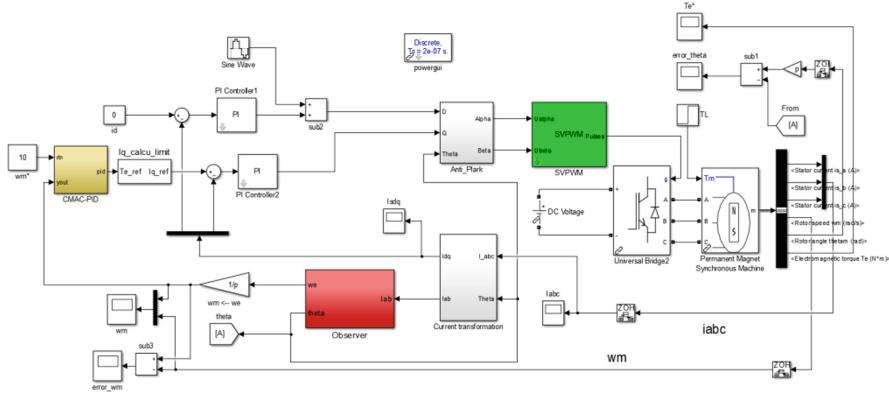


Fig. 3. System simulation block diagram in MATLAB.

The motor parameters of the system simulation are set as: $p = 2$, $L_d = 4.25 \text{ mH}$, $L_q = 8.25 \text{ mH}$, $R_s = 0.92\Omega$, $\psi_f = 0.45 \text{ Wb}$, $B = 0.008 \text{ N} \cdot \text{m} \cdot \text{s}$

The simulation conditions are set as: DC voltage $U_{dc} = 311 \text{ V}$, PWM switching frequency is 10 kHz, sampling period $T_s = 10 \mu\text{s}$, the frequency of the injected high-frequency voltage signal is 1 kHz, and the amplitude is 20 V.

To observe dynamic performance at low speed, the speed is set to 10r/min, no-load start, and the rated load torque is 4 N at $t = 0.6 \text{ s}$.

In the simulation, PMSM speed, rotor angle error and the magnetic electric torque response are displayed in Figs. 4 and 5. Figure 4 The red dotted line is the estimated speed, and the blue solid line is the measured speed.

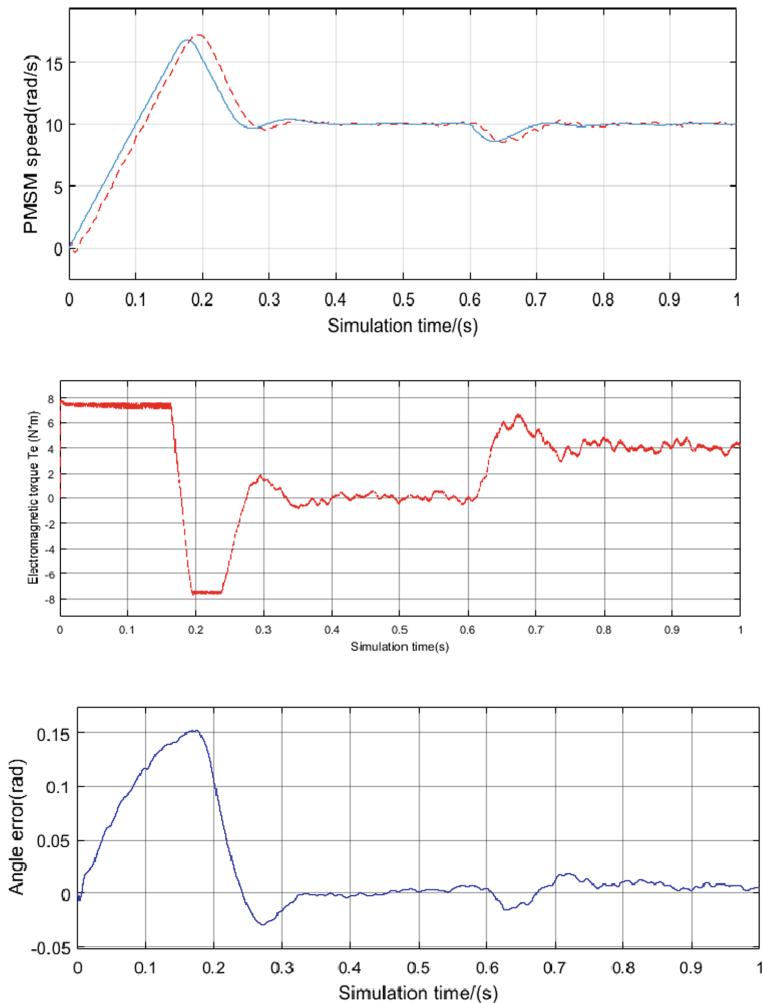


Fig. 4. Simulation results of PI.

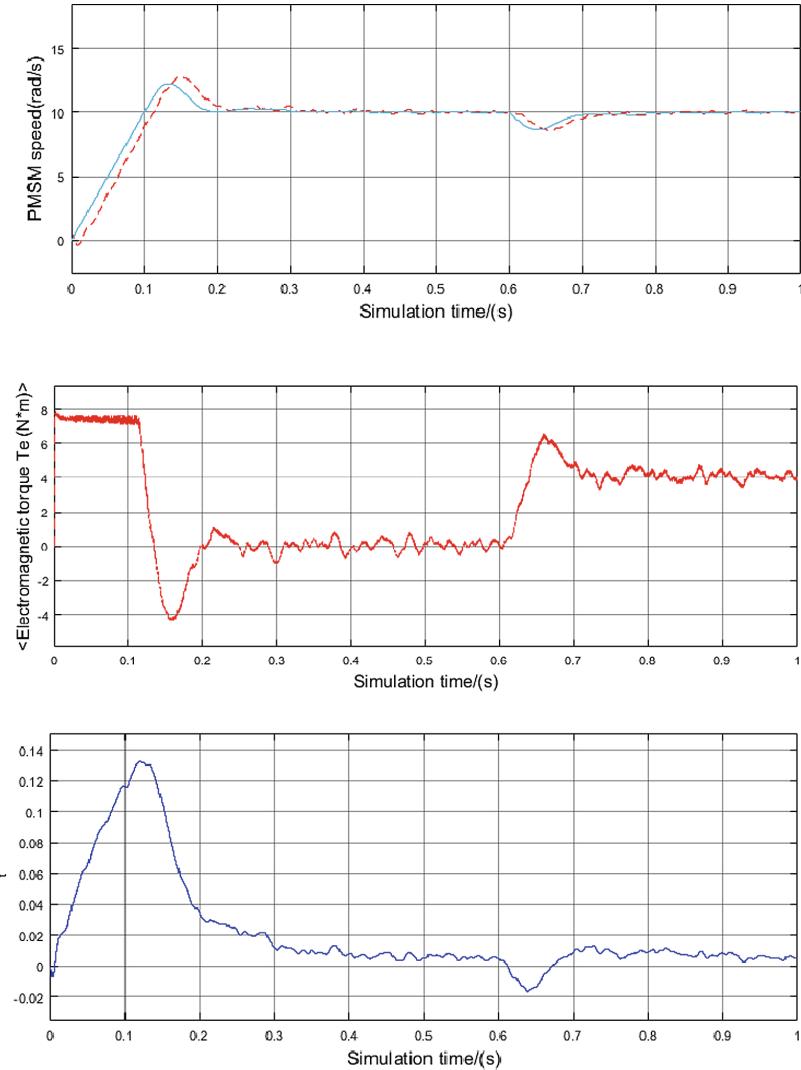


Fig. 5. Simulation results of CMAC-PID.

From Figs. 4 and 5, compared with the PI controller, the CMAC-PID controller reduces the overshoot time by 0.1 s and the overshoot by 50%. That proves that the CMAC-PID controller has better dynamic performance, small overshoot and strong robustness.

5 Conclusions

A CMAC-PID controller for PMSM speed-sensorless control is proposed. The simulation of permanent magnet synchronous motor under CMAC-PID controller is carried out in MATLAB. Simulations results show that CMAC-PID controller is faster than response to traditional PI controller, with small response, strong robustness and good anti-interference ability.

References

1. Wamique S, Kumar S, Nirala SK, et al (2013) Sensorless vector control of induction motor drive for electrical vehicle application. In: International conference on emerging trends in electrical engineering & energy management. IEEE
2. Yue X, Bai P, Yang R, et al (2016) Application of CMAC-PID in brushless DC motor control system. Instrum Tech Sens (7)
3. Niu H, Zhao R, Shang Z, et al (2006) Research on the control method of permanent magnet synchronous motor based on CMAC neural network. Electric Drive 36(5):19–21
4. Liu Y, Wang W (2006) Research on servo system control based on CMAC. Acta Aeronautica Sinica 27(3)
5. Tsai CH, Yeh MF (2009) Application of CMAC neural network to the control of induction motor drives. Appl Soft Comput 9(4):1187–1196
6. Li Y, Yang S (2018) PI speed control of permanent magnet synchronous linear motor based on neural network. Jiangxi Chem Ind 139(05):129–131
7. Yuan L, Hu B, Wei K, et al (2016) Modern permanent magnet synchronous motor control principle and MATLAB simulation. Beihang University Press



Uniform Demagnetization Fault Diagnosis in Permanent Magnet Synchronous Motors Based on the Back-EMF Gyration Radius

Yanjie Nie, Caixia Gao^(✉), and Ziyi Fu

School of Electrical Engineering and Automation, Henan Polytechnic University,
Jiaozuo 454000, China
gcx81@126.com

Abstract. Uniform demagnetization fault is troublesome, because it can profoundly impact the overall performance of permanent magnet synchronous motors (PMSMs). In this paper, the back electromotive force (Back-EMF) gyration radius method is used to diagnose the uniform demagnetization fault of the PMSM. First, a model of the Back-EMF gyration radius is established, which using time delay embedding (TDE) technology. Second, the Back-EMF gyration radius of PMSM with different operating speeds and different faults is calculated, and the rate of change of the back-EMF gyration radius is taken as the uniform demagnetization fault factor which is used to detect uniform demagnetization fault and fault degree. Finally, the simulation results show that proposed method can achieve PMSM uniform demagnetization fault detection and degree evaluation with a small amount of calculation.

Keywords: Back electromotive force (Back-EMF) gyration radius · Uniform demagnetization fault · Fault factor · Permanent magnet synchronous motor (PMSM)

1 Introduction

PMSM is widely used in electric vehicles, industrial robots, and aerospace fields, because of their high-power density, high efficiency, and ease of control [1]. However, permanent magnets are subjected to local or uniform demagnetization because of the combined effects of armature reaction, operating temperature rise and brittleness of sintered rare earth permanent magnet material [2–4]. Once the permanent magnet has irreversible demagnetization fault, the motor will have a serious heat, the overload capacity will decline, as will lead to performance degradation, and in serious cases the motor may be out of control and scrapped. The diagnosis of the early demagnetization fault of the motor can improve the reliability of the motor operation, guide the maintenance, prolong the service life and reduce sudden accident shutdown. Therefore, the study of PMSM demagnetization fault detection and accurate evaluation of fault degree is of great significance.

In recent years, many scholars have done a lot of research on PMSM demagnetization fault. In [5, 6], using advanced signal processing technology to extract stator

current harmonics from demagnetization fault, the local demagnetization fault diagnosis of PMSM permanent magnet is realized. In [7], the local demagnetization fault diagnosis of PMSM based on zero-sequence voltage component. In [8], proposes a detection technique for irreversible demagnetization using the harmonics characteristics of the back electromotive force (BEMF) in brushless dc motors (BLDCMs). However, these methods are not applicable in the case of uniform demagnetization fault, because they rely on the asymmetry caused by local demagnetization fault.

In order to diagnose the uniform demagnetization fault. In [9], a noninvasive approach to detect the uniform demagnetization fault for SPMSM based on the acoustic noise analysis from a back propagation neural network (BPNN) model is presented. But it is easily affected by environmental noise. In [10, 11], the change of PMSM electrical characteristics caused by the change of PMSM magnetic circuit state before and after demagnetization of permanent magnets is used as the diagnostic criterion for demagnetization fault of permanent magnets by high frequency signal injection, and the diagnosis of uniform demagnetization fault is realized. However, this method needs to reasonably determine the amplitude and frequency of the injected high-frequency signal, which affects the accuracy of diagnosis. In [12], According to acquiring the inductive potential of stator tooth flux by embedding the detecting coil on each stator tooth, and the diagnosis of uniform demagnetization fault is realized. However, this method increases the volume and cost of the motor by installing a number of detection coils in the motor manufacturing phase. In [13], by analyzing the gyration radius of the cogging torque signal, an index for diagnosing uniform demagnetization fault is obtained. However, due to the different structure of the motor, the cogging torque will have a complex waveform when in the fractional slot concentrated winding motor, therefore it is impossible to diagnose the uniform demagnetization fault through the gyration radius of the cogging torque.

Aiming at the PMSM permanent magnet uniform demagnetization fault, and this paper proposes a diagnosis method based on the back-EMF gyration radius. TDE technique to obtain the coordinate data of the differential phase space diagram of the Back-EMF, and the relationship between the rate of change of the back-EMF gyration radius and demagnetization degree at different velocities is analyzed. Simulation analysis results prove the feasibility and effectiveness of the proposed method.

2 Structure and Parameters of PMSM

This paper focuses on the uniform demagnetization fault, which in a surface-mounted PMSM with 66 poles and 72 slots. Table 1 shows the parameters of the motor. The motor structure is shown in Fig. 1. Note that stator uses fractional-slot concentrated teeth-separate winding, and each phase is composed of three branches in parallel, and each branch routes four coils in series. The magnetic pole adopts surface mounted equidistant distribution. The motor adopts NdFeB as the permanent magnet materials. The simulation analyzes the uniform demagnetization fault of PMSM, and realizes different demagnetization fault degree of PMSM permanent magnet by changing the coercivity (H_c) size.

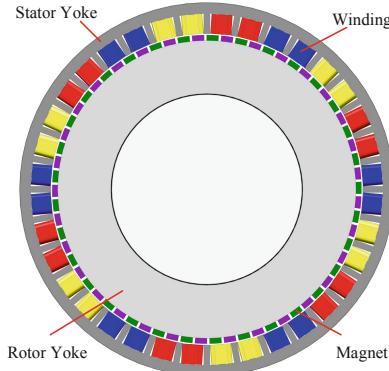


Fig. 1. Structure of the PMSM

Table 1. Primary parameters of the PMSM.

| Items | Values | Unit |
|-----------------------------|--------|-----------------|
| Out diameter of stator | 360 | mm |
| Inner diameter of stator | 300 | mm |
| Air-gap length | 1.2 | mm |
| Wire diameter of winding | 1.3 | mm |
| Area of slot | 154 | mm ² |
| Thickness of PM | 5.3 | mm |
| Angle of PM | 5 | ° |
| Axial length | 150 | mm |
| Rated power | 10 | kW |
| Rated speed | 200 | rpm |
| Number of phases | 3 | / |
| Number of coils | 36 | / |
| Coil turn | 48 | / |
| Parallel-circuits per phase | 3 | / |
| Slot-Pole combination | 72-66 | / |

3 The Back-EMF Gyration Radius Model

As a data mining technology, and time delay technique is often used to process time series data. Based on dynamic system theory, TDE method reveals hidden information in time series data [14]. To get the Back-EMF gyration radius, first, using TDE transformation processing on the time domain data, a Back-EMF delay sequence is obtained. Then, the coordinates of each point are obtained by using the delay sequence differential phase space transformation. The Back-EMF gyration radius is the square root of the average value of the sum of squares of the distance from each coordinate point to the center point.

3.1 The Differential Phase Space of Back-EMF

The TDE transformation of Back-EMF is shown as follows in (1) (2). $\Delta t(k-l)$ is abscissa and $\Delta t(k)$ is ordinate of a 2D diagram of the Back-EMF differential phase space, which is transformed from a one-dimensional time series into a two-dimensional phase space.

$$\Delta t(k) = t(k) - t(k-l) \quad (1)$$

$$\Delta T = \{\Delta t(k), k = l, l+1, \dots, N\} \quad (2)$$

Where $t(k)$ is the time series of the Back-EMF, ΔT is the difference sequence of delayed embedding of counter electromotive force.

3.2 Determination of Delay Time

The difference phase space diagram of the Back-EMF of health and different demagnetization fault degree under different delay time is shown in Fig. 2, and it can be seen that different delay times have a great influence on the size of the Back-EMF gyration radius. In order to determine the optimal delay time, this paper calculates and analyzes the Back-EMF difference phase space under different delay time [15]. When the delay time is delayed $l = 3$, and the difference phase space diagram of Back-EMF is more regular. Meanwhile, the calculated Back-EMF gyration radius has a more linear relationship with the demagnetization degree, and the precision is higher. Therefore, $l = 3$ was determined as the delay time of this study.

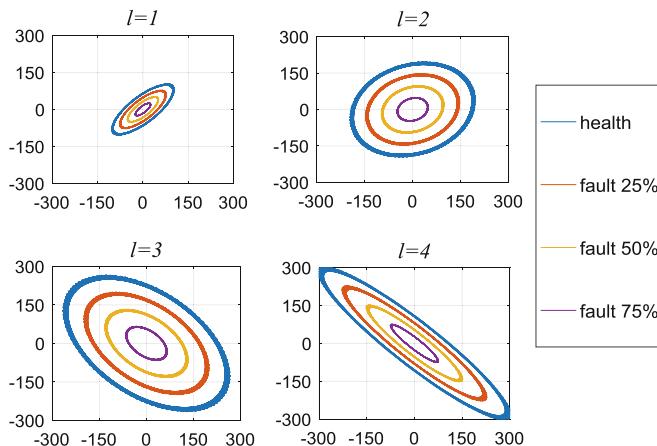


Fig. 2. The differential phase space diagram of different delay time

3.3 The Back-EMF Gyration Radius

The mathematical model of the Back-EMF gyration radius is shown as follows: (3) (4) (5).

$$x_0 = \frac{\sum_{k=l}^{N-l} \Delta t(k)}{N - l}, y_0 = \frac{\sum_{k=l+1}^N \Delta t(k)}{N - l} \quad (3)$$

$$d(k)^2 = [\Delta t(k) - x_0]^2 + [\Delta t(k - l) - y_0]^2 \quad (4)$$

$$r = \sqrt{\frac{\sum_{k=l+1}^N d(k)^2}{N - l}} \quad (5)$$

Where (x_0, y_0) is central coordinates, and r is the Back-EMF gyration radius, $d(k)^2$ is the distance between each point and the center of mass in the 2D diagram of the differential phase space of the Back-EMF.

4 Analysis of Simulation Result

4.1 Differential Phase Space Diagram Analysis of Back-EMF

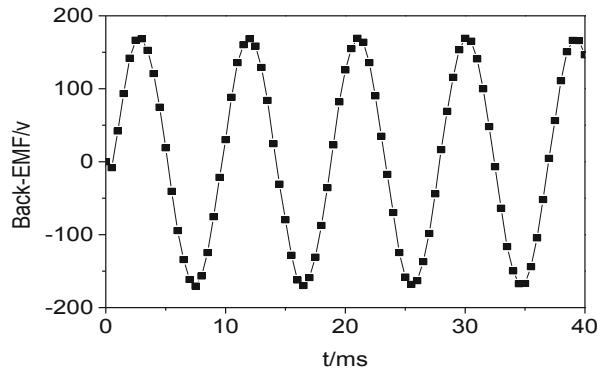
The time domain of the Back-EMF under health condition and the spectrum graph of its FFT analysis are shown in Fig. 3, and it can be seen that spectrum graph is simple. Meanwhile, the fundamental wave is the dominant component of the spectrum, and the amplitude of the fundamental wave is much larger than the harmonic amplitude.

The difference phase space diagram of Back-EMF of under different demagnetization degrees is shown in Fig. 4, and it can be seen that the area of the diagram and Back-EMF gyration radius decrease regularly with the increase of demagnetization fault degree. Meanwhile, the difference phase space diagram presents a regular oval, and the data point coordinates of the diagram are all on the edge of the circle, and the relation between the Back-EMF gyration radius and the demagnetization degree is obvious.

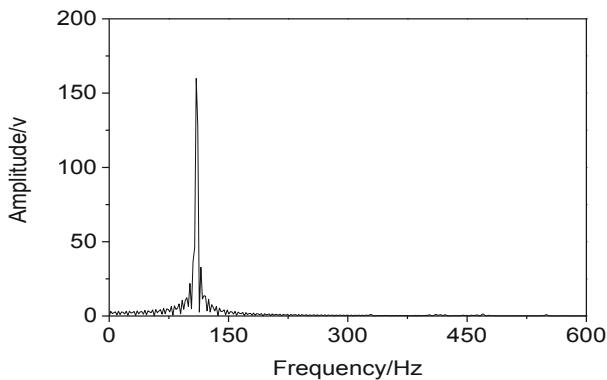
4.2 Extraction of the Uniform Demagnetization Fault Factor

The rate of change of the Back-EMF gyration radius is a quantity which the radius changes with the increase of demagnetization degree. The rate of change of the back-EMF gyration radius is taken as detect the uniform demagnetization fault factor. The fault feature is shown in formula (6):

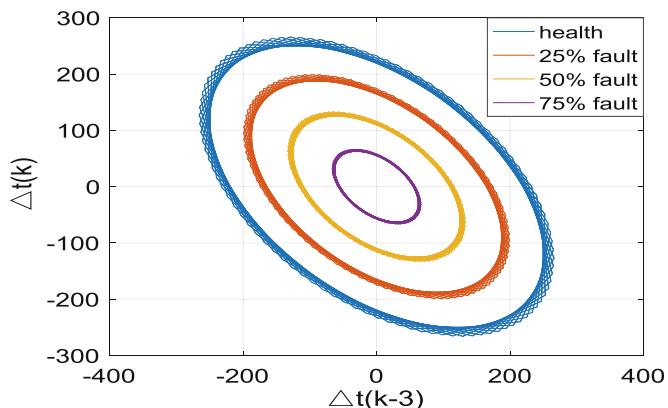
$$W = \frac{r_{health} - r_{fault}}{r_{health}} \quad (6)$$



(a) The time domain graph of Back-EMF



(b) The spectrum graph of Back-EMF

Fig. 3. The time domain graph and spectrum graph of the back-EMF**Fig. 4.** The Back-EMF differential phase space diagram under different demagnetization degrees

Where W is fault feature, r_{health} is the back-EMF gyration radius in health state, r_{fault} is the back-EMF gyration radius in fault state. The relationship between the degree of demagnetization fault at different speeds and the Back-EMF gyration radius is shown in Fig. 5, and it can be seen that the Back-EMF gyration radius decreases with the increase of demagnetization, and has a good linear relationship.

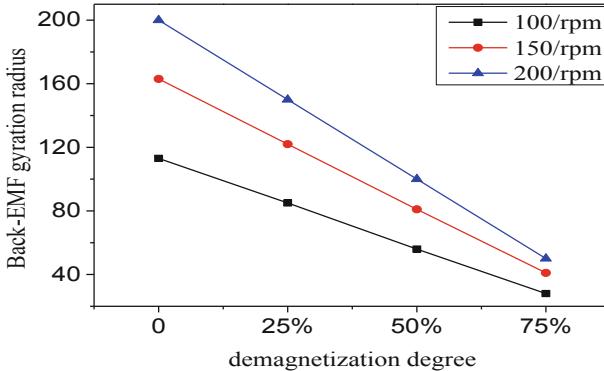


Fig. 5. The Back-EMF gyration radius of different demagnetization degrees

As shown in Table 2, and it can be seen that the rate of change of Back-EMF gyration radius at different speeds is basically consistent with the demagnetization degree of permanent magnet. Therefore, the rate of change of Back-EMF gyration radius can accurately detect and represent the degree of demagnetization.

Table 2. The rate of change of Back-EMF gyration radius

| Demagnetization degree | The rate of change(W) | | |
|------------------------|---------------------------|---------|---------|
| | 100/rpm | 150/rpm | 200/rpm |
| 0 | 0.1% | 0% | 0% |
| 25% | 25.2% | 25.2% | 25% |
| 50% | 50.4% | 50.1% | 50% |
| 75% | 75.3% | 75.4% | 75.1% |

5 Conclusion

This paper presents a fault diagnosis method for uniform demagnetization of PMSM based on the Back-EMF gyration radius. The change rule of Back-EMF differential phase space and Back-EMF gyration radius before and after the PMSM under different degrees of uniform demagnetization fault are analyzed. The analyzes results shows the Back-EMF gyration radius decreases linearly with the increase of demagnetization

faults of permanent magnets, and the rate of change of Back-EMF gyration radius is consistent with the degree of demagnetization, and the feasibility of the proposed method is proved.

Acknowledgments. This work was supported by the Key Science and Technology Program of Henan Province (No. 192102210230).

References

1. Boazz B, Pellegrino G, Vagati A et al (2014) Multipolar SPM machines for direct-drive application: a general design approach. *IEEE Trans Ind Appl* 50(1):327–337
2. Faiz J, Mazaheri-Tehrani E (2017) Demagnetization modeling and fault diagnosing techniques in permanent magnet machines under stationary and nonstationary conditions: an overview. *IEEE Trans Ind Appl* 53(3):2772–2785
3. Li X, Wang S (2017) Demagnetization analysis of the interior permanent magnet synchronous motor under different short circuit faults. *J China Coal Soc* 42(S2):626–632
4. Si J, Zhang L, Feng H et al (2017) Analysis of 3-D temperature field for surface-mounted and interior permanent magnet synchronous motor. *Electric Mach Control* 21(3):25–31
5. Wei Z, Ma H, Chen C et al (2017) Diagnosis method based on CWT-HHT of demagnetization fault for permanent magnet synchronous motor and analyzing its feasibility. *Electric Mach Control* 44(8):81–87
6. Li H, Chen T (2017) The local demagnetization fault diagnosis of PMSM based on fractal dimension. *Trans China Electrotech Soc* 32(07):1–10
7. Urresty J, Riba Ruiz J-R, Delgado M et al (2012) Detection of demagnetization faults in surface-mounted permanent magnet synchronous motors by means of the zero-sequence voltage component. *IEEE Trans Energy Convers* 27(1):42–51
8. Kim H-K, Kang D-H, Hur J (2015) Fault detection of irreversible demagnetization based on space harmonics according to equivalent magnetizing distribution. *IEEE Trans Magn* 51 (11):1–4
9. Zhu M, Hu W, Kar NC (2018) Acoustic noise-based uniform permanent-magnet demagnetization detection in SPMSM for high-performance PMSM drive. *IEEE Trans Transp Electrif* 4(1):1397–1405
10. Hong J, Hyun D, Lee SB et al (2010) Automated monitoring of magnet synchronous motors at standstill. *IEEE Trans Ind Appl* 46(4):1397–1405
11. Hong J, Hyun D, Kang T-J et al (2012) Detection and classification of rotor demagnetization and eccentricity faults for PM synchronous motors. *IEEE Trans Ind Appl* 48(3):923–931
12. Da Y, Shi XD (2013) A new approach to fault diagnostics for permanent magnet synchronous machines using electromagnetic signature analysis. *IEEE Trans Power Electron* 28(8):4014–4112
13. Nejadi-Koti H, Faiz J, Demerdash NAO (2017) Uniform demagnetization fault diagnosis in permanent magnet synchronous motors by means of cogging torque analysis. In: 2017 IEEE international electric machines and drives conference, pp 1–7
14. Bangura JF, Povinelli RJ, Demerdash NAO et al (2003) Diagnostics of eccentricities and bar/end-ring connector breakages in poly phase induction motors through a combination of time series data mining and time-stepping coupled FE-state-space techniques. *IEEE Trans Ind Appl* 39(3):1005–1013
15. Xiu C, Liu X, Zhang Y (2003) Selection of embedding dimension and delay time in the phase space reconstruction. *Trans Beijing Inst Technol* 23(2):119–224



Dynamic Gesture Capture, Location and Tracking Based on MeanShift Algorithm

Shenjian Gong, Changdi Li, Junyi Hou, and Lei Yu^(✉)

School of Mechanical and Electric Engineering, Soochow University,
Suzhou 215000, China
slender2008@163.com

Abstract. In this paper, dynamic gesture capture and tracking based on MeanShift algorithm for human-computer interaction system is proposed. In this method, the rectangular box representing the hand position is used as the control signal input and the initial detection target tracking. Thus the dynamic gesture video is effectively processed. In this paper, Python3 OpenCV library is adopted as the main processing library for experimental analysis. The initial capture and positioning of hand position and subsequent tracking experiments are carried out on a video of hand movement. Through the comparison and analysis with Camshift algorithm, the experimental results show that MeanShift algorithm achieves a good dynamic tracking effect of gesture position.

Keywords: MeanShift algorithm · Rectangular box · Gesture tracking · CamShift algorithm

1 Introduction

Human-computer interaction is a study of the interaction between the system and the user. At present, the mainstream human-computer interaction system control methods in the market are mainly divided into speech recognition control and image recognition control. Of course, there must be acoustic noise and image noise in the interactive environment. Under the same condition, the development of image filtering technology is superior, and the application scenes of speech recognition are few. The human-machine interaction technology based on machine vision collects image data through visual sensors, then by filtering and identification, detects and tracks objects, and links interactive feedback, which makes the human-machine interaction technology and link motion control interaction comfortable. With the development of machine vision, virtual reality, pattern recognition and target tracking, the application of these technologies also involves the design of human-computer interaction system and gesture tracking.

Gesture tracking algorithm is derived from target tracking technology. From the perspective of color image tracking method, it can be divided into three categories: tracking based on feature information, prediction tracking based on filtering method, and optimization tracking based on matching search. A tracking algorithm based on feature information, Lindeberg et al. proposed a tracking algorithm based on Scale invariant feature transform (SIFT), which is used to describe the target to be tracked in

motion [1–3]. Under the condition of uneven light, due to the pixel gradient becomes irregular, the same object in different light conditions may be unable to match the situation, and the scale of the operation objects invariant features and real-time matching. This has taller requirement to computational capacity. Computing equipment cannot fulfill the condition of real time with high probability. Predictive tracking based on filtering method, as described in the name, by capturing the motion trajectory of the front object, it is used to predict the position of the object at the next moment, so as to narrow the search scope and search in the small range where the object has a high probability. Because the search scope is greatly reduced, the computation is small, and the real-time requirements can be fulfilled, the capture of most motion trajectories is locally optimal. At present, the particle filter proposed by Alamsyahcan can achieve global optimization rather than local optimization [2], but at the same time, the real-time effect decreases with the continuous iteration of particle filter.

This paper presents an algorithm for gesture tracking control in a large screen interactive system. In this method, the initial gesture is captured, the rectangular box is used as the initial detection target tracking, the MeanShift is used in the image target tracking algorithm for target recognition, and the dynamic gesture video processing is used to complete the dynamic gesture location tracking of the interactive system [2–6].

2 Application of MeanShift Algorithm in Image Target Tracking

The MeanShift algorithm is a kernel density estimation algorithm that moves each point to the local maximum point of the density function, that is, the point where the density gradient is 0, also called the mode point.

The kernel density estimate at d-dimensional data point x is expressed as:

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (1)$$

In Eq. 1, $K(x)$ is a scaling function that satisfies the following four conditions:

$$\sup_{y \in R^d} |K(y)| < \infty = \quad (2)$$

$$\int_{R^d} |K(y)| dy < \infty \quad (3)$$

$$\int_{R^d} K(x) dx = 1 \quad (4)$$

$$\lim_{\|y\| \rightarrow \infty} \|y\|^d K(y) = 0 \quad (5)$$

In the formula, $\|\cdot\|$ represents the Euclidean distance, R^d indicates that the data set is the d dimension, and sup represents the upper bound of the data set, and all elements are smaller than the value.

When the kernel function uses the Epanechnikov kernel, the integral mean square error is minimal:

$$K_E(x) = \begin{cases} \frac{1}{2} c_d^{-1} (d+2) \left(1 - \|x\|^2 \right), & \|x\| < 1 \\ 0, & \|x\| \geq 1 \end{cases} \quad (6)$$

In the formula, c_d is the d-dimensional unit sphere volume. In addition to the Epanechnikov core is more commonly used, there is a commonly used core is the Gaussian kernel.

After that, since the cores used can be differentiated, the density gradient estimate is always equal to the gradient of the kernel density estimate:

$$\begin{aligned} \nabla \hat{f}_k(x) &= \nabla \hat{f}(x) = \frac{1}{nh^{d+2}} \sum_{i=1}^n (x - x_i) K' \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \\ &= \frac{1}{nh^{d+2}} \left[g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \right] \left[\frac{\sum_{i=1}^n x_i g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)} - x \right] \end{aligned} \quad (7)$$

The density estimate at point x is denoted by G :

$$\hat{f}_G(x) = \frac{c}{nh^d} \sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \quad (8)$$

The density gradient at this time is:

$$\nabla \hat{f}_k(x) = \hat{f}_G(x) \frac{2/c}{h^2} M_{k,G}(x) \quad (9)$$

Obtained by Eq. 9, the Meanshift vector is:

$$M_{k,G}(x) = \frac{2/c}{h^2} \frac{\nabla \hat{f}_k(x)}{\hat{f}_G(x)} \quad (10)$$

The Meanshift algorithm is divided into four steps for the cube clustering process:

- (1) Randomly select a center point in the unmarked data set.
- (2) Find all the points within a certain distance from the center point. These effective points form the set M , which belong to the cluster c . At the same time, the probability that the point in the set M belongs to c is increased by one.
- (3) Calculate the vector from the center point to the effective point and add all the vectors to get the deviation vector.
- (4) The center point plus the deviation vector can get a new center point, repeating 2, 3, 4 until the deviation vector is small, satisfying the iterative condition.

Figure 1 shows the step diagram of the Meanshift algorithm, continuously solving a vector, so that the center of the circle moves to the direction of the highest density of the data set. To put it simply, it is that each iteration finds the average position of the point inside the circle as the new center position, and finally converges to the place with

the highest data density. Firstly, the size and position of a search window are initialized, and then the position of the search window is adaptively adjusted according to the result obtained in the previous frame, thereby locating the center position of the target in the current image. Meanshift can effectively solve the problem of target deformation and occlusion. It has low requirements on system resources and low time complexity. It can achieve good tracking effect in a simple background.

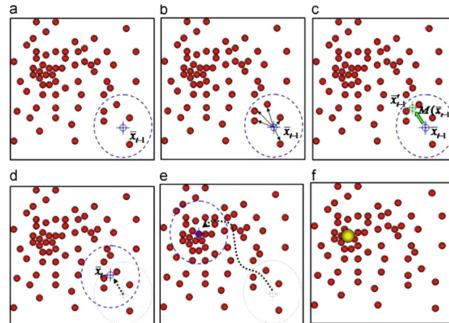


Fig. 1. The step diagram of the Meanshift algorithm

3 Dynamic Gesture Capture Positioning and Tracking Experiment

In this paper, the OpenCV library of Python3 is used as the main processing library. The initial capture and location of the hand position and the subsequent tracking experiment are performed on the video of a hand movement. The following is the experimental picture obtained by the relevant steps of the experiment and the result description of the picture.

3.1 Gesture Capture

Create a video object, specify a video directory, read the video image frame by frame, num_frame as a frame count variable, and increment the image variable by one for each frame read. The gesture capture operation is only valid when num_frame is 1.

- (1) Read the first frame image, the original picture is shown in Fig. 2:



Fig. 2. The first frame of the RGB image of the gesture video

- (2) The elliptical model is used to extract the skin color and then subjected to morphological optimization [7, 8]. After the processing, the four-connected algorithm is selected to select the maximum skin color connected domain, and the obtained result is shown in Fig. 3, which is the maximum skin color connected domain.

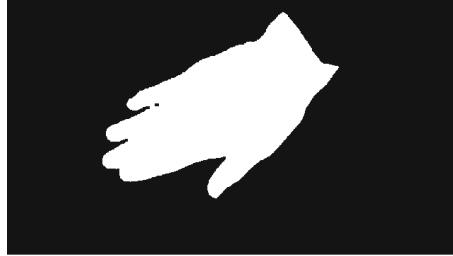


Fig. 3. Maximum skin color connected domain of the first frame image

- (3) The above two figures are subjected to image logic operation to obtain the original image of the largest skin color region of the original image. The resulting image was denoised by using bilateral filtering, and the results are shown in the Fig. 4. By comparing Fig. 2 with Fig. 4, it is found that after bilateral filtering, many pores in the skin color area of the hand are highly uniform because the edges are not obvious. In the overall gesture contour area, that is, the edge that needs to be detected in the next step, the bilateral filtering preserves the gray scale difference of the edge. As a pre-operation to extract the edge, the bilateral filtering not only achieves the effect of smoothing the image, but also highlights the edge to a large extent.



Fig. 4. Maximum skin connected domain image after bilateral filtering

- (4) The canny edge extraction is performed on Fig. 4, and the upper and lower thresholds are determined by the improved threshold algorithm for the grayscale image of Fig. 2, and the returned result is used as the edge upper and lower thresholds to extract the edges [8]. On the basis of the extracted edges, the screening condition is selected to be 1/3 of the longest contour length, and the

effective contour is selected. A rectangular box is used to fit the resulting effective edge, and the result is shown in Fig. 5.

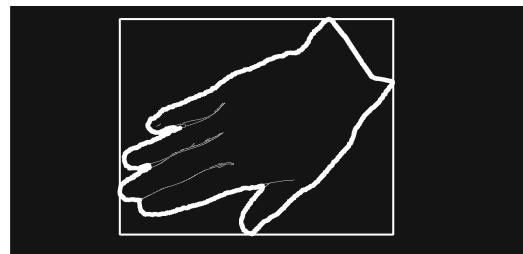


Fig. 5. Effective edge and rectangle fit image

First find the longest edge. The effective edge filtered out on the condition of 1/3 of the length is shown by the bold white line, and the other small edges are invalid. It can be found that the effective edge basically describes the shape of the hand, and the small edges are the details between the fingers, which has little effect on the description of the hand contour. And the fitted rectangle completely surrounds the effective contour, and the rectangular box can represent the position of the current gesture. The rectangular box is used as the search area on the image, and also serves as the search target of the subsequent video sequence, reflecting the color space information of the desired tracking target, providing an information basis for the subsequent tracking.

3.2 Gesture Tracking Comparison Experiments

After the previous correlation processing for the initial frame, as an initial condition, the application algorithm of MeanShift in image target tracking and Camshift are used for the subsequent video sequence, and the following is the result of processing some images. In order to clearly observe the movement change of the hand position in the processed image, selecting a picture with a certain time interval can reflect the change. Therefore, select the 1,200,400,600 frames in the video to view the gesture position tracking effect. The relevant results are shown in Figs. 6, 7, 8 and 9.

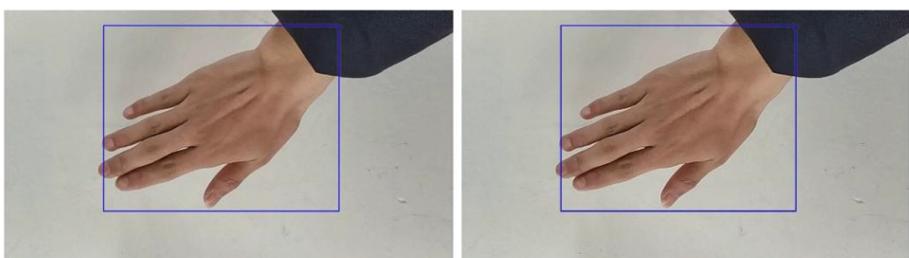


Fig. 6. Frame 1 tracking picture Meanshift (left) Camshift (right)

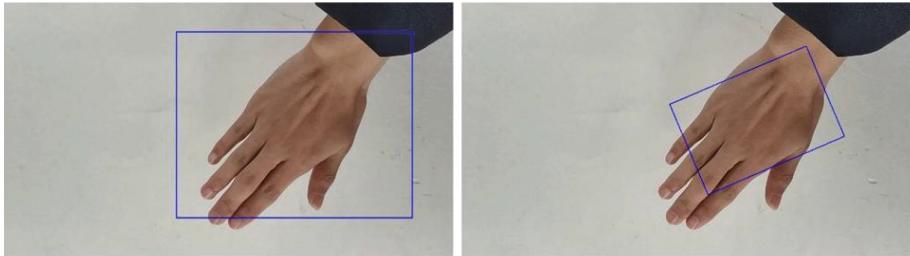


Fig. 7. Frame 200 tracking picture Meanshift (left) Camshift (right)

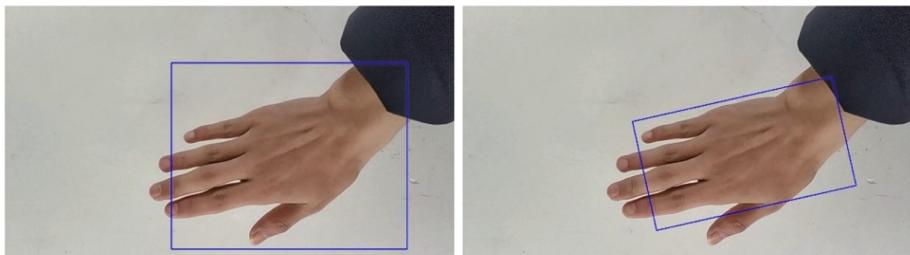


Fig. 8. Frame 400 tracking picture Meanshift (left) Camshift (right)

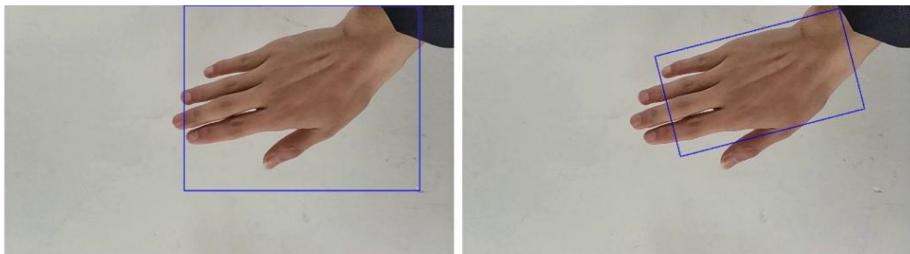


Fig. 9. Frame 600 tracking picture Meanshift (left) Camshift (right)

By observing the above picture results, when the hand is making some gestures on the wall, the MeanShift algorithm can complete the dynamic tracking of the gesture position, and the tracking effect is good. During the tracking process, the rectangular box can cover most areas of the hand. Although it is not completely surrounded by the hand, some fingers are not limited to the domain, but it is undeniable that the position of the rectangular box can be used to map the hand position and use it as the input of the control signal. However, this algorithm still has tracking defects. For example, when the hand becomes a fist, the rectangular frame will be too large, the adaptive size of the rectangular frame needs to be added, and the rectangular frame should be rotated when the hand is rotated. Therefore, the Camshift algorithm with rotation and scaling characteristics is selected for comparison experiments. The selected tracking target is

shown in Fig. 6. The two are the same, but in the subsequent tracking, the rotating rectangle can match the rotation angle of the hand. The algorithm is not optimized, and the box cannot cover the hand area completely. After the hand is rotated, as shown in Fig. 7, the size of the rotating box covers only the area of the back of the hand. Obviously, the position information obtained after tracking cannot be uploaded as control coordinates.

In the Meanshift algorithm tracking, the rectangular box represents the hand position as the control signal input method is feasible. For the fine control requirements, it is more necessary to locate the control finger in the rectangular frame, so as to obtain the precise input of the control signal. This experimental video is a simple background that achieves good tracking in a simple background with a simple background. Although it is usually a white wall when using the projection screen, it will have better adaptability. However, from the perspective of the extended performance of the algorithm, when the background is more complicated, or there are many pixel interferences similar to the target color, it will lead to tracking failure. From the degree of engineering application, the real-time performance and practical application of this algorithm can meet the requirements [8–11]. If you pursue the good performance of tracking, you can improve the tracking by adding prediction algorithms, feedback skin color detection or other algorithms to improve the above-mentioned defects, instead of determining the target from the color distribution alone.

4 Conclusion

This paper presents a gesture tracking control algorithm for human-computer interaction systems. By capturing the position of the gesture in the initial frame and then continuously tracking the initial target in the subsequent sequence of images, the position information during the tracking process is obtained and input as a control signal. Through the initial gesture capture, the rectangular frame is used as the initial detection target tracking, and the tracking algorithm is the application algorithm of MeanShift in the image target tracking. Finally, through the processing of dynamic gesture video and analysis of experimental results, it is shown that the MeanShift algorithm can better complete the dynamic tracking of gesture position.

Acknowledgments. The work is supported by the National Natural Science Foundation of China (61873176); Natural Science Foundation of Jiangsu Province, China (BK20181433); Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX18_0844; KYCX19_1927).

References

1. Lindeberg T (2012) Scale Invariant Feature Transform. Scholarpedia
2. Alamsyah D, Fanany MI (2013) Particle filter for 3D fingertips tracking from color and depth images with occlusion handling. In: International conference on advanced computer science & information systems. IEEE

3. Cheng Y (2002) Mean shift, mode seeking, and clustering. *IEEE Trans Pattern Anal Mach Intell* 17(8):790–799
4. Yu L, Li C, Fei S (2018) Any-wall touch control system with switching filter based on 3D sensor. *IEEE Sens J* 18(11):4697–4703
5. Sindhuja G, Renuka DSM (2016) Comparative analysis of mean shift in object tracking. In: Power, control, communication & computational technologies for sustainable growth. IEEE
6. Fasheng M (2005) Mean shift is a bound optimization. *IEEE Trans Pattern Anal Mach Intell* 27(3):471–474
7. Tan W, Dai G, Su H, et al (2012) Gesture segmentation based on YCb'Cr' color space ellipse fitting skin color modeling. In: Control and decision conference. IEEE, pp 1905–1908
8. Song Q, Lin G, Ma J, et al (2016) An edge-detection method based on adaptive canny algorithm and iterative segmentation threshold. In: International conference on control science and systems engineering. IEEE
9. Oka K, Sato Y, Koike H (2002) Real-time tracking of multiple fingertips and gesture recognition for augmented desk interface systems. In: Proceedings of the IEEE international conference on automatic face and gesture recognition, pp 429–434
10. Sato Y, Kobayashi Y, Koike H (2000) Fast tracking of hands and fingertips in infrared images for augmented desk interface. In: Proceedings of IEEE international conference on automatic face and gesture recognition. IEEE, pp 462–467
11. Zhang J, Chang W, Wu L (2010) Edge detection based on general grey correlation and LoG operator. In: International conference on artificial intelligence and computational intelligence. IEEE Computer Society, pp 480–483



Design and Implementation of Powered Parafoil System Based on FreeRTOS

Lin Guo, Qinglin Sun^(✉), Zengqiang Chen, Sai Chen,
and Hongchen Jia

School of Computer and Control Engineering, Nankai University,
Tianjin 300350, China
sunql@nankai.edu.cn

Abstract. In order to meet the low cost and high performance requirements of the powered parafoil system, a multi-task, multi-priority real-time embedded powered parafoil control system is designed. The hardware platform of the system consists of the microcontroller STM32F407 of the Cortex-M4 kernel and the microcontroller STM32F103 of the Cortex-M3 kernel, the software platform consists of the FreeRTOS real-time operating system and the ADRC algorithm. The system realizes the functions of real-time multi-task autonomous flight and remote control flight. In this paper, the hardware platform and software platform of the system are introduced in detail. Experiments of simulation show that using FreeRTOS to perform real-time multitask management of the powered parafoil system can make full use of CPU resources and improve the stability and real-time performance.

Keywords: STM32F407 · FreeRTOS · Real-time multitask · Powered parafoil system · Trajectory tracking control

1 Introduction

The powered parafoil system is an unmanned aircraft that uses the lift provided by the parafoil umbrella body and the propeller to generate thrust [1]. It has good controllability, and can realize the turning action by manipulating the parafoil rope to change the flight direction of the powered parafoil system [2]. The powered parafoil system can complete altitude cruising, climbing and other tasks, is widely used in precision air-drops in the military and aerospace fields [3, 4].

With the increasing use of powered parafoil system, the function is becoming more and more complex, the reliability and real-time requirements of the system are becoming higher and higher, which is directly related to the flight safety of powered parafoil system and the successful completion of flight mission.

The flight control system in [5, 6] uses the traditional embedded design method. The application is an infinite loop, the Application Programming Interface (API) function is called in the loop to complete the required operations. This large loop is called the background system. The interrupt service routine is used to handle asynchronous events of the system, which is the foreground system. This kind of sequential execution of the front-end and back-end system programs is relatively intuitive, but

because interrupts themselves spend time on the processing of breakpoints, a large number of interrupts will occupy resources, waste time, resulting in poor stability and real-time. The UAV flight control system software was developed using μ C/OS-II and Cortex-M3 microcontrollers in [7, 8]. Because of the limited internal resources of μ C/OS-II and the high consumption of ROM and RAM, the use of μ C/OS-II can ensure the real-time performance of real-time systems with high resource requirements, but it will increase the difficulty of developing flight control software.

The tasks of powered parafoil controller such as control quality calculation, ground station interaction, driving left and right steering gear and thrust motor have strict critical time. They need to be handled at appropriate time intervals. If this time is exceeded, the timing of control will be lost, which will have serious consequences.

In order to improve the real-time and stability of powered parafoil system, the FreeRTOS of embedded real-time multi-task operating system is introduced for design and management. FreeRTOS is a preemptive real-time multi-task kernel, which can achieve task scheduling and task management within a specified time, and meet the real-time and reliability requirements of powered parafoil system.

2 Design of Hardware System

The real-time multi-tasking powered parafoil system uses the ARM system as the control platform. The STM32F407IGT6 and STM32F103VCT6 chips are used as the core, and a main control unit and an auxiliary control unit are respectively designed, as shown in Fig. 1.

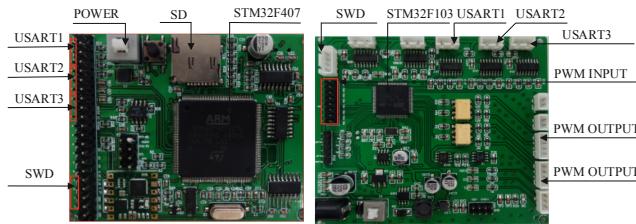


Fig. 1. Multi-tasking powered parafoil system control unit

The function of the main control unit is to receive the position information of the powered parafoil system sent by the GPS device, and calculate the control quality through the ADRC (Automatic Disturbance Rejection Control) algorithm; and transmit the control quality information to the auxiliary control unit STM32F103VCT6 through the serial port communication to drive the steering action. It can also communicate with the ground station through the full-duplex wireless serial port E62-TTL-1W to display the trajectory of the powered parafoil system in real time.

The function of the auxiliary control unit is to collect the steering angle information in real time through the photoelectric encoder, receive the control quality information of the main control unit in real time through the serial communication, output the PWM

signal to drive the left and right steering gears and the thrust motor, control the umbrella rope, realize the parafoil system turn, level flight and climb to complete trajectory tracking control; also can switch between working modes by PWM capture to achieve remote and autonomous flight.

Receive control information from the remote control unit through the Futaba receiver for remote control flight or autonomous flight selection. According to the above hardware devices, a multi-tasking powered parafoil system hardware platform is set up, as shown in Fig. 2.

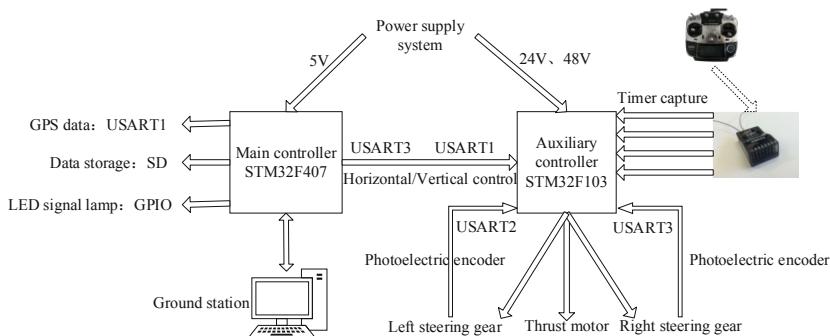


Fig. 2. Multi-tasking powered parafoil system hardware platform

3 Design of Software System

The name of FreeRTOS can be composed of two parts: Free and RTOS (Real Time Operating System), it can be seen that FreeRTOS is a free RTOS system [9].

FreeRTOS has the advantages of small kernel, strong real-time performance, high stability, fast task switching and good portability, and has functions such as task scheduling, task control, and task communication. Therefore, using the FreeRTOS operating system, complex parallel processing between multi-tasks of the powered parafoil system can be realized, ensuring the real-time and stability of tasks.

3.1 Software Design of Main Control Unit

The powered parafoil controller based on FreeRTOS divides the whole program into many tasks. Each task is independent. Even if one task has problems, it will not affect the operation of other tasks. Therefore, “task partitioning” of powered parafoil system is the key to the design of real-time multi-task powered parafoil system software. The powered parafoil system should consider two factors when dividing tasks. First, it must satisfy the “real-time” of tasks. Even in the worst case, the functions required for real-time in the system can be realized normally. In addition, the number of tasks in the system should be reasonably divided. If the number of tasks is too large, the functions implemented by each task are relatively simple. As a result, the scheduling operation of

the tasks and the communication between the tasks increase, which reduces the system operation efficiency and excessive resource overhead.

According to the above principle of task division, the powered parafoil application is divided into six tasks. The corresponding priority is set for each task, and the powered parafoil system is switched according to the priority of the task. The logic of each task in the main control unit is shown in Fig. 3. From Fig. 3, it can be seen that the functions of each task include:

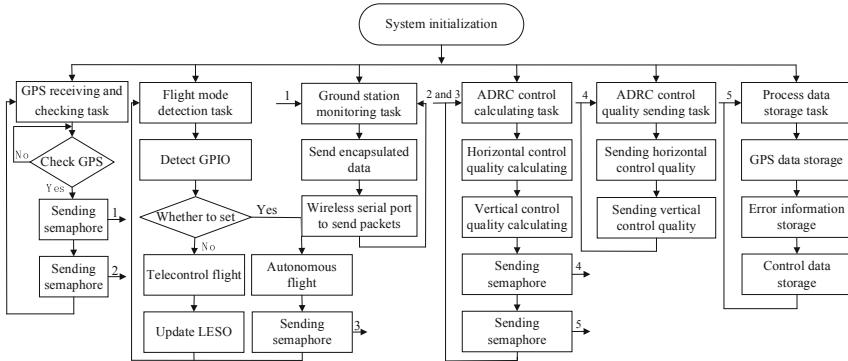


Fig. 3. Main control unit tasks logic

GPS receiving and checking task: priority is 8. GPS data is received by GPS antenna and transmitted to the main control unit through GPS card. The main control unit receives the GPS position data through the serial port interrupt, and then verifies the GPS data by the check bit. If the data is valid, the semaphore is sent to the ADRC control quality calculating task and the ground station monitoring task.

Flight mode detection task: priority is 7. The flight mode is determined by detecting the value of the STM32F407 pin GPIO2. When high level is detected, the semaphore is sent to the ADRC control quality calculating task in autonomous flight mode. When low level is detected, remote control mode is used.

Ground station monitoring task: priority is 4. When receiving the semaphore sent by GPS receiving and verifying task, the collected speed, height, longitude and latitude information will be encapsulated into data packets, and the data packets will be transmitted to the ground station monitoring system through the full duplex wireless serial port (E62-TTL-1W) connected by the main controller serial port.

ADRC control quality calculating task: priority is 3, which includes horizontal and vertical control quality calculation. When the powered parafoil system is in flight, the horizontal control quality is calculated according to the deviation of the horizontal plane between the current position and the target track (angle error and distance error), the horizontal control quality calculated by ADRC algorithm is used to adjust the course of the powered parafoil system, the specific calculation method is referred to in reference [10]. The vertical control quality is calculated according to the deviation between the set height and the current position height, and the powered parafoil height

is adjusted by using the vertical control quality calculated by ADRC algorithm, the specific calculation method is referred to reference [3]. After the calculation is completed, semaphore will be sent to ADRC control quality sending task and process data storage task.

ADRC control quality sending task: priority is 6, The priority of ADRC control quality sending task is higher than that of ADRC control quality calculating task, but it does not execute, and is always in a waiting state. When the semaphore sent by ADRC control quality calculating task is acquired, the horizontal and vertical control quality existing in the global array can be sent to the auxiliary controller through the serial port.

Process data storage task: priority is 5, FATFS is a common File Allocation Table (FAT) file system for small embedded systems, and can be easily transplanted into the powered parafoil system [11]. On this FAT file system, three files are created to store real-time longitude and latitude, horizontal and vertical control quality and errors. When the semaphore of the ADRC control quality calculating task is acquired, the data generated in the process can be written into the corresponding file.

According to the actual requirements of the main control unit of powered parafoil system, task modules with different functions are built on the basis of FreeRTOS operating system. According to FreeRTOS preemptive scheduling principle, tasks are scheduled according to the designed priority and communication mode, and the basic functions of the main control unit of the powered parafoil system are realized.

3.2 Software Design of Auxiliary Control Unit

According to the principle of task division of main control unit, the tasks in auxiliary control unit are divided into seven tasks. These tasks are managed by FreeRTOS operating system, and the functions of the system are accomplished together. The logic of each task in the auxiliary control unit is shown in Fig. 4. From Fig. 4, it can be seen that the functions of each task include:

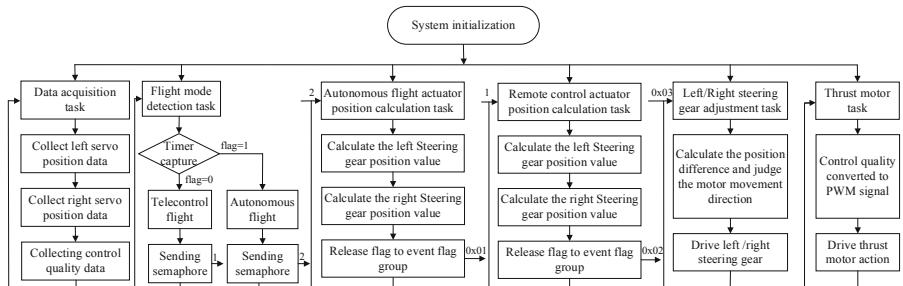


Fig. 4. Auxiliary control unit tasks logic

Data acquisition task: priority is 8. The position information of the left/right steering gear and the control information sent by the main control unit are received through serial interrupt. When the task is completing, the position data of the steering gear is analyzed from the steering gear arrays. At the same time, the data packets

received are analyzed by the protocol between the main control unit and the auxiliary control unit, the left/right steering gear and the thrust motor control quality are analyzed.

Flight mode detection task: priority is 7. The auxiliary control unit captures the PWM signal of the remote control through the input channel of the timer. According to the mode flag, it can be judged which flight mode is. If it is autonomous flight, it will send semaphore to the autonomous flight actuator position calculation task, on the contrary, it will send semaphore to the remote control actuator position calculation task.

Autonomous flight actuator position calculation task: priority is 6. When the task receives the semaphore from the flight mode detection task, it begins to execute. The zero position value of the left and right steering gear is used to subtract the position change value caused by the control quality of the steering gear, and the target position value of the steering gear is calculated. In this task, the flag 0×01 is issued to the event flag group of the left and right steering gear adjustment task.

Remote control actuator position calculation task: priority is 6, because remote control flight and autonomous flight are mutually exclusive tasks, they are set to the same priority. It can calculate the target position value of the steering gear, and issue the flag 0×02 to the event flag group of the left and right steering gear adjustment task.

Left/Right steering gear adjustment task: priority is 4/3, the initial value of event flag group is set to 0×03 , and the option is set to any flag position to run. In this task, comparing the position value of serial interrupt receiving and the position value calculated by the control quality determines whether the steering gear is a forward or reverse feeding rope.

Thrust motor task: priority is 5. This task is mainly to provide the powered of the parafoil system, and convert the control quality sent by the autonomous flight or the remote control flight into the PWM signal of the driving thrust motor to realize the height control.

The software platform of auxiliary control unit of powered parafoil system is built. The task modules are scheduled by FreeRTOS according to the set priority, which ensure that the driving mechanism can execute stably in real time.

4 Simulation Results and Analysis

In order to verify whether the powered parafoil system tasks based on the real-time multi-task operating system can operate reasonably and effectively, the semi-physical simulation test of the powered parafoil system is carried out. The hardware devices used in the semi-physical simulation are the same as the actual flight test, as shown in Fig. 5. The tracking target of the semi-physical simulation test of the powered parafoil system is a circular trajectory with an initial point of (0,0,500 m) (x, y, h, geodetic coordinate system) and a radius of 250 m.

When the virtual GPS latitude and longitude datas are transmitted to the main control unit, the main control unit switches the execution task according to the pre-emptive scheduling strategy. When the ground station monitoring task runs, the data generated in the process is encapsulated into a data packet and transmitted to the

ground station, and the ground station displays the motion track in real time after receiving the data, as shown in Fig. 6.

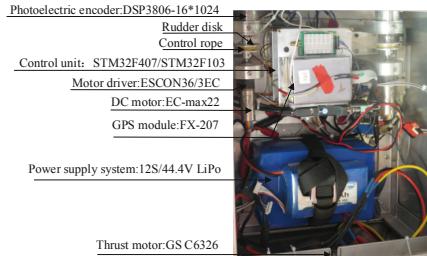


Fig. 5. Semi-physical simulation platform

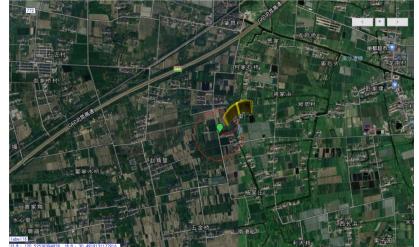


Fig. 6. Ground station displays trajectory

When the ADRC control quality calculating task receives the semaphore sent by the GPS receiving verification task and the flight mode detection task, it starts to calculate the horizontal and vertical control quality. After the ADRC control quality calculating task is completed, the process data storage task is started, the GPS data, the control quantity data, and the error data generated in the process are respectively stored in the corresponding files, as shown in Fig. 7.

| Control_u.TXT - 文本 | | Error_data | | GPS_data.TXT - 文本 | |
|----------------------|-----------|---|-------|-------------------|-------|
| 文件(F) | 编辑(E) | 文件(F) | 编辑(E) | 文件(F) | 编辑(E) |
| x-2, 61.8y75, 177, 3 | -0.025, 5 | \$GPGLL, 120003.75, 3030.0000, N, 12030.0217, E, 1, 10, 0, 9, 488.62, M, -8.80, M,, *00 | | | |
| x-2, 61.8y74, 843, 3 | 0.0174, 4 | \$GPGLL, 120000.00, 3030.0000, N, 12030.0233, E, 1, 10, 0, 9, 488.12, M, -8.80, M,, *00 | | | |
| x-0, 41.6y74, 76649, | -0.119, 4 | \$GPGLL, 120004.25, 3030.0000, N, 12030.0249, E, 1, 10, 0, 9, 487.74, M, -8.80, M,, *00 | | | |
| x-0, 41.6y74, 93399, | -0.026, 9 | \$GPGLL, 120004.50, 3030.0000, N, 12030.0265, E, 1, 10, 0, 9, 487.47, M, -8.80, M,, *00 | | | |
| x-0, 41.6y74, 9728, | -0.042, 6 | \$GPGLL, 120004.75, 3030.0000, N, 12030.0281, E, 1, 10, 0, 9, 487.23, M, -8.80, M,, *00 | | | |
| x-0, 41.6y74, 8834, | -0.041, 3 | \$GPGLL, 120005.00, 3030.0000, N, 12030.0297, E, 1, 10, 0, 9, 487.16, M, -8.80, M,, *00 | | | |

Fig. 7. Procedure data file storage

In the auxiliary control unit, when the left/right steering gear adjustment task and thrust motor task are executed, the resulting trajectory tracking and height tracking are shown in Figs. 8 and 9.

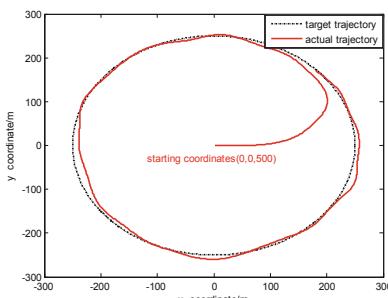


Fig. 8. Horizontal tracking trajectory

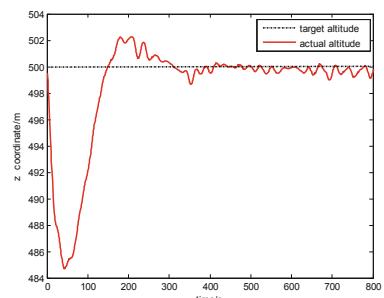


Fig. 9. Altitude tracking trajectory

After many simulation experiments, the FreeRTOS embedded multi-task powered parafoil system has stronger real-time performance and higher stability than the front and back powered parafoil system. The system can accurately and timely respond to the mode switching command sent by the remote control machine, and can realize real-time communication with the ground station, real-time control of the steering gear and the thrust motor, and the overall operation efficiency of the system is improved.

The system transplants the FreeRTOS real-time operating system to STM32F407 and STM32F103 respectively, uses the Keil μ Vision4 software to design real-time multi-tasking program modules, and builds a multi-tasking powered parafoil software control platform. The FreeRTOS real-time multitasking software structure enhances the modular design of the system, making the entire framework clear and effectively reducing the complexity of software design. Simulation experiments show that the system has stable performance and fast execution speed, which improves the real-time performance of the powered parafoil system.

Acknowledgments. This work is supported by National Natural Science Foundation of China under Grant (61273138), National Key Technology R&D Program under Grant (2015BAK06B04).

References

1. Sun Q, Chen S, Sun H, et al Trajectory tracking control of powered parafoil under complex disturbances. J Harbin Eng Univ 1–7
2. Guo L, Sun Q, Chen S, et al Design of powered parafoil system based on real-time multitask operating system. Comput Eng Appl 1–10
3. Tao J, Sun Q, Chen Q et al (2017) Linear active disturbance rejection height control for parafoil wing UAV. J Natl Defense Univ Sci Technol 39(6):103–110
4. Tan P, Sun Q, Chen Q (2017) Application of ADRC technology in power parafoil trajectory tracking control. J Zhejiang Univ 51(5):992–999
5. Xu L, Wang Z, Luo W (2017) Design and realization of the four rotor aircraft based on STM32. Autom Technol Appl 36(8):122–124, 135
6. Cao K, Ma B, Wang X (2017) Design of quadcopter aircraft control system. Appl Comput Syst 36(8):122–124, 135
7. Jiang Y (2009) Software design of small UAV flight control system based on μ C/OS-II. Nanjing University of Aeronautics and Astronautics
8. Xu J (2014) Design of embedded autonomous flight control system for multi-rotor UAV. Nanjing University of Information Engineering
9. Zuo Z, Liu J, Zhang Y (2017) FreeRTOS source code detailing and application development. Beijing University of Aeronautics and Astronautics Press
10. Jiao L (2011) Autonomous homing research based on parafoil air speculator system. Nankai University
11. Gao Z, Wang S, Xu W, et al (2017) Transplantation and application of FATFS in FreeRTOS embedded operating system. Mod Electron Technol 40(6):65–67, 70



Research on the Control of Permanent Magnet Synchronous Motor Based on High Frequency Injection and Optimized Sliding Mode Controller

Peng Bian, Bulai Wang^(✉), Qiangqiang Xu, Yuanyuan Yang,
and Na Lu

School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai, China
wruoyi@163.com

Abstract. Aiming at the shortcomings of current permanent magnet synchronous motor control, such as electromagnetic torque pulsation, flux pulsation, poor anti-interference performance and slow response speed. This paper introduces the control method of high frequency injection optimization sliding mode permanent magnet synchronous motor. Through the establishment of PMSM vector mathematical model. The design is based on high frequency injection and optimized sliding mode controller, MATLAB/Simulink was used for modeling and simulation analysis. It turned out that Compared with the traditional high frequency injection method and sliding mode controller, the vector control of permanent magnet synchronous motor based on high frequency injection and optimized synovial controller has stronger anti-interference, better reduction of chattering, which can improve the characteristics of the system more effectively and meet the needs of actual motor control.

Keywords: Permanent magnet synchronous motor · Vector control · Sliding mode controller · High frequency injection method · MATLAB/Simulink

1 Introduction

Permanent magnet synchronous motor (PMSM) has the advantages of high power density, high energy conversion efficiency, wide speed range, small size and light weight. It has been widely used in industry, civil and military fields [1].

Although the traditional PI control method can meet the speed regulation requirements of PMSM in a certain range, it has obvious defects in speed tracking accuracy and torque response, and is very vulnerable to environmental impact. Therefore, it is important to study and replace the traditional PI speed controller to achieve the desired control effect. Under this background, the sliding mode controller has been well developed. The optimized sliding mode controller used in this paper is improved under the original sliding mode controller to improve the system characteristics effectively.

Traditional sliding mode control is a control strategy in variable structure control system. The most essential characteristic of this control is the discontinuity of control, which is the switching characteristic of the system structure changing with time. In fact, it makes the system oscillate with small amplitude and high frequency along the prescribed trajectory under certain conditions. With the diversity of the working environment of motors in today's society, there will be more requirements for controlling motors. The advantages of tansigSMC-PI controller in this paper are high speed tracking accuracy, strong anti-load mutation recovery ability, good robustness to disturbance and system parameter changes, and good restraint to torque ripple [2].

2 Vector Mathematical Model of Permanent Magnet Synchronous Motor

Permanent magnet synchronous motor (PMSM) is widely used in many industrial fields. Because PMSM contains many uncertainties and is a multi-variable, strong-coupling nonlinear system, its control strategy and control method have become a research hotspot. These studies are based on reasonable mathematical models. In order to facilitate analysis, the following parameters are assumed when establishing mathematical models: (1) neglect the core saturation of the motor; (2) neglect the eddy current and hysteresis loss in the motor; (3) neglect all the spatial harmonics in the magnetic field; (4) the three-phase current in the motor is sinusoidal wave [4].

The mathematical model of PMSM in d-q coordinate system is based on the following assumptions: neglecting core saturation; ignoring eddy current and hysteresis loss; undamped winding on rotor; conductivity of permanent magnet material is 0; induction electromotive force waveform in phase winding is sinusoidal wave.

The voltage equation Formula (1)–Formula (2) and flux linkage equation formula (3)–Formula (4) of PMSM stator are as follows [5]:

$$U_d = R_s i_d + p\psi_d - w_r \psi_q \quad (1)$$

$$U_q = R_s i_q + p\psi_q - w_r \psi_d \quad (2)$$

$$\psi_q = L_q i_q \quad (3)$$

$$\psi_d = L_d i_d \quad (4)$$

By substituting the flux equation into the voltage equation, the following results can be obtained:

$$U_d = (R_s + pL_d)i_d - w_r L_q i_q \quad (5)$$

$$U_q = (R_s + pL_q)i_q - w_r L_d i_d + w_r \psi_f \quad (6)$$

Combination (5) and summation (6):

$$\begin{bmatrix} U_d \\ U_q \end{bmatrix} = \begin{bmatrix} R_s + pL_d & -w_r L_q \\ w_r L_d & R_s + pL_q \end{bmatrix} \begin{bmatrix} i_d \\ i_q \end{bmatrix} + \begin{bmatrix} 0 \\ w_r \psi_f \end{bmatrix} \quad (7)$$

Formula (7) is transformed into an equation suitable for building a model in the MATLAB/SIMULINK environment, that is.. $x = Ax + Bu$:

$$i_d = \frac{1}{L_d} U_d - \frac{R_s}{L_d} i_d + w_r \frac{L_q}{L_d} i_q \quad (8)$$

$$i_d = \frac{1}{L_q} U_q - w_r \frac{L_d}{L_q} i_d - \frac{R_s}{L_q} i_q - \frac{w_r \psi_f}{L_d} \quad (9)$$

Among: p —differential operator; R_s —Armature winding resistance; ω_r —Rotor angular velocity; ψ_f —Flux constant produced by permanent magnet; L_d —Self-inductance of D-coil; L_q —Self-inductance of q-coil.

The electromagnetic torque equation is:

$$T_e = P(\psi_d i_q - \psi_q i_d) = P[\psi_f i_q + (L_d - L_q)i_d i_q] \quad (10)$$

The mechanical motion equation of the rotor is as follows:

$$J \frac{d\omega_r}{dt} = P(Te - T_L - B \frac{\omega_r}{p}) \quad (11)$$

Among: P —Rotor pole logarithm; T_e —electromagnetic torque; T_L —Load torque; J —rotor inertia; B —Damping coefficient.

The mechanical motion equation of the rotor is as follows:

$$\Theta = \frac{d\omega_r}{dt} \quad (12)$$

3 Basic Principle of Optimizing Sliding Mode Control

Sliding mode control has control discontinuity, conventional control strategy and switching characteristics that make the system structure change with time. The principle of sliding mode control is to design the switching hyperplane of the system according to the dynamic characteristics required by the system, and to use the sliding mode controller to collect the state of the system from the hyperplane to the switching hyperplane. When the system reaches the switching hyperplane, its control function enables the system to reach the origin along the switching hyperplane [6].

The control method of PMSM in this paper is based on rotating high frequency and optimizing sliding mode speed controller. The system is characterized by small

amplitude and high frequency oscillation along the prescribed trajectory under certain conditions, which combines the optimized sliding mode speed controller with high frequency. The algorithm of tansigSMC-PI intelligent controller is as follows:

$$\dot{X} = f(\dot{x}, u, t); \quad (13)$$

$$\dot{X} - system\ state, x \in R^n; \quad (14)$$

$$u - control\ variable, u \in R^m; \quad (15)$$

$$s(x, t), s \in R^m; \quad (16)$$

$$s = cx - \omega \ among c > 0; \quad (17)$$

x-system state u-control variable c-Parameters to be designed ω -Actual speed.

In order to ensure that the three-phase PMSM drive system has good dynamic quality, the exponential reaching law method is adopted in the optimized sliding mode controller. The expression of the controller is as follows:

$$u = \frac{1}{D} [-ck + \epsilon tansig(s) + qs] \quad (18)$$

The reference current of the q-axis is calculated as follows:

$$i_q^* = \frac{1}{D} \int_0^t [-ck + \epsilon tansig(s) + qs] d\tau \quad (19)$$

$$D = \frac{2P_m \phi_f}{2J} \epsilon, \text{ q-Exponential reaching law constant;}$$

Compared with the traditional sliding mode controller, Tansig function is used to replace the traditional switching function to change the performance of the motor. The switching function of the traditional sliding mode controller is defined as follows:

$$sgn(s) = \begin{cases} 1, & s \geq 0 \\ 0, & s < 0 \end{cases} \quad (20)$$

The Tansig function used in this paper not only enhances the robustness of the sliding mode controller, but also hardly causes the phase delay of the system. The Tansig function is defined as follows:

$$th(x) = \frac{2}{(1 + e^{-2x}) - 1} \quad (21)$$

Compared with the traditional switching function, Tansig function can change the switching state more smoothly with a much smaller slope than the switching function, and achieve the same effect as the switching function, but it can obtain better motor performance. The control adopted includes integral terms, which can weaken the

chattering phenomenon and eliminate the steady-state error of the system and improve the control quality of the system [7].

4 System Simulation and Analysis Based on MATLAB

In order to verify the validity of this method, SPMSM is taken as the research object and the vector control strategy of $L_d = 0$ is adopted [8]. The speed controller is an optimized sliding mode speed controller and the current controller is a traditional PI controller. The simulation system is built by superimposing rotating high frequency injection and using MATLAB/Simulink. The system structure control block diagram is shown in Fig. 1.

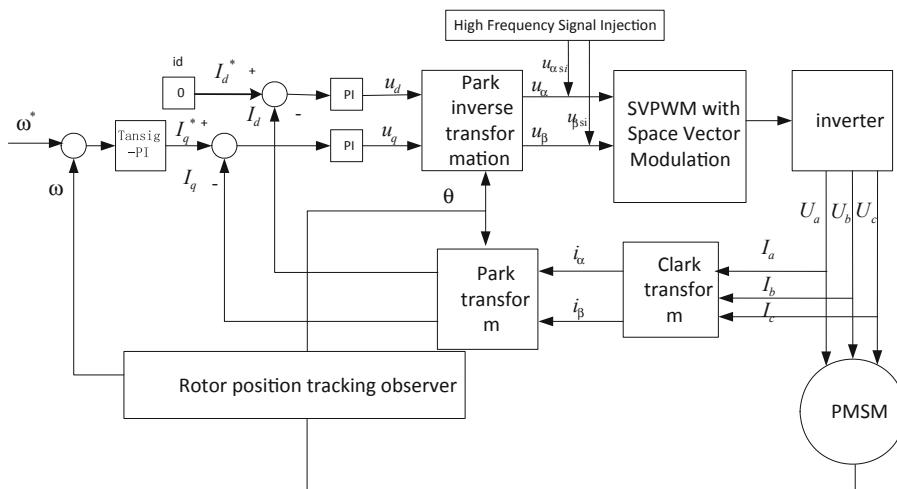


Fig. 1. System structure control block diagram

To verify the validity of the optimized sliding mode speed controller and rotating high frequency injection control method designed in this paper, according to the system control block diagram designed in this paper, the system is simulated by MATLAB/Simulink simulation software, in which the motor parameters are: pole logarithm $P_n = 2$; stator inductance $L_d = 5.2 \text{ mH}$, $L_q = 17.4 \text{ mH}$; stator resistance $R = 0.33\text{--}$; fluxf = 0.646 Wb; moment of inertia $J = 0.008 \text{ kg M}^2$; damping coefficient $B = 0.008 \text{ N M S}$. The simulation conditions are: DC voltage $U_{dc} = 311 \text{ V}$; PWM switching frequency $f_{pwm} = 5 \text{ kHz}$, using variable step size `ode45` algorithm, relative error 0.001, simulation time 0.4 s.

In addition, the amplitude of high frequency voltage signal $V_{in} = 20 \text{ V}$, frequency $f_{in} = 1000 \text{ Hz}$, i.e. $u_{\alpha phain} = V_{in} \cos(2\pi f_{in}t)$, $u_{\beta phain} = V_{in} \cos(2\pi f_{in}t + 90^\circ)$. Low-pass filter (LPF) is designed by Butterworth method with order 1 and passband frequency set to 150 Hz. The Butterworth method is also used in the design of

high-pass filter (BPF). The order is 2. The low-pass band-edge frequency is set to 987 Hz and the high-pass band-edge frequency is set to 1018 Hz.

Rotation speed $N = 100 \text{ r min}^{-1}$, initial load torque $TL = 0 \text{ N m}$, simulation time 0.4 s, when $t = 0.2 \text{ s}$ plus $TL = 2 \text{ N m}$ load, observe the motor speed, electromagnetic torque change curve. Figures 4 and 5 are the speed and electromagnetic torque curves of traditional sliding mode control PMSM under given speed and load respectively.

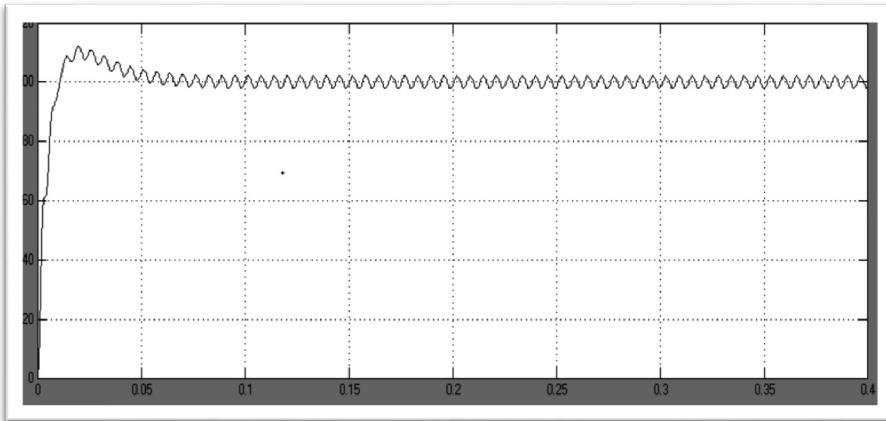


Fig. 2. Speed simulation results based on sliding mode and high frequency injection control

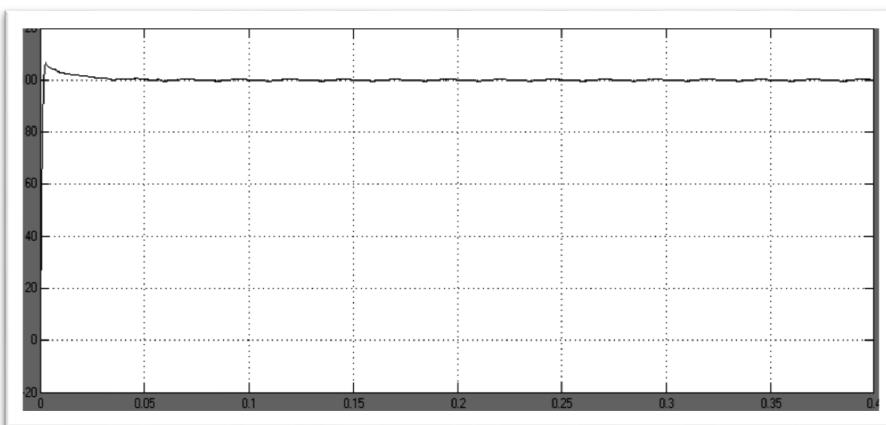


Fig. 3. Speed simulation results based on optimized sliding mode and high frequency injection control

When $t = 0.2 \text{ s}$, the load $TL = 2 \text{ N m}$ is added to observe the change curve of motor speed and electromagnetic torque. Figures 2 and 3 are the electromagnetic torque curves of traditional sliding mode control PMSM under given speed and load respectively.

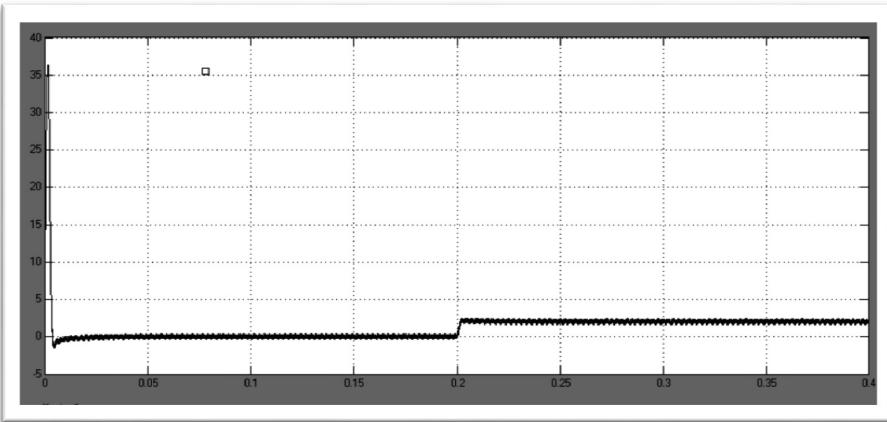


Fig. 4. Torque simulation results based on sliding mode and high frequency injection control

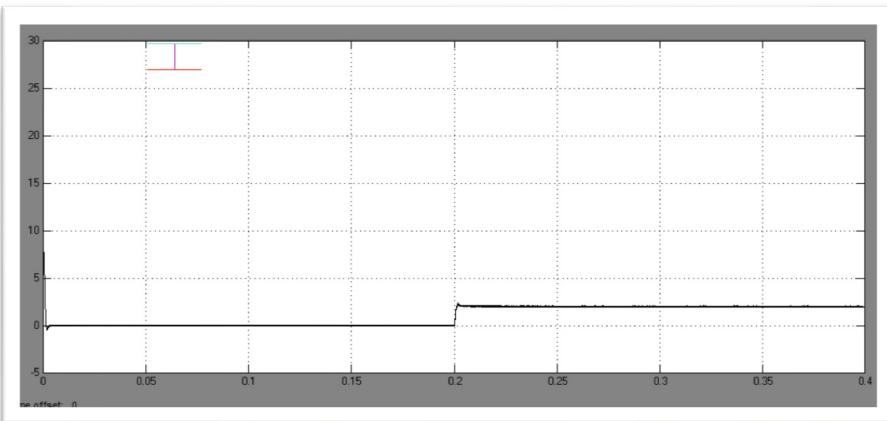


Fig. 5. Torque simulation results based on optimized sliding mode and high frequency injection control

As can be seen from the speed change curve in Fig. 2, it takes 0.02 s for PMSM based on sliding mode and high frequency injection control to achieve a given speed. The dynamic response speed is relatively slow, and there is a large overshoot before the motor speed reaches a given speed. Compared with Fig. 3, PMSM based on optimized sliding mode and high frequency injection control has a certain overshoot, but it only needs 0.01 s for the motor speed to reach the given speed. The dynamic response speed of the motor speed is obviously accelerated, and the speed is very stable, and the chattering is obviously reduced.

As can be seen from the electromagnetic torque variation curves in Figs. 4 and 5, when the load torque of the motor is 0 between 0 and 0.2 s and when the load torque is

2 N m suddenly added at 0.2 s, both control strategies can restore to a stable state, but the electromagnetic torque ripple of the conventional sliding mode and high frequency injection methods is greater, and the control methods based on optimized sliding mode and high frequency injection have better corresponding, and basically no torque ripple. It shows that the optimal sliding mode and high frequency injection control method in this paper have good torque corresponding ability.

References

1. Yan T, Liu J, Yang M, Ni T (2015) Research on control of permanent magnet synchronous motor based on high frequency injection method. *Electr Autom* 37(03):7–9
2. Wang Y, Lei J, Liu D (2016) Sensorless vector control of PMSM based on flux estimation. *J Shanghai Electric Power Inst* 32(02):175–180+187
3. Zhang H, Hu T, Wang B, Wanheng, Xu B, Zhang D (2018) An improved sliding mode variable structure position identification strategy for PMSM. *Micro-Special Motor* 46(12):42–46
4. Ji S, Wang X, Xu X, Sun W, Shi K (2019) Research on direct torque control of permanent magnet synchronous motor based on sliding mode control. *Electron Sci Technol* 45(07):1–6
5. Xie T, Gao G, Wang J (2018) Research on vector control system of PMSM based on sliding mode controller. *Motor Control Appl* 45(03):6–10
6. Qin F (2006) Sensorless control of PMSM based on the concept of power electronic system integration. Zhejiang University
7. Qin F, He Y, Jia H (2007) Sensorless operation of permanent magnet synchronous motor based on composite method of rotor position self-detection. *J Electr Eng China* 2007(03):12–17
8. Kim H, Son J, Lee J (2011) A high-speed sliding-mode observer for the sensorless speed control of a PMSM. *IEEE Trans Ind Electron* 58(9):4069–4077
9. Hu J, Qiu Y, Lu H (2016) Adaptive robust nonlinear feedback control of chaos in PMSM system with modeling uncertainty. *Appl Math Model* 40(19–20):8265–8275



Atmospheric Turbulence Rejection Control for a Geared Turbofan Engine

Yunxiao Liu, Zhongzhi Hu, and Jiqiang Wang^(✉)

Nanjing University of Aeronautics and Astronautics, Nanjing 210006, China
jiqiang.wang@nuaa.edu.cn

Abstract. The model of atmospheric turbulence is developed and added in front of a geared turbofan engine inlet to explore its influence on control performance for the low-pressure shaft speed. The PID controller and the linear active disturbance rejection controller (LADRC) are designed to compare the control effects of the two methods. The simulation results show that the LADRC, which can suppress the effects due to atmospheric turbulence, has better performance than the traditional PID controller.

Keywords: Atmospheric turbulence · LADRC · Geared turbofan engine · Disturbance rejection

1 Introduction

The aero-engine is an extremely complex nonlinear system. Due to the complexity of the controlled object and the variability of the external environment, the whole system is full of various kinds of disturbance. The disturbance can be divided into the process that cannot be modeled inside the system, the unknown system parameters, the change of the external environment, the noise of sensors or actuators, and so on [1]. According to the source of system disturbance, it can be divided into external disturbance and internal disturbance. External disturbance can be subdivided into atmospheric turbulence, lightning, EMI/HIRF, rain and snow weather, unidentified object collision, engine power extraction, the deflation of VBV, and so on. Internal disturbance can be classified into actuator components like fuel pumps subject to environmental disturbance, the deterioration of the health parameters of the turbine or compressor components and so on [2–4]. This paper studies the effects of changes in atmospheric temperature, pressure and Mach number caused by atmospheric turbulence on the closed loop control of the low-pressure shaft speed.

Research on atmospheric turbulence has been going on for a long time, and Nasstrom and Fairall [5] have developed a variety of models based on the Kolmogorov spectrum [6], but this method is difficult to create models in the time domain. Later, the von Kalman model was also used to approximate the Kolmogorov spectrum model. However, due to the fractional order of this model, there are some problems in the von Kalman model. Kopasakis sums up the pros and cons of the predecessors and develops more accurate models in the form of representative fractional orders of typical

atmospheric disturbances. This model can study the representative atmospheric turbulence model in the time domain by transfer function and disturbance frequency [7].

Domestic research on engine control for atmospheric disturbance is still in its infancy. Changes in the temperature, pressure, and Mach number of the gas in front of the inlet caused by atmospheric turbulence will lead to the changes in the gas parameters of various components within the engine, and then it will affect the control of engine speed and thrust. The basic purpose of the aero-engine control system is to ensure that the engine can operate stably and reliably under any environment and any working conditions. Therefore, the shaft speed needs to run as smoothly as possible according to the command state in the disturbance environment. The traditional PID control method has a long adjustment time and weak ability to suppress disturbance. ADRC is a control method proposed by Han Jingqing in the 1990s to solve a system with a large range of complex structural uncertainties [8, 9]. This method has been widely used in engineering. Many parameters of this method are difficult to choose with a simpler method, which limits its application in many practical systems. Later, LADRC method [10], which simplified the active disturbance rejection control method. Although the anti-disturbance performance of the LADRC is slightly lower, the method can still achieve good control effect and the settling time is short, which saves the application time of the engineering and is more favorable to the popularization and application of the LADRC [11, 12].

2 The Brief Description of Atmospheric Turbulence Model and LADRC

The LADRC of the engine low pressure shaft speed is established, and its structure is shown in Fig. 1. N_1 is the engine low pressure shaft speed, N_{1r} is the desired low-pressure shaft speed and W_f represents the fuel flow of the engine. Geared turbofan engine is a component level nonlinear model based on the T-MATS toolbox in the MATLAB [13]. The input conditions of the model are $H, Ma, \Delta T, \Delta Ma, \Delta P$. H is the height, Ma is the Mach number, and $\Delta T, \Delta Ma, \Delta P$ is the change in the temperature, Mach number and pressure calculated by the model of atmospheric turbulence.

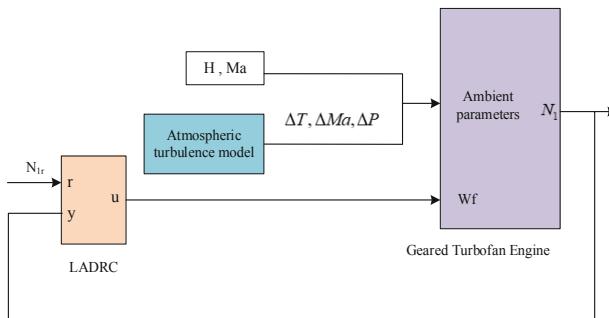


Fig. 1. Structure diagram of closed-loop control of low-pressure shaft speed

2.1 The Atmospheric Turbulence Model

The Kopasakis atmospheric turbulence model is used in this paper, which causes the changes in sound velocity, temperature, and pressure. The atmospheric disturbance model input in the time domain is simulated by the combination of unit amplitude sinusoids, the sinusoid's frequency is from 0 Hz to 200 Hz. These serve as inputs to the atmospheric disturbance transfer function. Atmospheric turbulence disturbance is shown in the Fig. 2.

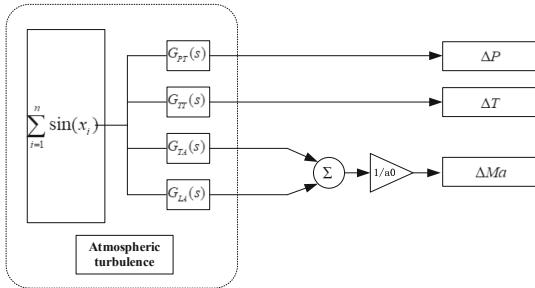


Fig. 2. The model of atmospheric turbulence

The Kolmogorov spectrum represents the spectral density of a random atmospheric turbulent flow field.

$$S_t(k) = \alpha_t \varepsilon^{2/3} k^{-5/3} \quad (1)$$

where $S_t(k)$ is the type of disturbance, α_t is a constant that corresponds to different disturbance forms, ε represents the eddy dissipation rate (m^2/s^3) and k is the wave number (cycles/m). Through a series of complex derivations, the expressions that apply to the flight envelope are finally derived.

The transfer function of atmospheric turbulence to longitudinal velocity is expressed as:

$$G_{LA}(s) = \frac{8.74(s/9.2 + 1)(s/55.0 + 1)(s/335.5 + 1)0 + 1)(s/335.5 + 1)}{(s/1.46 + 1)(s/30.1 + 1)(s/85.7 + 1)(s/1593.1 + 1)} \quad (2)$$

The transfer function of atmospheric turbulence to temperature:

$$G_{TT}(s) = \frac{41.75(s/33.0 + 1)(s/45.6 + 1)(s/602.4 + 1)}{(s/1.1 + 1)(s/25.1 + 1)(s/109.8 + 1)(s/816.3 + 1)} \quad (3)$$

Temperature change caused by atmospheric turbulence lead to the change in speed and its transfer function is:

$$G_{TA}(s) = \frac{M\gamma R}{a_0} \frac{20.88(s/33.0 + 1)(s/45.6 + 1)(s/602.4 + 1)}{(s/1.1 + 1)(s/25.1 + 1)(s/109.8 + 1)(s/816.3 + 1)} \quad (4)$$

where M is Mach number, γ is a specific heat ratio and the value is 1.4 generally, R is the gas constant and its value is 287J/(kg * k) and a_0 is local sound speed.

The transfer function of atmospheric turbulence to pressure is

$$G_{PT}(s) = \frac{37.96(s/33.0 + 1)(s/45.6 + 1)(s/602.4 + 1)}{(s/1.1 + 1)(s/25.1 + 1)(s/109.8 + 1)(s/816.3 + 1)} \quad (5)$$

The unit of the function G_{LA} , G_{TA} is m/s, G_{TT} is k and G_{PT} is kPa.

2.2 Linear Active Disturbance Rejection Controller

The linear active disturbance rejection controller is simplified from the active disturbance rejection controller. The ADRC is composed of three parts: TD (Tracking Differentiator), ESO (Extended State Observer), and LSEF (Linear State Error Feedback). The LADRC omits the TD link and changes the nonlinear function of the ESO to a linear function. The LSEF is also changed from the original nonlinear PD controller to the linear PD controller. When solving the actual problem, the existence of the nonlinear function leads to too many parameters and it is difficult to debug. The LADRC technology can not only ensure the accuracy of control and speed requirements, but also estimate the sum of the system's uncertain disturbances in real time and eliminate it [14]. The structure of LADRC is shown in the Fig. 3.

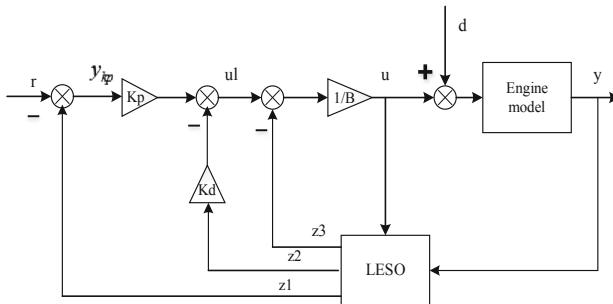


Fig. 3. Structure of LADRC

where R is the input of the system, y is the output signal, and d is the disturbance signal which is not necessary to know its specific expression.

The principle of LESO:

$$\begin{cases} \dot{z}_1 = z_2 - \beta_1(z_1 - y) \\ \dot{z}_2 = z_3 - \beta_2(z_1 - y) + Bu \\ \dot{z}_3 = -\beta_3(z_1 - y) \end{cases} \quad (6)$$

Where $\beta_1, \beta_2, \beta_3$ is tunable parameters in LESO, z_1, z_2, z_3 are estimation of the output signal y , the differential of y and unknown system disturbance. The most critical aspect of ADRC is ESO. Its main function is to estimate the total unknown factors of the whole system. These factors include various uncertainties, nonlinear factors of dynamic processes and disturbance caused by the external environment. The observer accurately observes these total disturbances and introduces them into the control loop to compensate, and then achieves the purpose of disturbance rejection. Therefore, these adjustable parameters determine the quality of the observer, and also affect the ability of anti-disturbance greatly.

The principle of the PD control:

$$\begin{aligned} u_t &= k_p(y_{kp} - z_1) - k_d z_2 \\ u &= (u_t - z_3)/B \end{aligned} \quad (7)$$

where k_p, k_d, B is the control parameter of PD, and its function is to correct the system after disturbance compensation to meet the requirements of various indicators of the system. According to the parameter configuration method provided by the literature [9, 15], the above six adjustable parameters can be conveniently selected. The parameter selection rules are as follows:

$$\beta_1 = 3w_o, \beta_2 = 3w_o^2, \beta_3 = w_o^3 \quad (8)$$

$$k_p = w_c^2, k_d = 2w_c \quad (9)$$

where w_o, w_c are the bandwidth of the observer and the controller. The LADRC only needs to adjust 3 parameters (w_o, w_c, B) to control the system.

3 Simulation Results

In order to verify the disturbance rejection effect of the LADRC on atmospheric turbulence, the influence of the atmospheric disturbance before the inlet on each parameter is simulated. The input of the atmospheric turbulence model in 40 s–45 s is shown in Fig. 4(a), This input is a superposition of 4 unit-sinusoidal functions. When the disturbance is present, there will be changes in the atmospheric temperature, pressure, and Mach number before the inlet at 40 s. As shown in the Fig. 4(b, c, d), the change in temperature is approximately $\pm 20\text{k}$, the change in pressure is $\pm 10 \text{ kPa}$ and the change in Mach number is ± 0.05 .

These disturbances are introduced into the geared turbofan engine model to compare the effects of PI control and LADRC control. In the case of 0–30 s, the input condition is $H = 0, Ma = 0$ and there is no disturbance. The commanded speed signal is from 5500 rpm to 6000 rpm to 5000 rpm as shown in Fig. 5. The parameters of the PI controller are designed according to the piece wise linearization of the component-level model, so its parameters are always changing. The LADRC parameters themselves have great robustness. After a design point is adjusted, the controller parameters of the point can be applied to a wide range of speed changes without much modification.

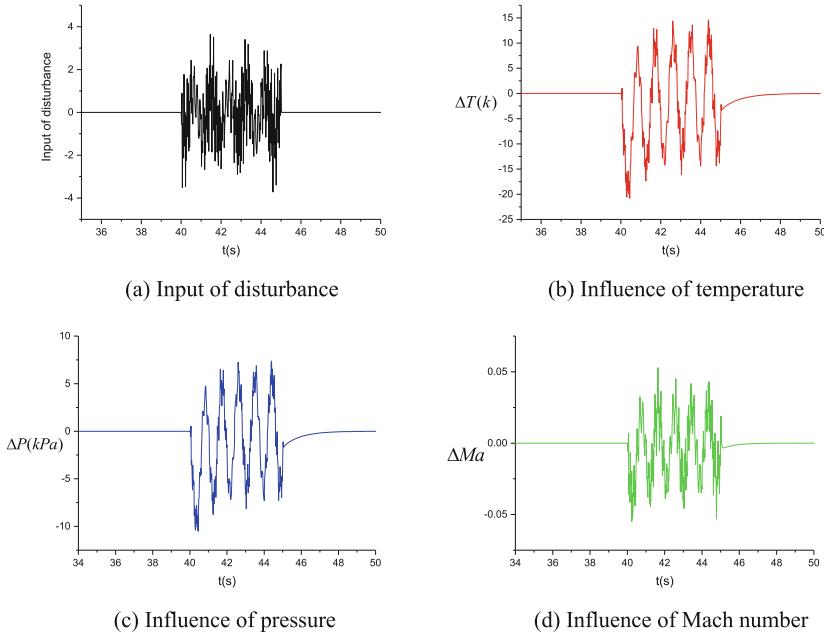


Fig. 4. Changes in various parameters caused by atmospheric turbulence

The three adjustment parameters of LADRC at this time are:

$$w_o = 14, w_c = 5.5, B = 13$$

$$k_p = 30.25, k_d = 11, \beta_1 = 42, \beta_2 = 588, \beta_3 = 2744$$

Comparing the control effects of the two controllers, it can be seen from the Fig. 5 that the settling time of LADRC is shorter, and there is basically no overshoot, while the PI controller is slightly overshoot and the settling time is longer.

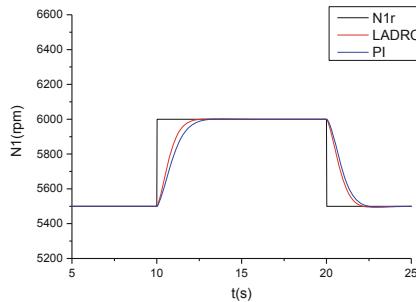
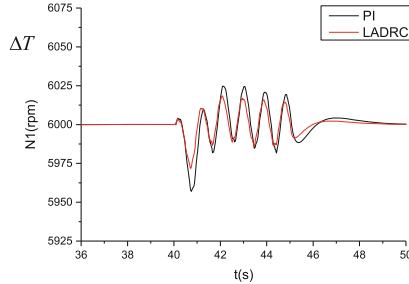
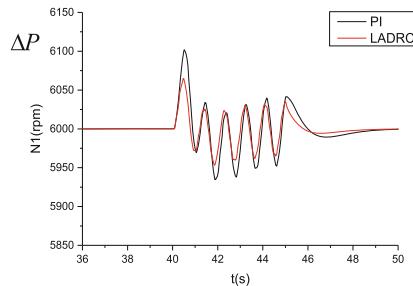


Fig. 5. Comparison of control effects between PI and LADRC without disturbance

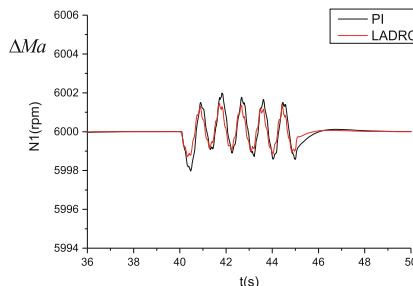
In the case of 40–45 s and $H = 0$, $Ma = 0$, testing the temperature disturbance and pressure disturbance. Testing the Mach number disturbance when the condition is $H = 0$, $Ma = 0.1$.



(a) Comparison of control effects under temperature disturbance



(b)Comparison of control effects under pressure disturbance



(c) Comparison of control effects under Mach number disturbance

Fig. 6. Comparison of the disturbance rejection between PI and LADRC

It can be seen from the Fig. 6 that due to the effect of disturbance compensation, the LADRC can quickly stabilize the speed and the fluctuation amplitude is small, and it is maintained as close as possible to the commanded speed. The control effect of low-pressure shaft speed is obviously better than PI control.

J is the index of disturbance rejection performance, which is the accumulation of the value of the low-pressure shaft speed deviating from the desired speed after the normalization in n seconds.

$$J = \sum_{t=0}^n \left| \frac{N_1 s, t - N_1 r, t}{N_1 r, t} \right| \quad (10)$$

where $N_1 s, t$ is the actual speed at time t and $N_1 r, t$ is the desired speed at time t (Table 1).

Table 1. Comparison of the disturbance rejection between PI and LADRC

| Disturbance form | J(PI) | J(LADRC) | Enhancement of J |
|------------------|--------|----------|--------------------|
| Temperature | 0.6795 | 0.4635 | 31.79% |
| Pressure | 1.6996 | 1.1428 | 32.76% |
| Mach number | 0.0437 | 0.0311 | 28.83% |

According to the calculation, the disturbance rejection ability of the LADRC is increased by about 30% compared with PI. Combined with the above simulations, LADRC can effectively suppress the atmospheric turbulence disturbance compared with PI control, which can quickly restore the low-pressure shaft speed to a stable state and avoid the unstable operation of the engine effectively.

4 Conclusions

In this paper, the effects of atmospheric turbulence on the low-pressure shaft speed of a geared turbofan engine and its control problems are studied. The results show that the changes of temperature, pressure and Mach number generated by atmospheric turbulence will cause different degrees of impact on the low-pressure shaft speed, and LADRC, compared with PID, can better suppress the effects due to the atmospheric turbulence and has better control performance, which can ensure safer engine operation. The future research work will be on modeling of other kinds of disturbance, developing corresponding LADRCs, and verifying the design in the hardware-in-the-loop test rig.

References

1. Zhang H (2016) Disturbance rejection control for several classes of nonlinear systems based on disturbance observers. Northeastern University, Mukden, pp 1–9
2. Kreiner A, Lietzau K (2002) The use of onboard real-time models for jet engine control. MTU Aero Engines, Germany
3. McNichols J, Barnes C, Wolff M et al (2013) Hardware-in-the-loop power extraction using different real-time platforms. Nat Biotechnol 21(4):422–427

4. Zhang L (2016) Anti-disturbance control for nonlinear systems based on disturbance observer. Ludong University, Shandong, pp 1–6
5. Nastrom GD, Gage KS (1985) A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *J Atmos Sci* 42(9):950–960
6. Kolmogorov AN (1991) Dissipation of energy in the locally isotropic turbulence. *Proc A* 434(1890):15–17
7. Kopsakis G, Kopsakis G (2015) Atmospheric Turbulence Modeling for Aero Vehicles: Fractional Order Fits: US8938377
8. Huang Y (2012) Xue W (2012) Active disturbance rejection control: methodology, applications and theoretical analysis. *J Syst Sci Math Sci* 32(10):1287–1307
9. Sun L, Wang R, Sun J et al (2011) Subsystem disturbance rejection control of turbo-shaft engine/helicopter based on cascade ADRC. *J Beijing Univ Aeronaut Astronaut* 37(10): 1312–1316
10. Li D, Gao Z, Chen X, et al (2011) Tuning method for second-order active disturbance rejection control. In: Proceedings of the 30th Chinese control conference, Yantai, China, 22–24 July 2011
11. Gao Z, Hu S, Jiang F (2001) A novel motion control design approach based on active disturbance rejection. In: Proceedings of the 40th IEEE conference on decision and control, Orlando, Florida, USA, December 2001
12. Chen Z, Sun M, Yang R et al (2013) On the stability of linear active disturbance rejection control. *Acta Automatica Sinica* 39(5):574–580
13. Chapman JW, Litt SL (2017) Control Design for an Advanced Geared Turbofan Engine. NASA TM-2017-219569
14. Mu C, Li S (2015) Turbofan engine high-pressure rotor speed linear since the immunity control. *Comput Simul* 32(7):365–368
15. Yuan D, Ma X, Zeng Q et al (2013) Research on frequency-band characteristics and parameters configuration of linear active disturbance rejection control for second-order systems. *Control Theory Appl* 30(12):1630–1640



Kalman-Filter-Group Based Aero-Engine Sensors Fault Diagnosis and Verification

Ying Liu, Jiqiang Wang^(✉), and Zhongzhi Hu

Nanjing University of Aeronautics and Astronautics, Nanjing 210016,
Jiangsu, China
jiqiang.wang@nuaa.edu.cn

Abstract. The normal operation of the aero-engine control system is inseparable from obtaining accurate sensor measurement signals. Many scholars have done researches on sensor fault diagnosis. In this paper, the state space equation is established based on the component level model of a turbofan engine. And the Kalman filter is designed to diagnose the hard faults of the engine sensors. Then the engine model and fault diagnosis methods are integrated into the full digital simulation platform FWorks for verification.

Keywords: Aircraft gas turbine engine · State variable model · Fault diagnosis · Kalman filter · Full digital simulation platform

1 Introduction

In the FADEC (Full Authority Digital Electronic Control) system of the engine, a large number of sensors are used for parameter measurement and state control. However, these sensors are easy to malfunction during engine operation because they are often in the harsh environment of high temperature and strong vibration. The performance of the automation system will be degraded even lead to catastrophic consequences. Therefore, when the sensor failure occurs, the control system needs to be able to diagnose the size and type of faults through the fault diagnosis system in time. Thusly, fault diagnosis of engine sensors and real-time mastery of engine health are important measures to ensure safe engine operation and extend engine life. It has engineering practical significance [1].

A study on the fault diagnosis of aero-engine sensors was carried out abroad as early as the middle of the 20th century. Wallhangen et al. used the analytic residual technology of the sensors for fault diagnosis, which improves the reliability of the aero-engine control system [2]. Naidu et al. have studied the sensor fault diagnosis method of control system based on neural network, which also showed good applicability in the verification of nonlinear system [3]. Yin et al. combined the neural network and Kalman filter to diagnose the aero-engine sensor. This was based on improved D-S evidence theory [4]. Yuan et al. proposed to use a cluster of Kalman filters for sensor fault detection and isolation, and reconstructing the adaptive model based on the information of the remaining health sensors [5]. Based on the research of the above scholars, this paper not only carries out the digital simulation of the fault diagnosis method of aero-engine sensor, but also validates it effectively on the full digital simulation platform.

2 Aero-Engine Model

The digital model of aero-engine is one of the most important foundations of control system design. By establishing a digital model to replace the real engine, the efficiency of research is greatly improved, and the cost is reduced. In this paper, the fault diagnosis of engine sensor is carried out by using the JT9D, a large bypass ratio twin-rotor turbofan engine designed and manufactured by P&W, as the research object [6].

The JT9D engine has two speed sensors, two resistive coupling sensors, five thermocouples, five pressure sensors, and two vibration sensors [7]. In this paper, eight of these sensors are selected for fault diagnosis. They are low-pressure rotor speed NL , high-pressure rotor speed NH , the high-pressure compressor outlet total temperature T_{t3} , the high-pressure compressor outlet total pressure P_{t3} , the high-pressure compressor inlet total temperature T_{t25} , the high-pressure compressor inlet total pressure P_{t25} , the low-pressure turbine outlet total temperature T_{t5} , and the fan outlet P_{t13} .

Aero-engine is a complex system with strong nonlinearity, multivariable, and time-varying. In this paper, the fault diagnosis of aero-engine sensor based on Kalman Filter Group is adopted, so the state variable model of the engine needs to be established. The nonlinear mathematical model of the engine is

$$\begin{cases} \dot{x} = f(x, u) \\ y = h(x, u) \end{cases} \quad (1)$$

where x is an n -dimensional state variable, y is an m -dimensional output variable, and u is an r -dimensional control variable. The nonlinear model is transformed into the following linear model according to the Taylor series expansion near a steady state point (x_0, u_0, y_0) :

$$\begin{cases} \Delta\dot{x} = A\Delta x + B\Delta u \\ \Delta y = C\Delta x + D\Delta u \end{cases} \quad (2)$$

In this paper, the low pressure rotor speed NL and high pressure rotor speed NH are selected as the state variables, the fuel quantity W_f is selected as the control variable, the output variables select the eight sensor signals mentioned above.

$$\begin{bmatrix} \Delta NL \\ \Delta NH \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \Delta NL \\ \Delta NH \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} \Delta W_f$$

$$\begin{bmatrix} \Delta NL \\ \Delta NH \\ \Delta T_{t3} \\ \Delta P_{t3} \\ \Delta T_{t25} \\ \Delta P_{t25} \\ \Delta T_{t5} \\ \Delta P_{t13} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \\ c_{51} & c_{52} \\ c_{61} & c_{62} \\ c_{71} & c_{72} \\ c_{81} & c_{82} \end{bmatrix} \begin{bmatrix} \Delta NL \\ \Delta NH \end{bmatrix} + \begin{bmatrix} d_{11} \\ d_{21} \\ d_{31} \\ d_{41} \\ d_{51} \\ d_{61} \\ d_{71} \\ d_{81} \end{bmatrix} \Delta W_f \quad (3)$$

In this paper, the combination of small perturbation method and fitting method is adopted to solve matrix A, B, C and D. That is, the initial value of the system matrix is obtained by using the small perturbation method first, and then correct it with the fitting method. After solving the matrix, it needs to be processed by a similar normalization method to avoid the large difference of the order of magnitude of each state variable. That is:

$$\begin{aligned} PNL &= \left(\frac{NL}{\sqrt{T_2}} \right) / \left(\frac{NL}{\sqrt{T_2}} \right)_{ds} \times 100\% \\ PT_{i3} &= \left(\frac{T_{i3}}{T_2} \right) / \left(\frac{T_{i3}}{T_2} \right)_{ds} \times 100\% \\ PP_{i3} &= \left(\frac{P_{i3}}{P_2} \right) / \left(\frac{P_{i3}}{P_2} \right)_{ds} \times 100\% \\ PW_f &= \left(\frac{W_f}{P_2 \sqrt{T_2}} \right) / \left(\frac{W_f}{P_2 \sqrt{T_2}} \right)_{ds} \times 100\% \end{aligned} \quad (4)$$

3 Introduction to Sensor Fault Diagnosis Method Based on Kalman Filter

When a sensor fails, there is a big error between the filtered value of Kalman filter and the observed value of other sensors. According to this feature, a cluster of Kalman filter is used to diagnose the fault of the sensor [8]. For the linearized state equation, the Kalman filter is designed by using the matrix A, B, C, and D.

The algorithm equation of the Kalman filter is as follows:

$$\Delta \hat{X}(0) = \Delta \bar{X}(0) = E(\Delta X(0)), P(0) = Var(\Delta X(0)) \quad (5)$$

$$\begin{aligned} \hat{X}(k+1|k) &= A(k)\hat{X}(k) + B(k)U(k) \\ \hat{Y}(k+1|k) &= C(k)\hat{X}(k+1|k) + D(k)U(k) \end{aligned} \quad (6)$$

$$P(k+1|k) = A(k)P(k)A(k)^T + Q(k) \quad (7)$$

$$K(k+1) = P(k+1|k)C(k+1)^T (C(k+1)P(k+1|k)C(k+1)^T + R(k+1))^{-1} \quad (8)$$

$$\hat{X}(k+1) = \hat{X}(k+1|k) + K(k+1)(Y(k+1) - \hat{Y}(k+1|k)) \quad (9)$$

$$P(k+1) = (I - K(k)C(k))P(k+1|k) \quad (10)$$

The difference between the input and the filtered output of the Kalman filter is called the residual, and the residual e of the i-th Kalman filter is expressed as:

$$e^i = \hat{Y}^i - Y^i \quad (i = 1, 2, \dots, m) \quad (11)$$

In the formula, \hat{Y}^i is the output of the filter and Y^i is the output of the model. The formula for calculating the sum of squared residual weights is as follows:

$$WSSR_i = (e^i)^T \cdot \left[\text{diag}(\sigma^i)^2 \right]^{-1} \cdot e^i \quad (j = 1, 2, \dots, i-1, i+1, \dots, m) \quad (12)$$

Among them, σ^j is the standard deviation of each sensor, which is the measurement of the characteristics of the sensor itself and the measurement noise. The WSSR is called the sum of residual weighted squares.

The principle of sensor fault diagnosis based on Kalman filter group is shown in Fig. 1:

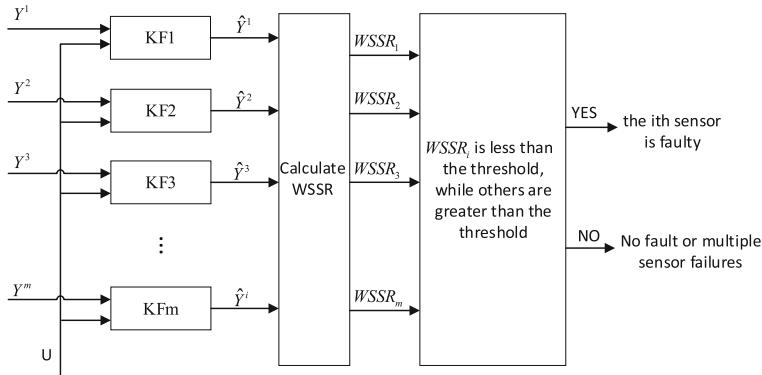


Fig. 1. Schematic diagram of a cluster of Kalman filters

The algorithm diagram is as follows:

- Step 1: Setting the initial value of the state quantity and covariance;
- Step 2: Sampling the current Y_k^i, U_k, K_k and matrix A, B, C, D;
- Step 3: Calculating \hat{Y}^i using Eqs. (5)–(10);
- Step 4: Calculating residual e of Kalman filters using Eq. (11);
- Step 5: Calculating WSSR using Eq. (12) to determine if the sensor is faulty.

In this paper, eight Kalman filters were designed for fault diagnosis of engine sensors by selecting seven outputs from the eight outputs mentioned above except the low-pressure rotor speed. The first filter uses all seven sensor signals as inputs, while the other seven filters only use six sensor signals. Therefore, if the i -th sensor fails, then only the value of the i -th WSSR is almost unchanged because it does not use the measurement information of the fault sensor, while the other Kalman filters use the information of the failed sensor, so the estimated results deviate from the actual

situation to varying degrees, and their WSSR values will change greatly, resulting in the diagnosis of fault sensors.

4 Software Platform Integration and Verification

4.1 Introduction to Platform and Integration Steps

The research and development of sensors, electronic controllers and actuators in the digital control system of aero-engines are carried out separately. It causes repeated iterative modification of the requirements of the engine control system integration. Besides, there are various unclear restrictions on the operation of the HIL platform. Therefore, the reliability of the engine model and fault diagnosis method needs to be verified before the integration of the HIL platform. This helps reduce the difficulty of error troubleshooting.

The FWorks platform (full-authority digital electronic control system collaborative design integrated platform) has two important features: one is to achieve multi-user collaborative design function, and the other is to allow users to replace control logic and components [9]. The FADEC simulation platform framework is shown in Fig. 2.

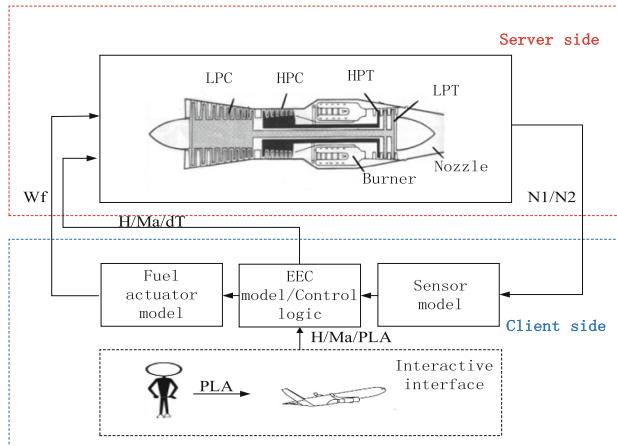


Fig. 2. Functional sketch of FWorks simulation platform

The FADEC simulation platform includes both the server side and the client side. The server side contains a model library and has functions of network connection and network communication. It provides an engine component level model for the users. The client side includes the model library of each component and baseline model of FADEC system. It also has functions of network communication, file transmission, human-computer interaction, etc. It is convenient for users to build FADEC system model and design control logic.

The requirement of the platform for the engine models is engineering files that the engine model needs to be written in C++. In this paper, the JT9D aero-engine model

based on T-MATS runs in Simulink environment, so it is necessary to transform the model format. This paper used the Code Generation tool to transform the JT9D model in Simulink into C++ code, which is eventually integrated into the server model Library of the FWorks platform. The integration process is shown in Fig. 3:

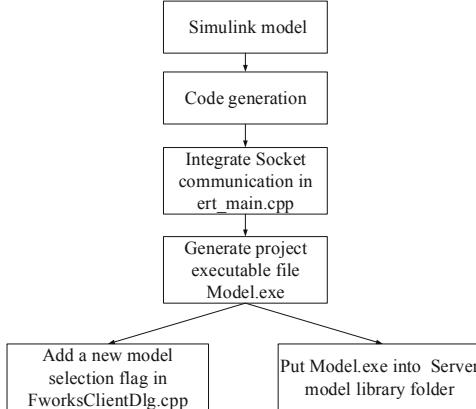


Fig. 3. JT9D model integration flow chart

4.2 Simulation Results

Taking the sensor NH as an example, it injects a 2% step fault at 20 s. The fault residual of each filter is shown in Fig. 4:

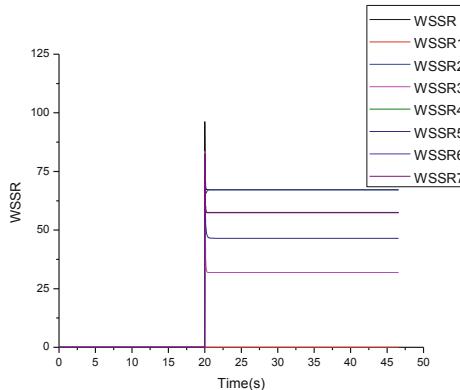


Fig. 4. Fault diagnosis results

It can be seen that the WSSR value which is not used the NH signal is almost unchanged, while the other WSSR values increase accordingly, so the sensor fault can be effectively diagnosed.

The comparison between Simulink simulation results and FWorks simulation results of NL is shown in Fig. 5:

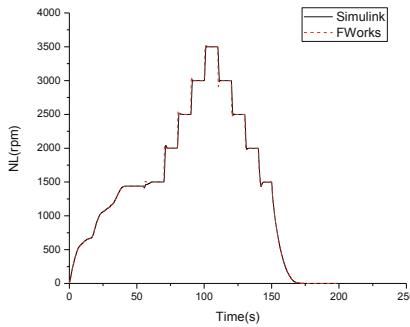


Fig. 5. Comparison of FWorks simulation and Simulink simulation

The simulation result of JT9D model before and after code conversion shows that Code Generation does not produce errors during the process of conversion. The code has good consistency with the original model.

5 Conclusion

In this paper, the linearization model of aero-engine is established. And a set of Kalman filter is designed based on this. It effectively diagnoses the sensor fault. Finally, the above model and diagnosis method are verified on the platform FWorks. The results show the effectiveness of code generation.

References

1. Kumar A, Viassolo D (2008) Model-based fault tolerant control. NASA/CR, pp 45–49
2. Wallhagen RE, Arpasi DJ (1974) Self-teaching digital-computer program for fail-operational control of a turbojet engine in a sea-level test stand. Contract 501:24
3. Naidu S, Zafiriou E, Mc Avoy TJ (1990) Use of neural networks for sensor failure detection in a control system. Control Syst Mag 10(3):49–55
4. Yin J (2001) Neural network and Kalman filter based fusion fault diagnosis for aero-engine sensor. Comput Meas Control 19(8):1936–1938
5. Yuan C (2006) Aero-engine adaptive model re-construction under sensor failure. J Aerosp Power 21:195–200
6. Hu X (2006) Handbook of small and medium-sized aero-engines in the world. Aviation Industry Press, Beijing, pp 443–445
7. Randal R, Daniel EV (2004) Towards in-flight detection and accommodation of faults in aircraft engines. In: AIAA, p 6463
8. Zhang X (2016) Design and validation of fault tolerant control system for a turbofan engine. Nanjing University of Aeronautics and Astronautics, Nanjing (2016)
9. Cao C, Shen J, et al (2017) FWorks: an integrated digital simulation platform for FADEC system. In: Proceedings of the 36th Chinese control conference, pp 10316–10321



Hypersonic Periodic Cruise Trajectory Optimization Based on Flexible Use of Pseudo-spectral Method

Haitao Gao¹, Mingwei Sun^{2(✉)}, and Zengqiang Chen²

¹ Anhui Science and Technology University, Bengbu 233000, Anhui, China

² Nankai University, Tianjin 300350, China

smw_sunmingwei@163.com

Abstract. Periodical cruise can raise the fuel efficiency of hypersonic vehicle. Therefore, it is crucial to seek optimal characteristic parameters of periodical trajectory to establish a feasible guidance law in practice. In this paper, we employ the popular, mature and efficient nonlinear optimization scheme, pseudo-spectral method (PSM), to optimize the hypersonic periodic cruise in an innovative way by utilizing the symmetric property of periodic curve. Two optimization strategies, combined and entire ones, with artificial constraints for the sake of symmetry are provided and embedded within the PSM to obtain the optimal numerical solution. Then the parameters of specific periodic functions are fitted through nonlinear square method according to these numerical solutions. The numerical results of these two proposed methods and the optimal steady-state method are compared and investigated, demonstrating the effectiveness of the proposed methods.

Keywords: Periodic cruise · Optimization · Hypersonic vehicle · Gauss pseudo-spectral method · Optimal control

1 Introduction

In the past two decades, the investigation of hypersonic vehicles is active all over the world for their long flight ranges, usually more than several thousand kilometers, in short periods. In most cases, hypersonic vehicles mainly operate at optimal steady-states (OSS) with constant altitudes and velocities for the guidance and control design convenience. Note that the cruise altitude and velocity are determined according to the principle of maximization of fuel efficiency. In other words, the flight range with respect to certain fuel consumption should be as long as possible. Further enhancing this efficiency is crucial to make room for more useful payload and to reduce the weight of the launch vehicle, which is an expected economical pattern. Therefore, a more favorable flight mode rather than the traditional steady-state cruise should be sought to raise the efficiency-cost of fuel consumption. This kind of research is important especially for the inter-continental hypersonic vehicles.

It is clear that the conventional steady-state cruise is the simplest form in the sense of optimality and this 2 degree-of-freedom (DOF) trajectory in terms of altitude and

velocity achieves a realizable guidance law at the loss of a certain range. The mathematically optimal trajectory in theory is an infinite DOF curve, which cannot be implemented in practice. Increasing the DOF of the cruise trajectory is the sole way to improve fuel efficiency, however should be performed within the achievable framework. In addition to the steady-state cruise, periodic cruise is a possible alternative to raise the fuel efficiency, which is an application of optimal periodic control (OPC) in the hypersonic cruise flights. Periodic control is a repetitive combination of the same control action in each cycle. Bryson [1] made use of energy state concept to solve minimum energy/maximum flight range problem. Schultz [2] derived the most economical velocity ratio of climbing and gliding stages by using the maximum principle. Barmen [3] also calculated the ratio mentioned above, and he found that additional reduction in fuel consumption can be achieved when the thrust is variable. Speyer [4] proved that steady-state cruise is non-optimal, laying the foundation for further OPC investigation. Chuang [5] utilized the shooting method to obtain persistent-oscillation type trajectory, which could reduce the fuel consumption by about 8% compared with the steady-state cruise. These results were quite attractive to enhance the economic performance of hypersonic vehicles, however subsequent developments about cruise optimization were rarely found in literature. In literature, the trajectory optimization is generally categorized as a constrained nonlinear optimization problem, which can be conventionally solved by using gradient-based methods, including direct and indirect schemes [6]. The indirect philosophy analytically derives the necessary condition and then obtains the optimal solution by solving these equations. Although it is mathematically elegant, the symbolical or even numerical indirect approach can rarely yield desirable solution for complicated problems in reality due to the curse of dimensionality. In recent years, the direct methods, especially pseudo-spectral methods (PSM), are becoming popular with the developments of computer and the maturity of the corresponding optimization software, GPOPS (Gauss pseudo-spectral optimization software), which is based on some effective nonlinear programming methods, such as SQP (sequential quadratic programming) [7]. The PSM, as an efficient direct optimization approach, has attracted a lot of attention. Although PSM has been widely used in trajectory optimization problems because of its maturity, there is no reports of its application in the periodic cruise optimization to our best knowledge. The major difficulty lies in the fact that the form constraints of periodic trajectory are beyond the framework of PSM, which is established based on the polynomial fitting in terms of Gauss nodes to deal with infinite-DOF curve optimization problem. Therefore, it is easy to implement a parameterized trajectory through proper guidance however the optimal characteristic parameters are difficult to obtain. At present, it is hard to find other means rather than PSM to solve this problem. Optimizing periodic cruise trajectory using PSM can provide a general, convenient, and effective methodology for practitioners in guidance design in light of the current popularity of GPOPS. Since the periodic cruise cannot be embedded into the GPOPS software directly, the problem is how to achieve this objective in a flexible manner.

In this paper, the hp-adaptive Gauss pseudo-spectral method (GPM) is used to obtain the periodic trajectory of HL-20 hypersonic aircraft by using two different philosophies, combined optimization strategy and entire optimization strategy, respectively. The respective constraints for both methods are presented with concrete

mechanical interpretation. Then we embed these selective constraints into GPOPS and obtain the optimal values at given Gauss nodes. To achieve periodic curve, these numerical results are fitted by using proper periodic functions with respect to the range. Mathematical simulations are conducted to verify the proposed approaches. The comparisons between the combined optimization and entire optimization methods are performed to provide insightful guidelines.

2 Mathematical Model and Problem Formulation

In this paper, a modified HL-20 hypersonic vehicle model is employed [8] to serve as a benchmark to interpret the proposed optimization method. At the trajectory optimization and guidance design stage, we assume that the attitude control can be realized ideally; in other words, there is no control error and the hypersonic vehicle can be regarded as a mass point. For simplicity, the spherical model of the Earth is neglected without affecting the essence of the optimization algorithm. The dynamic model for this mass point can be described as

$$\begin{cases} \frac{dh}{dt} = Ma \sin \gamma \\ \frac{dM}{dt} = \frac{T \cos \alpha - D - mg \sin \gamma}{ma} \\ \frac{d\gamma}{dt} = \frac{T \sin \alpha + L}{mMa} + \cos \gamma \left(\frac{Ma}{R_e + h} - \frac{g}{Ma} \right) \\ \frac{dm}{dt} = -\frac{T}{g I_{sp}} \\ \frac{dr}{dt} = Ma R_e / (R_e + h) \cos \gamma \end{cases} \quad (1)$$

where h is the altitude (m); M is the Mach number; γ is the elevation angle (rad); m is the mass of the flight vehicle (kg); r is the range (m); a is the normal sound speed and assumes the constant of 340 m/s for convenience; α is the angle of attack (rad); T is the thrust (N); L is the lift (N); $g \approx 9.81 \text{ m/s}^2$ is the gravitational constant; $R_e = 6378 \text{ km}$ is the radius of the Earth; I_{sp} is the specific impulse of the booster. The detailed information of this model can refer to [8].

According to the definition of periodic cruise, the period refers to the flight range rather than the flight time. Therefore, we can reformulate (1) as

$$\begin{cases} h' = \frac{dh}{dr} = \tan \gamma \left(1 + \frac{h}{R_e} \right) \\ M' = \frac{dM}{dr} = \frac{T \cos \alpha - D - mg \sin \gamma}{Ma^2 m \cos \gamma} \left(1 + \frac{h}{R_e} \right) \\ \gamma' = \frac{d\gamma}{dr} = \left(\frac{L + T \sin \alpha - mg \cos \gamma}{mMa^2 \cos \gamma} + \frac{1}{R_e + h} \right) \left(1 + \frac{h}{R_e} \right) \\ m' = \frac{dm}{dr} = -\frac{T}{g I_{sp} Ma \cos \gamma} \left(1 + \frac{h}{R_e} \right) \end{cases} \quad (2)$$

which is the foundation for later optimization.

In this paper, an optimal periodic cruise trajectory at a certain initial altitude needs to be found to minimize the fuel consumption per unit of flight range. Therefore, the cost function can be defined as

$$J = (m_i - m_t)/r_c \quad (3)$$

where m_i and m_t are the initial and terminal masses of one cycle of r_c .

3 Optimization Strategies for Periodic Cruise

Due to the mathematical form restriction of periodic cruise trajectory, GPM cannot be directly put into application. To solve this problem, we present several process and terminal constraints accompanying the expected periodic curve, or the necessary conditions, to shape the resultant optimal solution in a compulsive manner. The high symmetry properties of periodic curve are utilized here.

In this section, two optimization approaches, combined optimization and entire optimization, are provided, respectively. The primary difference between them is that the constraints are given in four parts according to the different stages of the periodic curve in the first one while only the terminal constraints are offered in the latter one. Thereafter, the GPM method is employed similarly for both schemes. Therefore, we only present the respective constraints below.

3.1 Constraints in Combined Optimization

In this strategy, a single cycle of the periodic trajectory is evenly divided into 4 parts with respect to its phase. Take the cosine curve as an example, we divide the curve into 4 segments, which are $[0, \pi/4]$, $[\pi/4, \pi/2]$, $[\pi/2, 3\pi/2]$, $[3\pi/2, 2\pi]$.

The symmetry properties should be fully guaranteed in the constraints in order to facilitate the fitting at the following stage. The state constraints in each stage can be refined further within the feasible range to shape the curve in an effective manner. We take the altitude as an example to interpret the philosophy. If h_0 is the initial altitude of each cycle or the first stage and h_{\min} is the possible lowest altitude within each cycle, h_0 is also the highest altitude within this cycle and can also be defined as h_{\max} . Then in the entire cycle, the constraints on the altitude can be roughly confined within $[h_{\min}, h_{\max}]$. To facilitate the periodic optimization process, the limits on the altitude in the first stage can be tightly formulated as $[(h_{\min} + h_{\max})/2, h_{\max}]$. In the second and third stages, the altitude constraints are still conservatively set as $[h_{\min}, h_{\max}]$ because of their far distances from the initial point. In the fourth stage, the constraints can also be tightly set as $[(h_{\min} + h_{\max})/2, h_{\max}]$ like the first stage due to the symmetry of the cycle. For the elevation angle, the state constraints are identical for the first and second stages while those of the third and fourth stages are same, because the maximal absolute value of γ can only emerge at the intersection of the first and second stages or the third and fourth stages. The constraints on the states of Mach number and mass are confined on the entire feasible interval of their minimal and maximal values. The path constraints here are selected as the change rates of states because many of them have explicit ranges of

values during each phase for a periodic cycle. For example, the change rates of the altitude and the elevation angle, h' and γ' , can be confined as $[-h_r, 0]$ and $[-\gamma_r, 0]$ for the first phase, respectively, where h_r and γ_r are their maximal feasible change rate in terms of range. According to the symmetry of the cycle, the combinations of the change rates of the altitude and the elevation angle for the second to the fourth phases can be summarized as $\{[-h_r, 0], [0, \gamma_r]\}$, $\{[0, h_r], [0, \gamma_r]\}$, and $\{[0, h_r], [-\gamma_r, 0]\}$, respectively. It should be noted that a set of tight constraints is helpful to accurately seek the required optimal periodic solution as long as there are feasible solutions. On the other side, the ranges of change rates of the Mach number and the mass are relatively rough, which can be described as $[-M_r, M_r]$ and $[-m_r, 0]$, respectively. The boundary constraints are mainly used to ensure continuity of consecutive phases, that is to say, the terminal states of each phase should be equal to the corresponding initial states of next phase. Without loss of generality, the control constraints are described with regard to the limits for the corresponding control variables, that is $\alpha_{\min} \leq \alpha \leq \alpha_{\max}$ and $0 \leq s \leq 1$ for the two control variables in this case.

3.2 Constraints in Entire Optimization

In the abovementioned combined optimization strategy, multiple constraints are added to compulsively shape the cycle like a cosine periodic curve. It can be seen that too many constraints are needed to shape the curve, which imposes great computational complexity for an optimizer. Here we attempt to propose another entire optimization strategy to simplify the optimization problem. For the entire optimization, no artificial segmentation is needed at all and only the terminal constraints that ensure a continuous cycle are considered. In a word, we only specify that the initial and terminal boundary states are identical with additional reasonable path constraints. The path constraints can also employ the ones used in the combined optimization method or roughly the change rates limit for the states. Since there is no phase segmentation, the state and path constraints can be provided in a unified form. The corresponding constraints are listed in Table 1, where x_l ($x = h, M, \gamma, m, r; l = i, t$) are the initial or terminal value of state x of each cycle. Here, the constraints are quite compact.

Table 1. Constraints of the entire optimization.

| State constraints | Boundary constraints | Path constraints |
|--|---------------------------|-------------------------------------|
| | $h_i = h_t = h_0$ | |
| $h \in [h_{\min}, h_{\max}]$ | $M_i = M_t = M_0$ | $h' \in [-h_r, h_r]$ |
| $\gamma \in [-\gamma_{\max}, \gamma_{\max}]$ | $\gamma_i = \gamma_t = 0$ | $\gamma' \in [-\gamma_r, \gamma_r]$ |
| $M \in [M_{\min}, M_{\max}]$ | $m_i = m_0$ | $M' \in [-M_r, M_r]$ |
| $m \in [m_{\min}, m_0]$ | $r_i = r_0$ | $m' \in [-m_r, 0]$ |
| | $r_t = r_f/4$ | |

3.3 Periodic Fitting of the Optimal Solution

No matter the combined or the entire optimization approach is utilized, the optimal solution by using GPOPS is just a set of state and control variables at finite Gauss nodes. This solution is based on the Lagrange polynomial fitting scheme, which has high-order form and is difficult to implement in the onboard computer. In addition, this polynomial form is unable to reveal the periodic characteristics of the trajectory we hope for. Therefore, we adopt the periodic fitting of optimal solution by using proper order cyclic function. By doing so, we can achieve an explicit parameterized guidance law for practitioners. Because the periodic functions are nonlinear with respect to the characteristic parameters, we employ nonlinear least square method to identify them.

4 Numerical Examples and Analysis

In this section, numerical examples are provided to illustrate the effectiveness of the proposed methods. We attempt to investigate the following aspects, the optimality, the accuracy, and the computational complexity. Optimality is the scope of performance improvement comparing with the traditional optimal steady-state cruise. Here the performance is just the average fuel consumption. The comparing results can explain the practical significance of the periodic cruise and the proposed method. The accuracy investigation can compare the two proposed schemes in terms of the effectiveness of periodic fitting because of the different setup of constraints. The computational complexity offers an insightful view for the efficiency of the two proposed methods.

4.1 Optimal Velocities of Steady-State Cruise at Different Altitudes

At first, the optimal velocities of steady-state cruise at different altitudes should be obtained to provide a foundation for further periodic optimization and comparison. The optimal velocity can be obtained through trajectory optimization by letting the altitude rate, the Mach number rate and the elevation angle to be zeros. In this method, we transform all dynamic equations into algebraic equations in terms of h , M , α and s in the steady situation. Thereafter, we traverse a range of M to find the optimal solution through nonlinear least square method. Finally, the real optimal M at a certain altitude can be obtained by sorting these sampling results in terms of fuel consumption, which corresponds to the optimal M . The optimization results for OSS cruise with altitude range [42 km, 45 km] are shown in Table 2. It is clear that the optimal cruise velocity is larger at higher altitude with more fuel consumption, being compatible with our empirical knowledge.

Table 2. Steady-state cruise optimization results.

| Altitude (km) | Mach number | Fuel consumption (kg/km) |
|---------------|-------------|--------------------------|
| 45 | 14.536 | 1.566 |
| 44 | 13.980 | 1.564 |
| 43 | 13.413 | 1.569 |
| 42 | 12.830 | 1.580 |

4.2 Combined Periodic Cruise Trajectory Optimization

Firstly, the combined periodic optimization is carried out utilizing the initial altitudes and Mach numbers as shown in Table 2, and the corresponding results are shown in Table 3. It can be seen that periodic cruise can improve fuel efficiency comparing with the corresponding optimal steady-state cruise, and this improvement is more evident at higher altitudes. The optimization times are all about 10 s.

Table 3. Combined optimization results.

| Altitude (km) | Fuel consumption per kilometer (kg/km) | Cycle range (km) | Improvement in fuel consumption (%) | Optimization time consumption (s) |
|---------------|--|------------------|-------------------------------------|-----------------------------------|
| 45 | 1.566 | 980 | 2.72 | 7.29 |
| 44 | 1.564 | 878 | 2.14 | 7.45 |
| 43 | 1.569 | 961 | 2.20 | 11.39 |
| 42 | 1.580 | 980 | 1.79 | 7.69 |

Next, a periodic fitting is implemented to the optimal solution; otherwise, this numerical high-order polynomial obtained by using GPM is difficult to be realized in practice. We take the 43 km optimization results as examples.

It should be noted that the values of state and control at the discrete Gauss time nodes are variables to be optimized in the GPOPS and they are optimized as a NLP problem with the help of gradient based SQP method. These optimal discrete points at 43 km are shown in Fig. 1. It is obvious that the control variables, α and s , have some outlier-like points. To achieve a periodic cruise guidance law, we attempt to fit the discrete optimal altitudes in a periodic form. It is well known that trigonometric functions are typical periodic functions. Therefore, we hope to fit these points by using some types of cosine functions. Here we consider the second-order harmonic form. Since the harmonic functions are nonlinear in terms of its characteristic parameters, we employ nonlinear least square method to obtain these parameters. For the 43 km case, the fitting formula is

$$h = 3785 \cdot \cos(0.006945 \cdot r_{\text{km}}) - 66.65 \cdot \cos(2 \cdot 0.006945 \cdot r_{\text{km}}) + 39250 \quad (4)$$

The original altitude data and the fitting result are illustrated in Fig. 2. It is clear that the fitting matches the original optimization result perfectly.

4.3 Entire Periodic Cruise Trajectory Optimization

For the entire optimization scheme, the initial values of the states and the constraints are identical to those of the combined optimization method. The results of entire optimization are presented in Table 4. Despite similar optimization results, the optimization time consumption is all about 3 s.

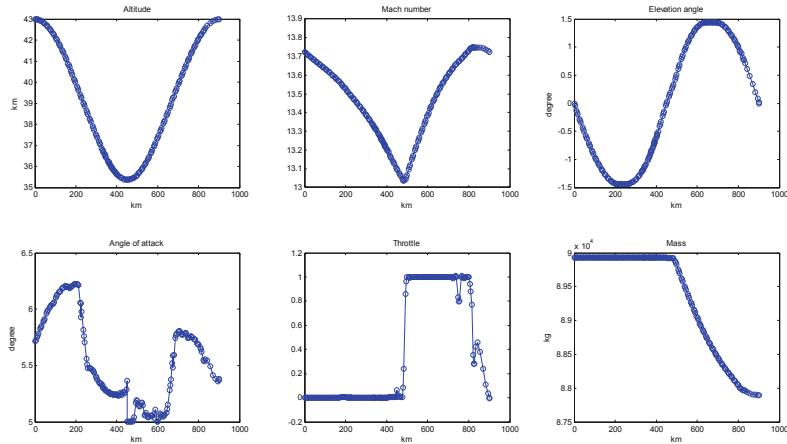


Fig. 1. Combined optimization results of one cycle at the altitude of 43 km.

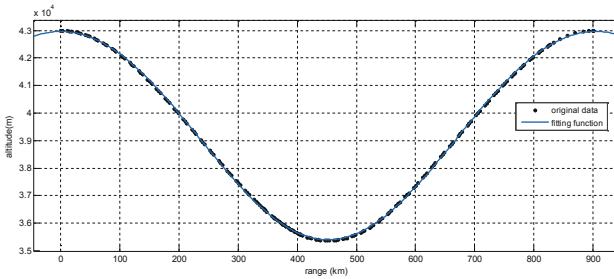


Fig. 2. Second-order harmonic fitting of 43 km combined optimized trajectory.

Table 4. Entire optimization results.

| Altitude (km) | Fuel consumption per kilometer (kg/km) | Cycle range (km) | Improvement in fuel consumption (%) | Optimization time consumption (s) |
|---------------|--|------------------|-------------------------------------|-----------------------------------|
| 45 | 1.5203 | 980 | 2.94 | 3.50 |
| 44 | 1.5246 | 878 | 2.54 | 3.75 |
| 43 | 1.5356 | 961 | 2.13 | 3.38 |
| 42 | 1.5530 | 843 | 1.72 | 3.30 |

For this entire optimization philosophy, the 43 km results are shown in Fig. 3. Comparing these results with those of Fig. 1, it can be found that the resultant control variables or the fast variables, the angle of attack and throttle, are smoother than those of the combined optimization strategy. Then we implement fitting for the optimal discrete altitudes. The result of the second-order harmonic fitting is

$$h = 4087 \cdot \cos(0.007459 \cdot r_{\text{km}}) - 306.2 \cdot \cos(2 \cdot 0.007459 \cdot r_{\text{km}}) + 39300 \quad (5)$$

which is shown in Fig. 4.

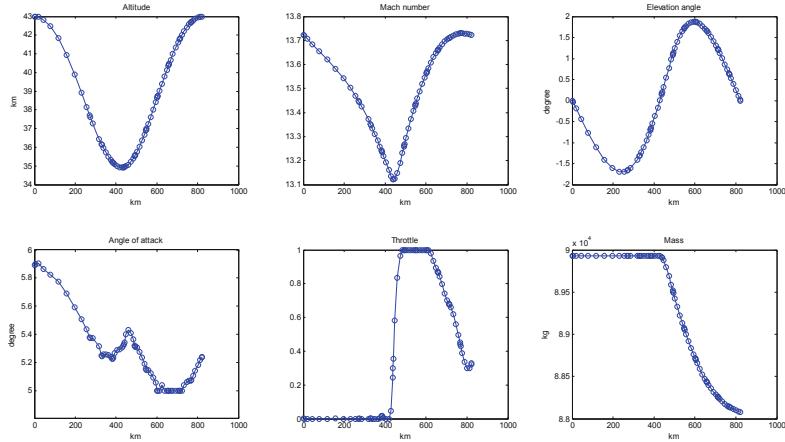


Fig. 3. Entire optimization results of one cycle at the altitude of 43 km.

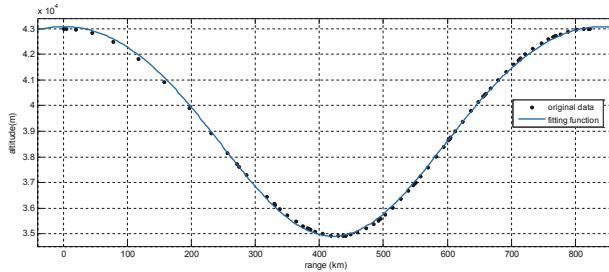


Fig. 4. Second-order harmonic fitting of 43 km entire optimized trajectory.

4.4 Comparison of Two Optimization Strategies

Here, we compare the two proposed periodic cruise optimization methods, the combined and the entire approaches, with regard to the optimality, accuracy, and computational complexity.

Firstly, either the entire optimization or the combined optimization we utilize, the achievable results are better than the results of the optimal steady-state cruise in terms of fuel consumption. In addition, the improvement in fuel consumption is more remarkable with higher initial altitude according to Tables 3 and 4.

Secondly, the optimal discrete altitudes obtained through both methods can be well fitted by utilizing second-order harmonic functions. Moreover, the results of the combined optimization are equipped with high fitting accuracy. However, it is a

trade-off when using the combined optimization approach, wherein outliers in the control variables emerge. However, these outliers have negligible effects on the altitude and range.

Thirdly, the optimization computational complexity is also different for these two strategies. The average optimization time of the combined optimization is 11.67 s, while that time of the entire optimization is only 3.30 s. This difference results from the different numbers of optimization phases and constraints.

5 Conclusion

In this section, numerical examples are provided to illustrate the effectiveness of the proposed methods. We attempt to investigate the following aspects, the optimality, the accuracy, and the computational complexity. Optimality is the scope of performance improvement comparing with the traditional optimal steady-state cruise. Here the performance is just the average fuel consumption. The comparing results can explain the practical significance of the periodic cruise and the proposed method. The accuracy investigation can compare the two proposed schemes in terms of the effectiveness of periodic fitting because of the different setup of constraints. The computational complexity offers an insightful view for the efficiency of the two proposed methods.

Acknowledgments. This work was supported by National Natural Science Foundation of China (Nos. 61573197, 61573199, 61603051 and 61603056), and the University Youth Elite Support Plan Foundation of Anhui Province (No. gxyq 2017045).

References

1. Bryson AE, Desai MN, Hoffman WC (1969) Energy-state approximation in performance optimization of supersonic aircraft. *J Aircr* 6(6):481–488
2. Schultz RL, Zagalsky NR (1972) Aircraft performance optimization. *J Aircr* 9(1):108–114
3. Barman JF, Erzberger H (1976) Fixed-range optimum trajectories for short-haul aircraft. *J Aircr* 13(5):748–754
4. Speyer JL (1976) Nonoptimality of the steady-state cruise for aircraft. *AIAA J* 14(11):31–38
5. Chuang C-H, Speyer JL (1997) Periodic optimal hypersonic scramjet cruise with constraints. *J Spacecraft Rockets* 34(2):165–171
6. Betts JT (1998) Survey of numerical methods for trajectory optimization. *J Guid Control Dyn* 21(2):165–171
7. Rao AV, Benson DA, Darby CL, Patterson MA, et al (1998) Algorithm 902: GPOPS, A MATLAB software for solving multiple-phase optimal control problems using the Gauss pseudospectral method. *ACM Trans Math Softw* 37(2):Article 22, 39 p
8. Chuang C-H, Speyer JL (1987) Periodic optimal hypersonic scramjet cruise. *Optimal Control Appl Methods* 8(3):231–242



Research on Global Stability Control Strategy of Under-Actuated Crane System Based on Equivalent Input Interference

Xiangsheng Liu^(✉), Zhengxin Zhou, Yunxia Jiang, Yuanyuan Yang,
and Qingqing Chen

Sanda University, Shanghai, China
lxssit@163.com

Abstract. Arm to anti-swing control of underactuated crane system, a new control strategy is proposed to ensure the global stability by changing the nonlinear part into equivalent input interference. Based on homeomorphic coordinate transformation, the crane system can be modeled by the combination of linear and nonlinear model, and put forward its global stability conditions. A controller including a state observer and an equivalent input interference predictor is designed to guarantee the global stability of the crane system. The proposed control strategy can effectively avoid the impact of measure errors and noise made by speed transmitter. Simulation results show that the work efficiency of the crane system is improved by a higher positioning accuracy and anti-swing control.

Keywords: Under-actuated cranes system · Equivalent input interference · Homeomorphic coordinate transformation · Global stability control

1 Introduction

Under-actuated cranes are a typical representative of under-actuated mechanical systems. In recent years, the control problems of under-actuated cranes have received extensive attention. Due to the operation of the crane and its acceleration and deceleration, the cargo will oscillate under the action of inertia, which will cause safety problems and affect the final hanging efficiency. Therefore, how to restrain the swing of the load during the running of the crane and the stability of the entire driving process is a research hotspot to ensure the safety of the system and improve the overall working efficiency of the crane [1].

In response to the above problems, scholars at home and abroad have done a lot of research work. Some use traditional control methods, such as double closed loop PID control, PID control, and some apply modern control theory, such as optimal control, trajectory planning [2–4], input shaping, energy/pассив based methods [5], partial feedback linearization, saturation control, sliding mode control, etc. In the literature, He Kongde et al. used the LQR optimal control method to analyze and control the swing of the crane due to the swing phenomenon caused by the inertia force and the external load disturbance during the working process of the gantry crane, and adopted

the Matlab software. The simulation results show that the control method can effectively realize the global stability control of the crane. Unfortunately, this article does not take into account the nonlinear characteristics of the crane system, because the actual crane is a strongly coupled nonlinear system, the nonlinear part is an important feature of the crane system, so this method is limited in practical applications.

Most of the above methods are based on the accurate linear part of the mathematical model design and analysis of the two-degree-of-freedom crane system, while ignoring the important nonlinear part of the crane mathematical model, and thus is limited in practical applications. When some research results are applied to crane systems, some more demanding assumptions or work environments must be proposed. In response to the above problems, inspired by the literature [6, 7], A control strategy based on equivalent input interference is proposed. The crane system is subjected to system coordinate transformation, which becomes a linear and nonlinear part through the separation system. At the same time, the nonlinear part is equivalent to the virtual interference at the input of the system. This not only preserves the nonlinear dynamics of the crane, but also reduces the difficulty of designing the under-executive nonlinear system controller.

This paper consists of the following three parts: 2 is the global stability control strategy design of underactuated crane system based on equivalent input, control performance simulation verification analysis. Finally, the work done in this paper is analyzed and summarized.

2 Underactuated Controller Design Based on Equivalent Input Interference

2.1 Description of the Problem

The physical structure model of a two-degree-of-freedom crane is shown in Fig. 1.

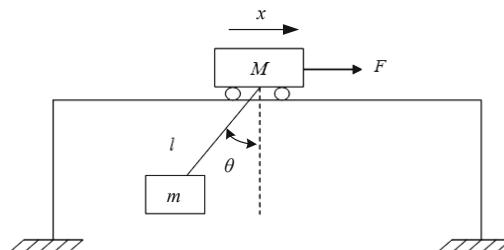


Fig. 1. Crane two-dimensional plane model

Assume that the length of the wire rope remains unchanged during transportation, while the friction between the car and the platform and the air resistance are ignored, that is $\dot{l} = \ddot{l} = 0$, $\mu = 0$, the following dynamic equation can be obtained:

$$\begin{bmatrix} m_2 l^2 & m_2 l \cos \theta \\ m_2 l \cos \theta & m_1 + m_2 \end{bmatrix} \begin{bmatrix} \ddot{\theta} \\ \ddot{x} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -m_2 l \dot{\theta} \sin \theta & 0 \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{x} \end{bmatrix} + \begin{bmatrix} m_2 g l \sin \theta \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ F(t) \end{bmatrix} \quad (1)$$

Among them: m_1, m_2 is the weight of the trolley and the weight of the hoist, l is the length of the sling, $F(t)$ is the input for the system, $x(t), \theta(t)$ are the swing level of the crane and the cargo angle with respect to the vertical direction.

2.2 Two-Degree-of-Freedom Under-Driven Crane System Homeomorphic Coordinate Transformation

The dynamic equation of the two-degree-of-freedom underactuated crane system is:

$$\begin{bmatrix} m_{11}(\theta, x) & m_{12}(\theta, x) \\ m_{21}(\theta, x) & m_{22}(\theta, x) \end{bmatrix} \begin{bmatrix} \ddot{\theta} \\ \ddot{x} \end{bmatrix} + \begin{bmatrix} C_{11}(\theta, x) & C_{12}(\theta, x) \\ C_{21}(\theta, x) & C_{22}(\theta, x) \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{x} \end{bmatrix} + \begin{bmatrix} G_1(\theta, x) \\ G_2(\theta, x) \end{bmatrix} = \begin{bmatrix} 0 \\ F(t) \end{bmatrix} \quad (2)$$

$M(\theta, x) = \begin{bmatrix} m_{11}(\theta, x) & m_{12}(\theta, x) \\ m_{21}(\theta, x) & m_{22}(\theta, x) \end{bmatrix} \in R^{4 \times 4}$ is the inertia matrix of the system.

It can be seen from the literature that although the two-degree-of-freedom underactuated crane system cannot fully linearize the state feedback, can be used as the input of the system, and it is designed to linearize the driving state quantity. Inspired by this, it is a crane. System (5) introduces the following transformation:

$$\Gamma \left\{ \begin{array}{l} z_1 = \theta_1 + \alpha(x) \\ z_2 = \frac{\partial L(\theta, x)}{\partial \dot{\theta}} = m_{11}(\theta, x)\dot{\theta} + m_{12}(\theta, x)\dot{x} \\ z_3 = x \\ z_4 = \dot{x} \end{array} \right. \quad (3)$$

Then

$$\dot{z} = h(z) + B_z u \quad (4)$$

$$\begin{aligned} h_1(z) &= [\frac{z_2}{m_{11}(\theta, x)} + (\frac{d\alpha(x)}{dx} - \frac{m_{11}(\dot{\theta}, \dot{x})}{m_{12}(\dot{\theta}, \dot{x})})z_4] \Big|_{\zeta=\Gamma^{-1}(z)} \\ h_2(z) &= [\frac{1}{2}(\dot{\theta}, \dot{x})^T \frac{\partial M(\dot{\theta}, \dot{x})}{\partial \theta} - G_1(\dot{\theta}, \dot{x})] \Big|_{\zeta=\Gamma^{-1}(z)} \\ h(z) &= [h_1(z), h_2(z), z_4, 0]^T \\ B_z &= [0, 0, 0, 1]^T \\ \zeta &= [\theta, x, \dot{\theta}, \dot{x}]^T \end{aligned} \quad (5)$$

In this paper, the choice of function $\alpha(\cdot)$ is very important. For the two-degree-of-freedom under-actuated crane, this paper chooses: $\alpha(x) = 0$, The system is approximated by a linearization of balanced position $z = 0$: $\dot{z} = A_z z + B_z u$.

Let $\sigma = h(z) - A_z z$, Then the system state equation of the new coordinate transformation can be obtained:

$$\begin{cases} \dot{z} = A_z z + B_z u + \sigma \\ y = C_z z \end{cases} \quad (6)$$

among them:

$$A_z = \begin{bmatrix} 0 & \frac{1}{ml^2} & 0 & -\frac{1}{l} \\ -mgl & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B_z = [0, 0, 0, 1]^T, C_z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\sigma = \begin{bmatrix} \frac{z_4 - \cos z_1 \cdot z_4}{l} \\ \frac{-z_2 \cdot \sin z_1 \cdot z_4 + ml \cdot \sin z_1 \cdot \cos z_1 \cdot z_4^2}{l} - mgl \cdot (\sin z_1 - z_1) \\ 0 \\ 0 \end{bmatrix}$$

So the linear part and the non-linear part of the system are separated as above by coordinate change. In this case, if the non-linear higher order term of the heavy machinery system is equivalent to the virtual input interference of the system, the linear system control design method can be used to design the controller.

If the crane system is controllable, the system can control the matrix, then $s = 0$. Only if the following conditions are met If the crane system is controllable, the system can control the matrix, then $s = 0$. Only if the following conditions are met $\left[\frac{m_{12}(\theta, x)}{m_{11}(\theta, x)} \frac{\partial G_1(\theta, x)}{\partial \theta} - \frac{\partial G_1(\theta, x)}{\partial x} \right] |_{\theta, x=0} \neq 0$, Similarly, the system observability matrix can be obtained, then the system can be viewed if and only if the following conditions are met $-\frac{1}{m_{11}(0)} < 0$, Similarly, the system observability matrix can be obtained, then the system can be viewed if and only if the following conditions are met $-\frac{1}{m_{11}(0)} < 0$.

2.3 Controller Design and Analysis

Suppose there is an equivalent input disturbance σ_e corresponding to the disturbance term σ such that the new equation is:

$$\begin{cases} \dot{z} = A_z z + B_z(u + \sigma_e) \\ y = C_z z \end{cases} \quad (7)$$

A control system based on equivalent virtual input interference. For traditional control system, the separation theorem shows that the pole of the closed-loop system is

the pole of the state feedback subsystem plus the pole of the observer, so the separation law is used for the equivalent input disturbance estimator, feedback controller, state the design of the observer, its schematic diagram is shown in Fig. 2.

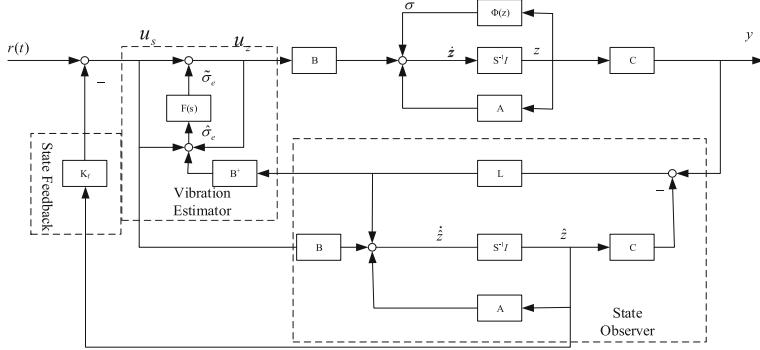


Fig. 2. Equivalent input interference controller structure diagram

(1) Equivalent input interference predictor design

In this paper, Longberg state observer is used to effectively estimate the real-time value of the equivalent input disturbance σ_e , as follows:

$$\dot{\hat{z}} = A_z \hat{z} + B_z u_z + L_z (y - C_z \hat{z}) \quad (8)$$

$u_s = u_z + \sigma_e$, L_z is observer gain, ensure that $A_z L_z C_z$ is a stable value. Let $\Delta z = \hat{z} - z$, Then $\dot{\hat{z}} = A_z \hat{z} + B_z u_z + B_z \sigma_e + (\Delta \dot{z} - A_z \Delta z)$, Suppose there is an $\Delta \sigma$ satisfied

$$\Delta \dot{z} - A_z \Delta z = B_z \Delta \sigma \quad (9)$$

Then by (9), (10)

$$\dot{\hat{z}} = A_z \hat{z} + B_z (u_z + \hat{\sigma}_e) \quad (10)$$

Among them $\hat{\sigma}_e = \sigma_e + \Delta \sigma$, get $B_z (u_z + \hat{\sigma}_e - u_s) = L_z (y - C_z \hat{z})$, Simultaneously multiply both ends of the upper type B_z^T , get $\hat{\sigma}_e = B_z^+ L_z (y - C_z \hat{z}) + u_s - u_z$, take the value of $\hat{\sigma}_e$ at this time as an estimate of the virtual input interference $\hat{\sigma}_e$, among them $B_z^+ = (B_z^T B_z)^{-1} B_z^T$. In addition, The filter FZ(S) is introduced to adjust the estimated range of the estimator to ensure accurate estimation of the virtual input interference estimate.

$$F_z(s) = \frac{1}{T_z s + 1} \quad (11)$$

T_z is the time constant.

(2) Feedback controller design

In view of the lack of an effective tool for solving the time-varying Riccati equation, the following performance index functions are used:

$$J_{K_z} = \int_0^{\infty} [z^T(t)Q_z z(t) + u_s^T(t)R_z u_s(t)] dt \quad (12)$$

The definition of the equivalent input interference shows that the control law is:

$$u_z = u_s - \tilde{\sigma}_e \quad (13)$$

(3) State observer design

The purpose of using the state observer is to construct an approximate state for the original system for feedback control law design, and second, to construct an equivalent interference estimate to compensate for the influence of the disturbance on the system. Here we adopt a high gain observer. The observer gain L_z is designed as

$$L_z = K_{z1}^T, K_{z1} = \frac{1}{2} R_{L_z}^{-1} C_z S \quad (14)$$

Where $Q_{L_z} = Q_{L_z}^T \geq 0, R_{L_z} = R_{L_z}^T \geq 0$ is the weighting matrix.

3 Simulation and Analysis

Aiming at the two-degree-of-freedom plane underactuated crane model, the control algorithm in this paper is used for simulation analysis. The simulation experiment is divided into 7 groups. Experiments 1, 2 and 3 are to verify that the control algorithm has the same horizontal moving distance P_{dx} in the trolley. The control effect on the crane is different when the specified time t_f is different or the same horizontal movement distance P_{dx} is different for the specified time t_f . Experiment 4 is used to verify the control effect of the external disturbance caused by the control algorithm under different operating conditions during the crane work; Experiment 5 is verified The superiority of the control algorithm in this paper over the fuzzy PID control algorithm.

Here, the control object parameters used in this article:

$$m_2 = 5 \text{ kg}, l = 1 \text{ m}, m_1 = 5 \text{ kg}, g = 9.8 \text{ m/s}^2$$

- (1) The purpose of simulation one to two simulation is to verify the control effect of the designed control algorithm on the crane system under different driving conditions of different driving horizontal distance P_{dx} or different specified time t_f .

Simulation 1: Expectation $p_{dx} = 30 \text{ m}$, $t_f = 50 \text{ s}$

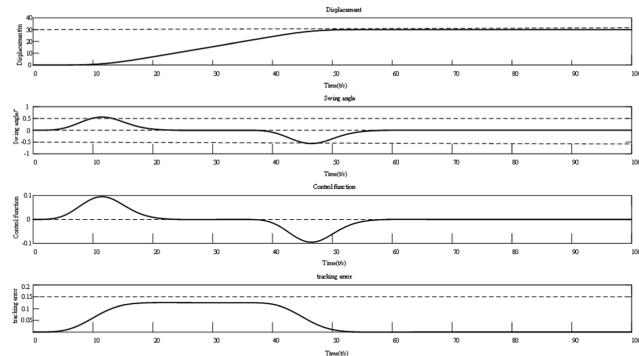


Fig. 3. Simulation results under $p_{dx} = 30 \text{ m}$, $t_f = 50 \text{ s}$ condition

Simulation 2: Expectation $p_{dx} = 30 \text{ m}$, $t_f = 30 \text{ s}$

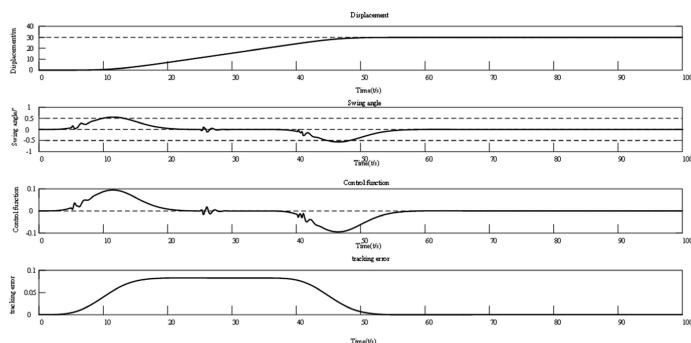


Fig. 4. Simulation results under $p_{dx} = 30 \text{ m}$, $t_f = 50 \text{ s}$ condition

Experimental simulation analysis: It can be seen from Figs. 3 and 4, when the crane is transported horizontally, the cargo swing angle remains at least 2° , when the crane reaches the desired position within the specified t_f time. swing angle has almost decayed to 0. The whole control process is relatively smooth and the tracking error control is small. At the same time, compared with Figs. 3 and 4, it can be concluded that the farther the moving distance is, the larger the swing angle is in the same time, but the swing angle is within 2° of the control range, and the control smoothing does not produce a controller. saturation.

- (2) Under the same simulation conditions as Simulation (1), increase the external interference and verify the effectiveness of the designed control algorithm.

Simulation 3: Add a Gaussian white noise perturbation of a weight angle for one second at $t_f = 5 \text{ s}$, $t_f = 25 \text{ s}$, $t_f = 40 \text{ s}$, expectation $p_{dx} = 30 \text{ m}$, $t_f = 50 \text{ s}$:

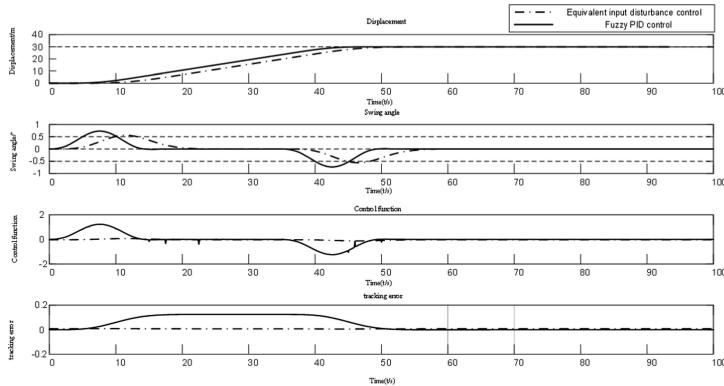


Fig. 5. Simulation results under $p_{dx} = 30 \text{ m}$, $t_f = 50 \text{ s}$ condition and adding Gaussian white noise to the swing angle

Simulation analysis: It can be seen from Fig. 5, when the cargo is running and the external influence increases, the equivalent input disturbance can be estimated and suppressed, and received when $t_f = 5 \text{ s}$, $t_f = 25 \text{ s}$, $t_f = 40 \text{ s}$. In the case of 1 s Gaussian white noise interference, the swing angle can quickly recover to near zero without affecting the positioning accuracy and anti-shake effect of the crane.

In addition, the comparative analysis of simulation (2) and simulation (1) shows that when the heavy object is greatly affected by external influences, it can not only return to normal in a short time, but also does not affect the tracking of the control car trajectory. Achieve global stability control.

- (3) Compared with the fuzzy PID control effect, the effectiveness of the de-signed control algorithm is verified.

Simulation 4: Fuzzy PID control and the equivalent input interference control effect of this paper are shown in Fig. 6.

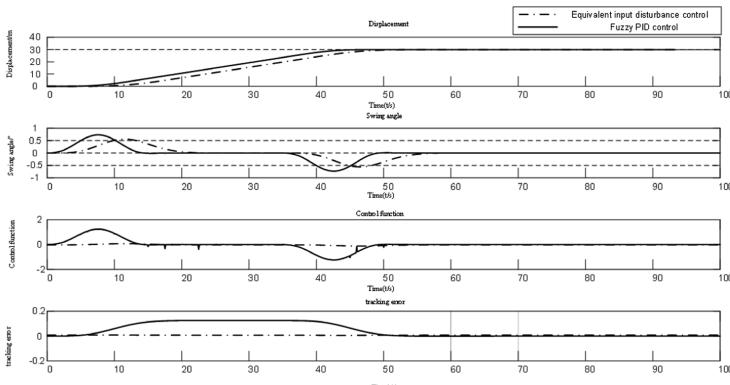


Fig. 6. Comparison of fuzzy PID and equivalent input interference control effects (Expectations: $p_{dx} = 20 \text{ m}$, $t_f = 50 \text{ s}$)

Simulation analysis: It can be seen from Fig. 6 that the equivalent input interference control effect is smooth, the control force is soft, and the swing angle of the weight can be controlled within a small range. However, the control error is slightly higher than the fuzzy PID control. Under normal circumstances, the smoothness of the control process is high and the control tracking accuracy is high and contradictory. To pursue high tracking accuracy, it is necessary to reduce the smoothness. This paper uses equivalent input interference to achieve a smooth control effect, although the tracking accuracy is reduced. However, as can be seen from the simulation results, the post-error is almost zero after $t_f = 50$ s.

4 Conclusions

In this paper, a new global stability control method for under-actuated cranes is designed. While preserving the nonlinear characteristics of the crane system, it can accurately locate the horizontal direction of the crane and effectively suppress and eliminate the swing of the load. The method can realize global stability control only by using the position information, which can not only reduce the controller cost, but also avoid the interference of the Velocity sensor error and measurement noise on the control, and thus has significant practical application significance. The method proposed in this paper only needs to design the controller by using the position information of the crane. The simulation results show the effectiveness of the designed method. Next we will carry out experimental verification and improve the control method to improve the system's response speed and tracking accuracy.

References

1. Liu L (2014) Summary of research on anti-swing control method of under-actuated shore container crane. *Autom Instrum*
2. Chen H, Fang Y, Sun N (2017) A swing constrained time-optimal trajectory planning strategy for double pendulum crane systems. *Nonlinear Dyn* 2017(7):1–12
3. Ma B, Fang Y, Zhang X (2008) Adaptive tracking control for an overhead crane system. *IFAC Proc Vol* 41(2):12194–12199
4. Wang J, Noda Y, Inomata A (2015) Straight transfer control system using PI control and trajectory planning in overhead traveling crane. In: IEEE/SICE international symposium on system integration. IEEE, pp 522–527
5. Liu X, Yang Y (2009) Observer-based robust predictive control of singular systems with time-delay and parameter uncertainties. *J Syst Eng Electron* 20(3):590–597
6. Spong M (1995) The swing up control problem for the acrobot. *IEEE Control Syst Mag* 15(1):49–55
7. Xie W (2013) Global asymptotic tracking of asymmetrical underactuated surface vessels with parameter uncertainties. *J Control Theory Appl* 11(4):608–614



Truck Size Measurement System Based on Computer Vision

Zhengguang Xu and Meiling Yang^(✉)

School of Automation and Electrical Engineering,
University of Science and Technology Beijing, Beijing 100083, China
934729954@qq.com

Abstract. According to the application background of the size measurement of trucks in cement plants, the car size measurement system was developed according to the monocular vision principle. The key frame is extracted by the image difference method, then sub-pixel fine positioning is performed by the Zernike moment method. The edge connection is performed by the Progressive Probabilistic Hough Transform algorithm, and the internal and external parameter matrix obtained by the camera calibration is combined to complete the automatic measurement of the truck's volume. By comparing with the measured data, the measurement accuracy meets the requirements of the factory. This also verifies the feasibility and accuracy of this measurement method.

Keywords: Computer vision · Three-dimensional measurement · Sub-pixel edge extraction · Camera calibration

1 Introduction

Cement is one of the main building materials. Driven by the rapid development of China's national economy, the demand for cement in industrial and agricultural production is growing. However, some factories are still relying on manual loading of cement. The dust on the scene is flying. The pollution is serious. And it directly harms the health of the workers. So FengBo Cement Company located XinXiang proposed to design an automatic loading system to achieve the goal of ensuring safe production.

The automatic loading system consists of a robotic arm, a camera and a control computer. The worker manually clicks on the camera to open the camera. The camera transmits the image of the cemented vehicle on the scene to the control computer. The computer processes the image and calculates the length, width and height of the compartment, and then transmits the size information to the PLC controller of the robot arm. The robot control arm places the cement in the planned position inside the compartment to complete the automatic loading process [1]. The author uses the monocular vision technology to study the non-contact rapid detection system of the truck size parameters to achieve accurate measurement of truck size parameters.

2 Truck Target Detection

Moving target detection is to remove the background content from the video image and detect the target of the motion. This system uses the method of inter-frame difference [2] to perform moving target detection of images. According to the law of on-site vehicle entry and exit, we use the inter-frame difference to judge whether the vehicle stops, and after extracting the vehicle, the key frame is extracted for the next operation. The reason why the last frame information is used instead of the video stream information is because the site environment is more complicated. The cement loaded on the car is sometimes similar to the ground texture, and the lights will suddenly turn on or off at night. It has a great impact on the accurate positioning of the target.

Use the following steps to perform moving target detection, which detects that the cement truck enters the work area and stops.

Step1: According to the characteristics of the image, we perform inter-frame difference operation on each frame image and the previous frame image.

Step2: The average gray scale is calculated for the difference map after the difference between frames, and a threshold T is set. If the average gray value after the difference is greater than the threshold T , it indicates that the truck is traveling in the lane.

Step3: When the image of several consecutive frames no longer changes significantly, that is, the frame difference is less than the threshold, the car may stop and be detected. At this point we can extract this frame image for further frame edge determination.

The effect picture after key frame extraction is shown in Fig. 2(a).

3 Image Preprocessing

3.1 Image Region of Interest Extraction

Since the camera is fixed, the work area of the cement truck is also fixed, and the edge of the frame is also within this range, so that the vehicle stop area in the image is used as the image processing area of interest. In order to improve the calculation speed and eliminate other interference factors, we cut out this small range including the vehicle as the action area for the next image processing. The extracted region of interest is shown in Fig. 2(b).

3.2 Image Noise Reduction

Digital images are often contaminated with different types of noise, such as Gaussian noise and impulse noise. Gaussian noise is often caused by electrical noise from the sensing device, while salt and pepper noise, black and white points in the image, is typically produced by image sensors, channel transmission, and decoding processes. Noise removal plays an important role in image processing, and it has a direct impact on image segmentation, registration, feature extraction and image recognition. Signal

and noise always exist at the same time. Protecting the signal and suppressing noise, that is, improving the signal-to-noise ratio, has always been a hot issue in image preprocessing [3].

Observing the image, it can be seen that the image of the truck mainly has Gaussian noise. Therefore, this paper uses a Gaussian filter of size 5 to smooth the image. Gaussian filtering has little effect on the texture features of the image itself, and the noise is suppressed while maintaining the image edge information. The effect of image processing is shown in Fig. 2(c).

3.3 Rough Positioning of the Truck's Edge

At the present stage, the traditional image edge detection algorithm is more mature. The differential operation is the main method of traditional edge detection and extraction. According to the characteristic that the edge has a step change, the edge point is located at the point where the first-order differential is the largest or the smallest, or the second-order differential is the zero-cross or the second-order differential. The first-order differential operators include the Robert operator [4], the Sober operator, the Prewitt operator [5], the Kirsch [6], and the Canny operator [7]. Second-order differential operators include Laplacian operators, LOG operators [8], and so on. This paper chooses the Canny operator which is the most widely used. The detection effect diagram is shown in Fig. 2(d).

Steps of the Canny edge detection algorithm are as follows:

- Step1: Using Gaussian smoothing filter convolution noise reduction;
- Step2: Using the finite difference of the first-order partial derivative to calculate the amplitude and direction of the gradient;
- Step3: Non-maximum suppression of the gradient amplitude;
- Step4: Detecting the edge with a double threshold algorithm.

3.4 Precise Positioning of the Truck's Edge

It can be seen from Fig. 2(d) that the detection effect is not ideal, and some important edges are not connected, so the author uses the subpixel edge detection method to perform fine positioning based on the canny edge detection. The sub-pixel edge detection algorithms currently studied can be summarized into three types: interpolation method [9], fitting method [10], and moment method [11]. The interpolation method is simple, but is susceptible to noise. The fitting method is not sensitive to noise, but due to the complexity of the model, its solution speed is slow. The moment method is simple to calculate and an analytical solution can be obtained. The Zernike moment method is much smaller than the space moment method because only three templates need to be calculated. Therefore, I chose to use Zernike moments for edge positioning. The effect picture is shown in Fig. 2(e).

The principle of subpixel edge detection using the Zernike moment method is as follows:

The ideal step edge model of the two-dimensional pixel is shown in Fig. 1. The ideal sampling area is a unit circle, where k is the step height, h is the background gray,

and the angle between the edge normal and the x-axis is φ , the unit circle The vertical distance from the center to the edge is l .

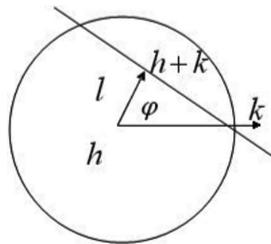


Fig. 1. The ideal step edge model of the two-dimensional pixel

- Step1: Establish three moments Z_{00}, Z_{11}, Z_{20} corresponding to four parameters such as φ, h, k, l etc., wherein the integral kernel functions corresponding to the three moments are $V_{00} = 1$, $V_{11} = x + iy$, $V_{20} = 2x^2 + 2y^2 - 1$;
 Step2: Solving the above three moments and obtaining the parameters in the moment definition, the step edges can be represented by φ, h, k, l ;
 Step3: Compare the obtained parameters with the set thresholds to finally accurately locate the edges of the image.

In the unit circle ($x^2 + y^2 = 1$), according to the moment definition method proposed by Zernike, the complete orthogonal basis represented by the complex polynomial is introduced, and the four parameters of the ideal step edge model can be calculated, according to the moment definition method proposed by Zernike, the complete orthogonal basis represented by the complex polynomial is introduced, and the four parameters of the ideal step edge model can be calculated.

$$\varphi = \arctan\left(\frac{I_m[Z_{11}]}{R_e[Z_{11}]}\right) \quad (1)$$

$$l = \frac{Z_{20}}{Z'_{11}} = \frac{Z_{20}}{Z_{11}} e^{-j\varphi} \quad (2)$$

$$k = \frac{3Z'_{11}}{2(1-l^2)^{3/2}} = \frac{3Z_{11}}{2(1-l^2)^{3/2}} e^{-j\varphi} \quad (3)$$

$$h = \frac{Z_{00} - \frac{k\pi}{2} + karcsinl + kl\sqrt{1-l^2}}{\pi} \quad (4)$$

In the discrete case, the moment can be calculated by convolving the template with the gray level of the image. Considering the template effect, if the unit circle is used for

sampling, in the $N * N$ pixel region, when the template is moved over the image for convolution, the template covers N^2 pixels in the center of the template, and the unit radius becomes N at this time. So the vertical distance l calculated on the unit circle needs to be enlarged by $N/2$ times, so the calculation formula of the sub-pixel coordinates of the pixel point:

$$\begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \frac{N}{2} l \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix} \quad (5)$$

3.5 Compartment Edge Connection

After using the edge detection algorithm to extract a series of pixel points along the target boundary, extracting meaningful features from it is an important part of computer vision measurement. The feature points are defined as points where the brightness of the two-dimensional image changes sharply or points with maximum curvature values on the edge curve of the image, also called corner points, and other features include lines, curves, contours, and the like. The author uses Hough transform to realize the feature fitting. Hough transform is an algorithm that uses the global features of the image to connect the edge pixels to form the boundary of the closed region. Under the condition that the shape of the region is known in advance, the boundary can be conveniently obtained and discrete edge pixels can be connected. Edge pixels are connected. Based on the prior knowledge of the characteristics of the truck, such as the car is roughly rectangular, aspect ratio, etc., combined with the cumulative probability Hough transform to extract the body features. The extraction results are shown in Fig. 2(f).

4 Camera Calibration and Carriage Coordinate Calculation

The relationship between the three-dimensional geometric position of a point on the surface of a space object and its corresponding point in the image is determined by the geometric model imaged by the camera. These geometric model parameters are camera parameters. Under most conditions, these parameters must be obtained through experiments and calculations. This process is called camera calibration. Classic calibration methods, in order to get accurate standard point coordinates, usually need to make calibration reference objects (such as checkerboard, inscribed circles, etc.), and the size of these calibration templates is generally equivalent to the target to be tested. When the size of the target to be measured is relatively large, it is difficult and costly to produce a large-scale, high-precision calibration template, and sometimes it is impossible to operate in practice. Therefore, this paper proposes an improved three-dimensional calibration scheme for the field environment. Under the premise of not affecting production, the use of the site's points for maximum stereo calibration. First set the world coordinate system, use the total station to measure the world coordinates of more than a

dozen non-coplanar points in the field, and manually select the image coordinates of the calibration point on the image with the mouse, and use the two-step calibration algorithm to obtain the image coordinate system to the world coordinates. The mapping relationship of the system. For reasons of space, it will not be repeated here.

The coordinates of the eight key points after the edge extraction are substituted into the internal and external parameter matrix obtained by the camera calibration to obtain the world coordinates of the corner points of the car. Finally, the length, width and height of the car are calculated by using the distance between the points.

5 Experimental Results and Error Analysis

5.1 Experimental Results

This article is based on the VS2013 platform and is experimented with the OpenCV open source image library. To verify the feasibility and accuracy of the method, we write the program according to the above method and perform the experiment after debugging. Image processing results are shown in Fig. 2.

Finally, five different trucks were selected as experimental subjects. Calculate the length, width, height and volume of the truck and define the relative error as the percentage difference between the calculated volume and the measured volume. The experimental results are as follows (Table 1):

The error is within 5% allowed by the factory.

5.2 Analysis of Experimental Results

By analyzing the on-site environment, hardware system construction and software system of the calibration scheme, the sources of system errors are summarized as follows:

① Errors caused by the performance of experimental equipment

Such as: the measurement error caused by the resolution of the CCD camera device, the higher the resolution, the higher the measurement accuracy; the image edge extraction error caused by the uneven brightness of the light source; the inaccuracy of the standard block size and the uneven reflection degree can also bring measurement error; lens distortion cannot be completely corrected to cause image nonlinear distortion error and so on.

② Algorithm error

The edge extraction method obtains the error that the coordinates of the edge point and the real coordinates of the point cannot be avoided.

③ Calibration error

In the calibration process of the image, the nonlinear distortion of the CCD and the calibration algorithm will also bring certain calibration deviations.



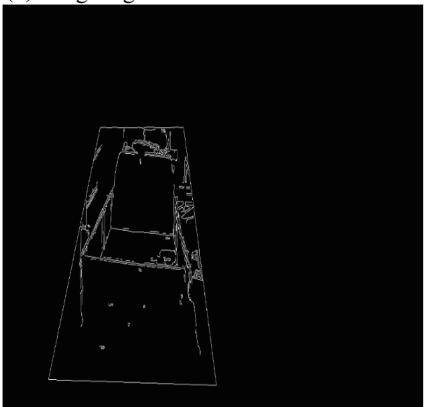
(a) Key frame extraction



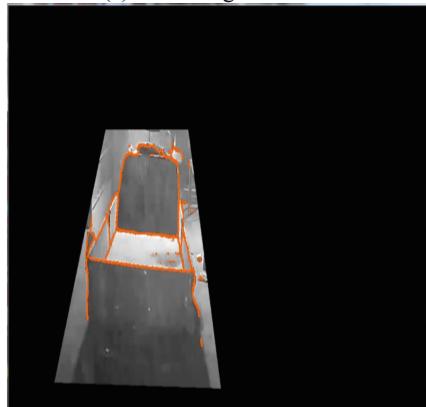
(b) Image region of interest extraction



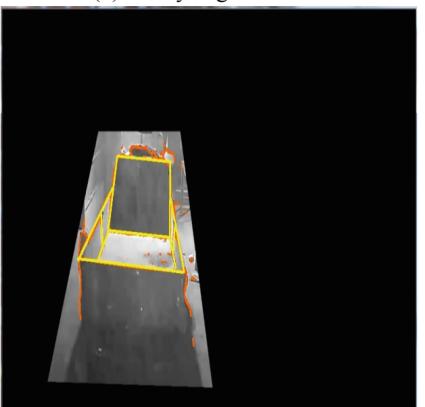
(c) Truck target detection



(d) Canny edge detection



(e) Subpixel edge detection



(f) Edge connection

Fig. 2. Image processing results

Table 1. Comparison between actual data and calculated data.

| Number | Calculated data/cm | | | Actual data/cm | | | Relative error/% |
|--------|--------------------|-------|--------|----------------|-------|--------|------------------|
| | Length | Width | Height | Length | Width | Height | |
| 1 | 679.5 | 249.6 | 58.4 | 680 | 250 | 60 | 2.9 |
| 2 | 399.6 | 221.5 | 51.8 | 400 | 220 | 50 | 4.2 |
| 3 | 598.8 | 221.2 | 59.7 | 600 | 220 | 60 | 0.2 |
| 4 | 550.5 | 210.6 | 53.9 | 560 | 200 | 55 | 1.4 |
| 5 | 300.7 | 168.8 | 51.1 | 300 | 170 | 50 | 1.7 |

6 Conclusion

This paper presents a method for measuring the size of a car in a complex background. Firstly, the key frame is extracted by the image difference method, then the edge detection is performed by the combination of the canny operator and the Zernike moment method. The edge connection is performed by the Hough probability transformation algorithm, and the internal and external parameter matrix obtained by the camera calibration is used to complete the automatic volume of the compartment. Comparing with the measured data, the measurement accuracy meets the requirements of the factory, and it can provide accurate data support for subsequent cement placement location planning and manipulator path planning. This also verifies the feasibility and accuracy of this measurement method.

References

1. Liao M (2016) Design of control system for cement automatic loading machine. Hubei University of Technology, Hubei
2. Black MJ, Jepson AD (1997) EigenTracking: robust matching and tracking of articulated objects using a view-based representation. *Int J Comput Vis* 26(1):64–84
3. Arce GR, McCloudlin MP (1987) Theoretical analysis of the MAX/Median filter. *IEEE Trans Acoust Speech Signal Process* 35(1):60–69
4. Roberts L (1963) Machine perception of three-dimensional solids. MIT Press, Cambridge
5. Prewitt JM (1970) Object enhancement and extraction. In: Picture Processing and Psychopictorics. Academic Press, New York, pp 75–149
6. Kirsch RA (1971) Computer determination of the constituent structure of biological images. *Comput Biol Res* 4(3):315–328
7. Canny JF (1986) A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell PAMI* 8(6):679–698
8. Marr DC, Hildreth EC (1980) Theory of edge detection. *Proc Roy Soc Lond Ser B Biol Sci* 207(1167):187–217
9. Li YS, Young TY (1988) Subpixel edge detection and estimation with a microprocessor controlled line scan camera. *IEEE Trans Ind Electron* 35(1):105–112

10. Zheng Y, Qian W, Luo J (2007) Study on subpixel edge localization method based on curve fitting. *Opt Technol* (03):386–389
11. Chang-Hasnain CJ, Harbison JP (1991) Dynamic, polarization, and transverse mode characteristics of vertical cavity surface emitting lasers. *IEEE J Quantum Electron* 27(6):1402–1409



Dual Space Vector Control Study of Two Stage Matrix Converter Fed PMSM

Huida Duan^{1(✉)} and Xin Peng²

¹ Changchun Institute of Technology, Changchun 130012, China
huida_duan@126.com

² Beihua University, Jilin 132021, China

Abstract. Two-stage matrix converter (TSMC) is the most advanced power converters which provide direct ac-ac power conversion without needing a dc link. Permanent magnet synchronous motor (PMSM) has attracted more and more scholars' interest in industrial applications due to their outstanding. In this paper, the AC adjustable-speed system based on the PMSM which is driven by TSMC is proposed. The dual space vector modulation (DSVM) method for three phases TSMC is studied in detail. In addition, the vector control technique is applied for the TSMC fed PMSM drive. To verify the effectiveness and feasibility of the proposed control methods, some simulations using MATLAB have been performed. The simulation results show that the proposed AC adjustable-speed drivers system has good performance and high efficiency.

Keywords: TSMC · DSVM · PMSM · Vector control

1 Introduction

In recent researches for direct frequency conversion technique focus is made on using a matrix converter (MC). The main reason for the interest in MCs is that it provides a compact solution for a four quadrant frequency converter producing sinusoidal input and output waveforms without passive components in the DC-link [1]. The matrix converter topologies can be divided into two types: direct matrix converters (DMCs) and two-stage matrix converters (TSMCs). Both of these converters are able to generate input and output waveforms with same quality. However, in some applications, the TSMCs may be preferred to direct matrix converters due to safe commutation switches and compact circuit and to reducing the number of power semiconductor switches [2]. Growing interest may be observed in the PMSMs as the best choice of machines in several applications due to its high power density, high torque-to-inertia ratio, small torque ripple and precise control in the low speed range, possibility for torque control at zero speed, high efficiency, and small size. Instead of the inverter with DC-link reactive energy storage components in the vector-controlled drive system, the use of the two-stage matrix converter made the drive system capable of operating in all four-quadrant regions. It was demonstrated that the two-stage matrix converter topology is an alternative solution to the commonly used voltage source inverter with diode rectifier and is characterized by higher performance of output signals, controllable input power factor, lower volume, inherent bidirectional power flow, and long lifetime due to the

absence of reactive components in the intermediate DC-link [3]. The topology of three phase two-stage matrix converter is shown in Fig. 1.

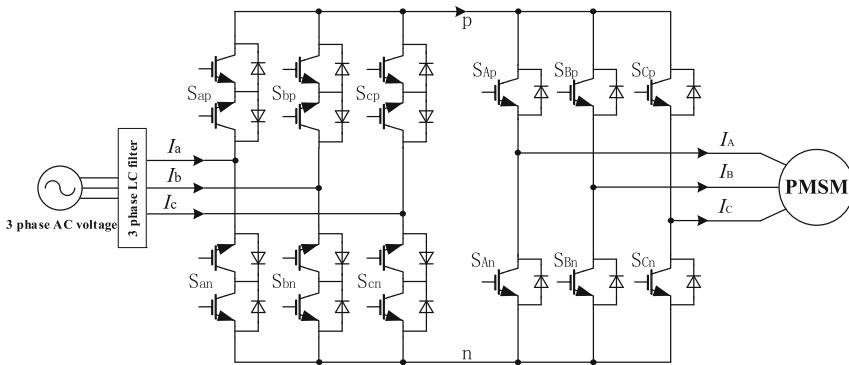


Fig. 1. The topology of three phase TSMC

2 Modulation of the TSMC

The two-stage matrix converter functions are equivalent to those of direct matrix converter. The TSMC topology consists of front-end current source rectifier (CSR) and a rear-end voltage source inverter (VSI). Dual Space Vector Modulation (DSVM) is applied for both the CSR and VSI. The DSVM control technique of TSMC is based on the instantaneous space vector representation of input and output voltages and currents [4].

2.1 SVM for the Rectifier Stage

For safe operation of the system, the six bidirectional switches of the rectifier stage can have only nine allowed combinations to avoid open circuit at the DC-link rails. The nine combinations can be divided into six non-zero input current vectors which are called active vector I_1 – I_6 and three zero input current vectors which are called zero vector I_0 [5].

The CSR input current vector diagram is shown in Fig. 2. The space vector of the desired input current I_l^* can be approximated by two adjacent vectors as shown in Fig. 3. The duty cycles of the switches vectors are written as

$$\begin{cases} d_\gamma = T_\gamma/T_S = m_C \cdot \sin(\frac{\pi}{3} - \theta_C) \\ d_\delta = T_\delta/T_S = m_C \cdot \sin(\theta_C) \\ d_{0C} = T_{0C}/T_S = 1 - d_\gamma - d_\delta \end{cases} \quad (1)$$

where, T_s is the period of a switching cycle; d_γ , d_δ and d_{0C} are duty cycles of two active vectors and one zero vector, respectively. θ_C is the angle between input phase current vector and starting edge of its locating sector, as shown in Fig. 3.

The m_C is the current modulation index, $0 \leq m_C \leq 1$ and defines the desired current transfer ratio such as

$$m_C = \frac{I_L^*}{I_{DC}} \quad (2)$$

where, I_{DC} is the current of the intermediate DC-link.

In order to maximize the intermediate DC-link voltage to increase utilization of the input voltage, the current modulation index m_C is often fixed to unity and the voltage modulation index m_V is variable according to a required overall voltage transfer gain.

In the SVM, it is synthesized by time averaging from a selection of adjacent vectors in the set of converter output vectors in each sampling period. The synthesis of the input currents and DC-link voltages for a switching cycle within the current sector VI is chosen as an example:

$$\begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} = d_\gamma \cdot I_1 + d_\delta \cdot I_2 = \left(d_\gamma \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} + d_\delta \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \right) \cdot \begin{bmatrix} I_{DC+} \\ I_{DC-} \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} V_{DC+} \\ V_{DC-} \end{bmatrix} = \left(d_\gamma \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + d_\delta \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \cdot \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} \quad (4)$$

where, d_γ and d_δ are the duty cycles of the active vector I_1 and I_2 , respectively. V_{DC+} and I_{DC+} are the voltage/current of the positive rail of the intermediate DC-link, respectively. V_{DC-} and I_{DC-} are the voltage/current of the negative rail of the intermediate DC-link, respectively.

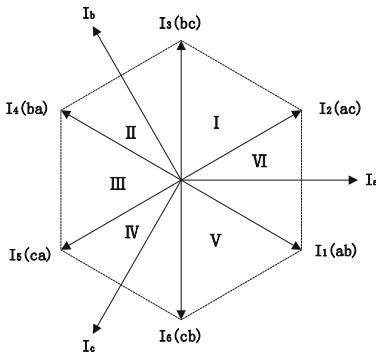


Fig. 2. Sector of input current space vector

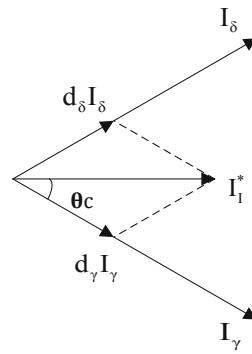


Fig. 3. Synthesis of input current space vector

2.2 SVM for the Inverter Stage

The inverter stage of the TSMC has a standard 3 phase voltage source inverter topology consisting of six unidirectional switches. The inverter switches can have only eight

allowed combinations to avoid a short circuit through three half bridges for safety. The eight combinations can be divided into six nonzero output voltage vectors which are active vector V_1-V_6 and two zero output voltage vectors which are zero vector V_0 . The VSI output voltage vector diagram is shown in Fig. 4. The space vector of the desired output voltage $V^* o$ can be approximated by two adjacent vectors as shown in Fig. 5. Similarly, the duty cycles of the switching state vectors are calculated by [5].

$$\begin{cases} d_\alpha = T_\alpha/T_S = m_V \cdot \sin(\frac{\pi}{3} - \theta_V) \\ d_\beta = T_\beta/T_S = m_V \cdot \sin(\theta_V) \\ d_{0V} = T_{0V}/T_S = 1 - d_\alpha - d_\beta \end{cases} \quad (5)$$

where, d_α , d_β and d_{0V} are duty cycles of two active vectors and one zero vector, respectively. θ_V is the angle between output phase voltage vector and starting edge of its locating sector, as shown in Fig. 5.

The m_V is the voltage modulation index, $0 \leq m_V \leq 1$ and defines the desired voltage transfer ratio such as

$$m_V = \frac{\sqrt{3} \cdot V_o^*}{V_{DC}} \quad (6)$$

where, V_{DC} is the voltage of the intermediate DC-link.

The synthesis of the output phase voltages and DC-link currents for a switching cycle within the voltage sector VI is chosen as an example.

$$\begin{bmatrix} V_A \\ V_B \\ V_C \end{bmatrix} = d_\alpha \cdot V_6 + d_\beta \cdot V_1 = \left(d_\alpha \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} + d_\beta \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \right) \cdot \begin{bmatrix} V_{DC+} \\ V_{DC-} \end{bmatrix} \quad (7)$$

$$\begin{bmatrix} I_{DC+} \\ I_{DC-} \end{bmatrix} = \left(d_\alpha \cdot \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} + d_\beta \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \right) \cdot \begin{bmatrix} I_A \\ I_B \\ I_C \end{bmatrix} \quad (8)$$

where, d_α and d_β are the duty cycles of the active vector V_1 and V_2 , respectively.

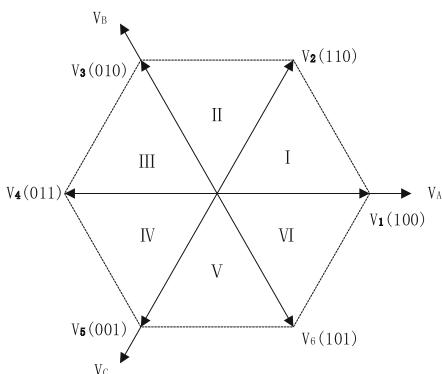


Fig. 4. Sector of output voltage space vector

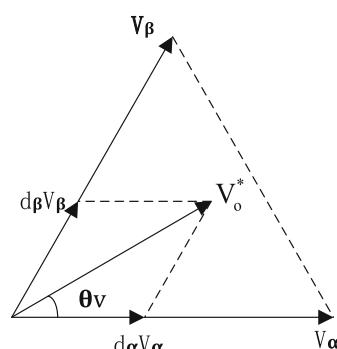


Fig. 5. Synthesis of output voltage space vector

2.3 Synthesis of SVM for the Entire TSMC

When the V_α and V_β are applied to the first current vector I_γ , two new vectors, $V_\alpha - I_\gamma$ pair and $V_\beta - I_\gamma$ pair, are created and duty cycles of new vectors becomes $d_{\alpha\gamma}$ and $d_{\beta\gamma}$, respectively, as defined in formula (9). When the V_α and V_β are applied to the second current vector I_δ , two new vectors, $V_\alpha - I_\delta$ pair and $V_\beta - I_\delta$ pair, are created and duty cycles of new vectors becomes $d_{\alpha\delta}$ and $d_{\beta\delta}$, respectively. And during the remaining part of the switching period T_s , zero vector is applied [6].

$$\begin{cases} d_{\alpha\gamma} = d_\alpha \cdot d_\gamma = m_V \cdot \sin(\frac{\pi}{3} - \theta_V) \cdot \sin(\frac{\pi}{3} - \theta_C) = \frac{T_{\alpha\gamma}}{T_s} \\ d_{\alpha\delta} = d_\alpha \cdot d_\delta = m_V \cdot \sin(\frac{\pi}{3} - \theta_V) \cdot \sin(\theta_C) = \frac{T_{\alpha\delta}}{T_s} \\ d_{\beta\gamma} = d_\beta \cdot d_\gamma = m_V \cdot \sin(\theta_V) \cdot \sin(\frac{\pi}{3} - \theta_C) = \frac{T_{\beta\gamma}}{T_s} \\ d_{\beta\delta} = d_\beta \cdot d_\delta = m_V \cdot \sin(\theta_V) \cdot \sin(\theta_C) = \frac{T_{\beta\delta}}{T_s} \\ d_0 = 1 - d_{\alpha\gamma} - d_{\alpha\delta} - d_{\beta\gamma} - d_{\beta\delta} = \frac{T_0}{T_s} \end{cases} \quad (9)$$

where, $d_{\alpha\gamma}$, $d_{\alpha\delta}$, $d_{\beta\gamma}$ and $d_{\beta\delta}$ are duty cycles of corresponding new vectors; d_0 is the duty cycle of zero vector.

3 Vector Control of TSMC-PMSM

Simulation has been done using Matlab/Simulink software to show the validity and feasibility of the proposed method. The simulation has been implemented on both 3 phase RL load and PMSM. Firstly, the TSMC fed RL load is simulated. Figures 6 and 7 show the input phase voltage and current waveforms, output line voltage and current waveforms, respectively. It shows that the effectiveness of dual space vector modulation strategy for the TSMC. Finally, and most importantly, a closed-loop technique, namely vector control, has been implemented in the TSMC fed PMSM drive. Figure 8 shows a schematic block diagram of the vector controlled PMSM driven by TSMC, where the rotor position and rotor speed of the PMSM are required to be measured by shaft encoder [7, 8]. During the implementation of the vector control, the direct-axis component reference of stator current $i_d^* = 0$ is employed in this paper.

The motor speed, ω is measured and compared with the demanded speed ω^* . The resulting speed error is then processed by a proportional-integral (PI) controller to produce an i_q^* demand, which in the constant torque region is proportional to the torque demand, providing that the system is field-oriented, the technique used is also called vector control. The synchronous frame d-q current regulator utilizes the references and feedback currents to generate the voltage reference signals for the TSMC controller [9].

The detailed simulation process is as followed: the reference speed was first changed from zero to 100 rad/s. After steady-state operation for 0.5 s, the speed command was given as -100 rad/s. Following the steady-state operation for 1 s again, it has been changed to 100 rad/s to prove four-quadrant process of the drive system [10]. And the motor load torque reference was changed from 10 N m to -10 N m at the time of 1 s during the simulation process. As can be seen from the simulation, the

speed of PMSM has tracked its reference commands very well, as demonstrated in Fig. 9. The corresponding electromagnetic torque and stator current waveforms of i_{sd} and i_{sq} are illustrated in Figs. 10, 11 and 12, respectively.

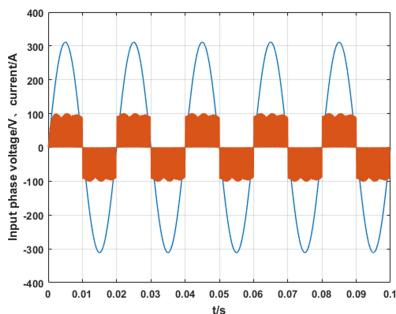


Fig. 6. The input phase voltage and current

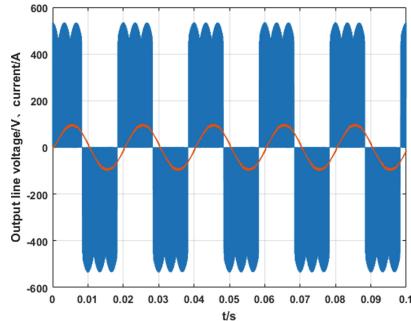


Fig. 7. The output line voltage and current

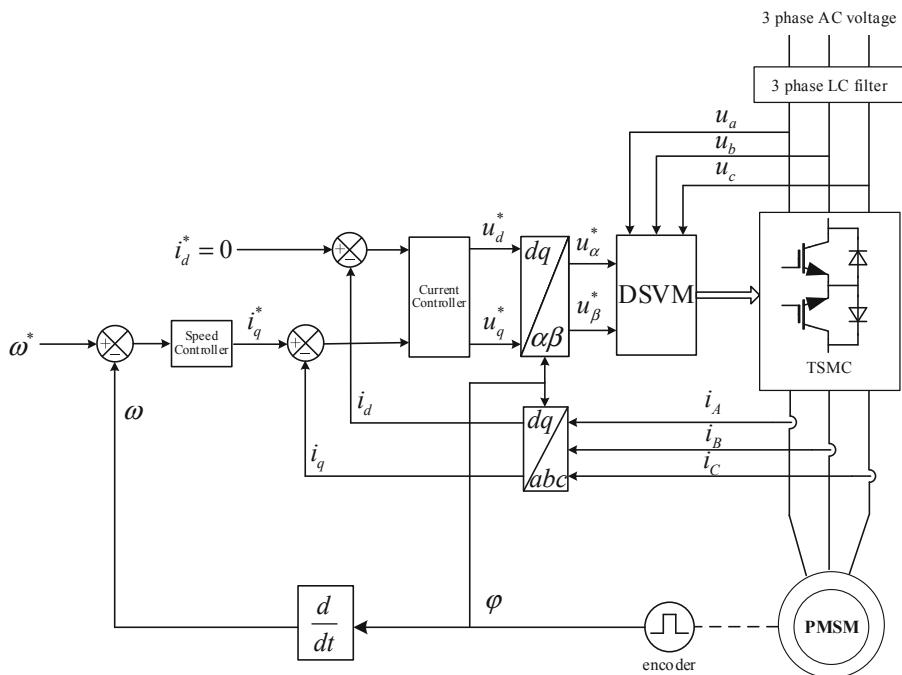
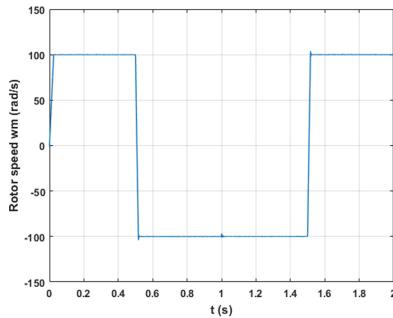
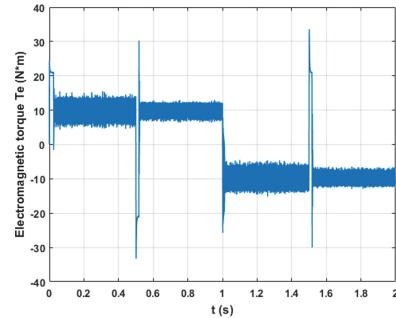
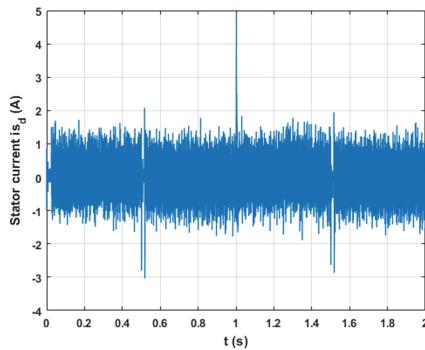
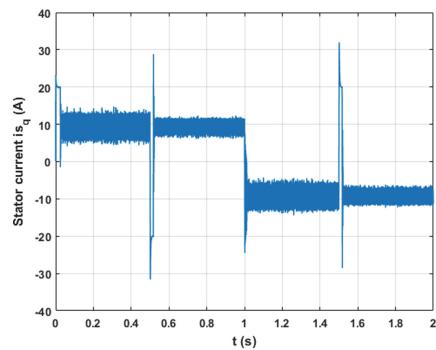


Fig. 8. Schematic diagram of the vector controlled PMSM driven by TSMC

**Fig. 9.** Rotor speed**Fig. 10.** Electromagnetic torque**Fig. 11.** Stator current d-axis component**Fig. 12.** Stator current q-axis component

4 Conclusion

This paper represents a dual space vector modulation (DSVM) method for the two-stage matrix converter that feeds PMSM drive. The vector control strategy was implemented in the PMSM driven by TSMC. From the results of the simulation, the use of the two-stage matrix converter made the drive system indeed capable of operating in all four quadrant regions, therefore, a high energy efficiency and very rapid speed response could be achieved by employing the TSMC.

Acknowledgments. This work is supported by the Fund Project of Jilin Education Department (2015162) and Jilin Provincial Department of Science and Technology (20190101018JH).

References

1. Natarajan R, Baskaran B, Iruapparaj G (2014) Simulation and speed control of induction motor fed by indirect matrix converter. *Int J Dev Res* 4(3):565–569
2. Gang L (2015) Realization of SVM algorithm for indirect matrix converter and its application in power factor control. *Adv Power Electron* 2015(740470):1–10

3. Szczesniak P, Urbanski K, Fedyczak Z (2016) Comparative study of drive systems using vector-controlled PMSM fed by a matrix converter and a conventional frequency converter. *Turk J Electr Eng Comput Sci* 24(3):1516–1531
4. Chen YC, Liu ZW, Zou YW (2016) Research on SVM strategy for matrix converter and its simulation. *WSEAS Trans Circ Syst* 15(5):35–40
5. Bhavithra SK, Balasubramonian M (2016) Matrix converter fed induction motor drive. *Middle-East J Sci Res* 24(3):776–781
6. Abdalla AN, Ghoni R, Zakaria NF (2011) Simulation model of space vector modulated control matrix converter-fed induction motor. *J Appl Sci* 11(5):768–777
7. Lin WJ, Liu DL, Wu QX (2012) Matrix converter fed field-oriented control of permanent magnet synchronous motor. *Appl Mech Mater* 5(4):1–8
8. Sünter S, Clare JC (2000) Feedforward indirect vector control of a matrix converter-fed induction motor drive. *Int J Comput Math Electr Electron Eng* 19(4):974–986
9. Yakut YB, Sünter S, Özdemir M (2016) Simulation of matrix converter-fed permanent magnet synchronous motor with neural fuzzy controller. *Eur J Tech* 6(2):178–187
10. Yakut YB, Sünter S, Özdemir M (2017) Control method for driving two permanent magnet synchronous motors connected in parallel fed by a matrix converter. *Int Refereed J Eng Sci* 2(9):48–66



Visual Pose Estimation Based on the DenseNet Network

Lixin Lu^(✉) and Weicun Zhang

University of Science and Technology Beijing, Beijing 100083, China
lixinlu42@yeah.net

Abstract. An end-to-end neural network model based on DenseNet was designed to estimate the pose of the camera in this paper. The picture frame captured by the camera and the camera position (3-dimensional space coordinates) and pose (quaternion) corresponding to the picture frame are the inputs to the network model. Through the neural network model, the spatial structure information and the higher-layer features in the image are trained and learned, so that the network model finally outputs the 7-dimensional vector representing the camera position (3-dimensional space coordinates) and the pose (quaternion). Due to the pose estimation constraint of the network, the training effect of the model is guaranteed, and the pose estimation ability of the network is improved. The trained model is validated on the StMarysChurch Dataset. The experimental results show that the network model has good performances in accuracy and shorter training time.

Keywords: Deep learning · Neural network · DenseNet · Pose estimation · Regression network

1 Introduction

With the development of deep learning, it has been applied to solve related problems in more and more application fields. The great success of deep neural networks in the field of computer vision has provided new ideas for the study of visual localization. At present, there are mainly two methods of visual localization:

- Based on traditional geometric schemes. With the calibration camera model, the inter-frame estimation method based on sparse features or dense features restores relative translation and rotation by constantly tracking image features.
- End-to-end neural network model training scheme based on deep learning. Deep learning can directly obtain the mapping relationship between image and camera motion by training data, which can be extended to visual localization, providing new solutions and ways for visual localization problems.

Compared with the traditional inter-frame estimation method based on sparse features or dense features, the end-to-end neural network model method based on deep learning does not require feature extraction, nor feature matching and complex geometric operations, and no cumulative error. Each frame of new pictures will have the same mean error, which makes the method based on deep learning more intuitive and

concise, while the traditional geometric scheme will increase the cumulative error with the increase of time and distance.

In computer vision and image processing, a feature is a set of information related to a computing task, which depends on the specific application. Features may have special structures in the image, such as corners, edges, or block objects. An intuitive feature extraction method is to identify different image corner points and determine their correlation. In this case, the corner points are features what they call. The simple corner point still cannot meet the demand in the actual situation. So the researchers designed many more stable local image features such as SIFT, SUFT and ORB. Those fully consider various problems in the image transformation process, but also bring a large amount of calculation. Generally, the calculation on the CPU has an influence on real-time performance.

More and more research is based on the end-to-end deep learning model with the increasing demand for real-time requirements. For example, Kendall et al. [1] designed the network model PoseNet based on the deep neural network. The model is based on the GoogleNet [2] network model, and the final classification layer is changed to the linear regression fully connected layer of the 7-dimensional vector. A fully connected layer of size 2048 is inserted in front of the fully connected layer in order to form a localized feature vector, which enables subsequent linear regression operations, thereby changing the GoogleNet classification network to the regression network PoseNet. A pose estimation model end-to-end learning proposed by Qiao et al. [3] uses only image frames as input to merge optical flow extraction networks and pose estimation networks, and reconstructs optical flow fields using self-encoding networks to improve the robustness of optical flow characteristics. Valad et al. [4] proposed VLocNet, which is a CNN network for 6-DOF global pose regression and odometry estimation of continuous images. VLocNet designs a new loss function to use the relative pose information during training to constrain the search space to obtain a consistent pose estimate. DispNet which proposed by Mayer et al. [5] was modified on the structure of FlowNet [6]. In general, both FlowNet and DispNet are based on the optical flow method which is two or more images sequentially taken by a single camera along any orbit. The overall idea of DispNet is to include a contraction part and an expansion part. The contraction part is mainly composed of convolution, which is used to extract the higher-layer features of the two pictures in depth, and the expansion part intelligently restores the result to a high pixel feature map from coarse to fine. These methods are based on the modification of the older convolutional neural network, and deal with the corresponding image higher-layer features, the accuracy has reached the highest value of the old network, no modified experiments for the newer DenseNet. Therefore, this paper optimizes the network based on DenseNet and adds BN layer and Relu layer between blocks to optimize training results.

2 Position Estimation Network Design

2.1 Position Estimation Problem Description

After giving a single image frame $\mathbf{I} \in R^3$, we try to estimate the transformation of the camera's central pose p relative to the initial moment by training the deep neural network $F(\mathbf{I})$. The output of the network model is pose $p = [x, q] \in R^7$, where $x \in R^3$ is the 3-dimensional spatial displacement vector of the camera center, and $q \in R^4$ is the quaternion vector of the camera center point. The work of this paper is to use the deep learning model to learn the function $F(\mathbf{I})$.

2.2 Network Structure Design

The pose estimation problem is essentially a regression problem, the input training image frame is used to integrate the image features, and the 6-degree-of-freedom pose vector fitting is performed in high dimension. The basic idea of the network in this paper is to use the convolutional layer to extract higher-layer features from shallow to deep, map the original data to the hidden layer feature space, integrate the high-dimensional image information with the global average pooling layer and fully connected layer, map the learned distributed feature representation to the sample mark space, and then the logistic regression of the pose information is achieved.

The pose estimation network design structure is shown in Fig. 1. It draws on the basic ideas of DenseNet to achieve its own regression effect.

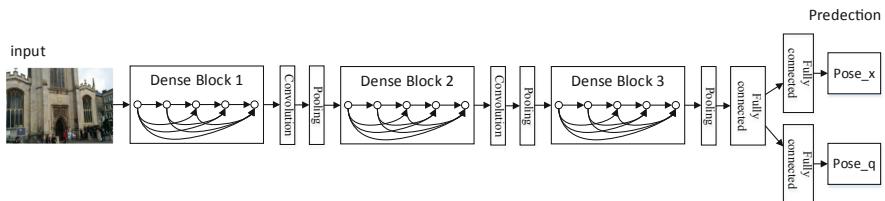


Fig. 1. Position estimation network structure design.

The pose estimation network structure has three structural blocks, each of which contains 12 layers of convolution, which are composed of 6 pairs of 1×1 and 3×3 convolutional layers, plus a 1×1 convolutional layer and a final fully connected layer between the blocks are added, and a total of 40 layers are formed.

The core idea of the DenseNet network is to connect all the layers in the network block, that is, the input of one layer is obtained by splicing the output of all layers before the layer, rather than a normal neural network, only adjacent layers have input and output connections. However, the size of this splicing will be very large, which will lead to a very large amount of training, resulting in a serious decline in training speed. DenseNet uses the bottleneck operation to solve this problem by adding a 1×1 convolution layer in front of each 3×3 convolution layer in the block to reduce the

input dimensions of each layer. When a 1×1 convolution kernel is used to reduce the input dimensions of each layer in the block, the model parameters will be greatly reduced, the corresponding calculation amount will be reduced, and the training speed will naturally increase rapidly. Although the internal dimensions of the Block are degraded, the output dimension of the Block is also large. Therefore, the transition layer structure is designed in the Dense. The transition layer structure is also a 1×1 convolution layer, which is used to reduce the input dimension between each Block, further reducing the amount of calculation and improving the training speed.

2.3 Loss Function

In a general classification network, the network only needs to output a probability vector whose number of dimensions and number of sample categories are equal. The softmax function is usually used to convert the output of the network into a probability vector, and then the loss of the network is calculated by the cross entropy loss function. DenseNet network is this classification network. The network designed in this paper requires the pose information to be predicted. It is necessary to predict the continuous pose information instead of the predicted probability vector. Therefore, the network proposed in this paper is based on DenseNet to modify softmax to a fully connected layer to return pose information, which belongs to the regression network. The pose information is a multi-dimensional vector, the distance between normal vectors represented by Euclidean distance usually. The L2 norm of the pose information in this paper is used to express the loss function as follows:

$$\text{loss}(I) = \| \hat{x} - x \|_2 + \beta \| \hat{q} - \frac{q}{\| q \|} \|_2 \quad (1)$$

where β is a scaling factor to balance the position information x and the attitude quaternion q to make the final convergence value of the loss error smaller.

After experimental tests, the β value is in the range of 500–2000, and the training effect is better. The spin set is on the unit sphere of the quaternion space. However, the Euclidean loss function does not keep q on the unit sphere. However, we found that during training, q became very close to \hat{q} so that the difference between the spherical distance and the Euclidean distance became irrelevant. For the sake of simplicity, and to avoid impediments and unnecessary constraint optimization, this paper chooses to omit the spherical constraint.

When verifying the training error, it is necessary to convert the attitude quaternion into the rotation angle information. In general, the conversion formula of the attitude quaternion $q = [q_0, q_1, q_2, q_3]$ to the Euler angle $[\alpha, \beta, \gamma]$ is as follows:

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} \arctan \frac{2(q_0q_1 + q_2q_3)}{1 - 2(q_1^2 + q_2^2)} \\ \arcsin 2(q_0q_2 - q_1q_3) \\ \arctan \frac{2(q_0q_3 + q_1q_2)}{1 - 2(q_2^2 + q_3^2)} \end{bmatrix} \quad (2)$$

where α, β, γ is the angle at which the X-Y-Z axis rotates around the fixed coordinate axis.

However, our error angle is a 3-dimensional rotation angle, not the angle mapped to each coordinate axis, so we use the rotation angle error formula as follows:

$$\Delta\theta = 2 * \arccos\left(\frac{\hat{p}}{\|\hat{p}\|_2} \otimes \frac{p}{\|p\|_2}\right) * \frac{180}{\pi} \quad (3)$$

where \otimes means that all corresponding elements of the vector are multiplied and summed.

2.4 Network Training and Parameter Setting

The end-to-end pose estimation network proposed in this paper adopts DenseNet as the prototype, so we can load the pre-training model provided by DenseNet, initialize the pose estimation model parameters through Transfer Learning, and initialize the parameters of added fully connected layer using Xavier algorithm. Finally, a 7-dimensional pose vector is obtained.

We performed different scale scaling experiments on the original image for training and testing. Before downsampling to 256 pixels, enlarging the input is equivalent to cropping the input, which increases the spatial resolution of the input pixel, but does not increase the pose estimation performance. This suggests that context and view are more important than resolution for pose estimation. However, in order to reduce the amount of calculation and reduce the GPU memory usage, we downsample the original image to 256 pixels, and randomly crop it on this basis. Experiments have shown that random cropping does not affect camera pose. We construct a pose estimation network and initialize network parameters by fine-tune part of DenseNet's network layer and parameters. The Momentum optimization algorithm is used to optimize the network. The learning rate is 0.001 and is reduced to one tenth of every 50 iterations. The weight decay is 0.0001, the nesterov momentum parameter is 0.9, the β in the loss function is 2000, and the other parameters use the default values.

3 Dataset

The StMarysChurch Dataset was produced by a series of picture frames with pose information collected by the Cambridge University Research Group handheld Apple smartphone for the performance of the visual algorithm network in pose estimation. The pose information automatically generated by the sfm [7] technique is used in the acquisition process, which saves a lot of time and effort compared to manual labeling. Figure 2 is a picture of the StMarysChurch Dataset. The environment is complex and the texture features are general. Using such a picture as a data set can enhance the network generalization performance and robust performance, and make the network model test and application results more stable.

The StMarysChurch Dataset can be downloaded from the official website link. StMarysChurch is a church at Cambridge University. The data set is made by



Fig. 2. StMarysChurch Dataset.

researchers taking pictures from outside the church and taking pictures. This dataset includes a total of 2019 images, a train set of 1488 images, and a test set of 531 images. The annotation data file contains two files: `dataset_train.txt` and `dataset_test.txt`. Each file contains storage address information of each picture and three spatial position data with respect to the initial time XYZ axis and angle rotation data expressed by four pose quaternions. When training the network model, we read the image in the annotation file and the pose label corresponding to the image, and use the image as the input of the network, and the pose label as the input of the network loss function.

4 Training and Test Results

The GPU model used in our training is GeForce GTX 1060 6G version, which is a mid-range graphics card and can run this program well. It takes about 8 h to run 150 iterations. The training time is within the tolerance range, and the 6G memory supports the memory usage when the program is running.

The loss value curve for each epoch is shown in Fig. 3. Since the loss value of each epoch takes the mean value, the curve is much smoother than the mean of each batch. With each learning rate decreasing, the loss curve also drops significantly, which means that the decay of the learning rate has a significant effect on the network convergence in a certain loss interval, and then to the second learning rate decay (100 iterations). The decrease in the loss value is not as obvious as the first time, indicating that the learning rate decay is less effective in the interval where the loss value is lower.

Figure 4 shows the loss verification curve of pose x , we set a verification test every two iterations to see the convergence of the validation set loss. In this way, if the loss convergence of the training process is good and the loss convergence of the testing process is not good, the parameters can be modified immediately and retrained without waiting for a long time to complete the training process and then test the convergence effect of the loss on the new test set. It turns out that the loss curve on the validation set is very flat and finally converges to 1.88, which is similar to the PoseNet results, but

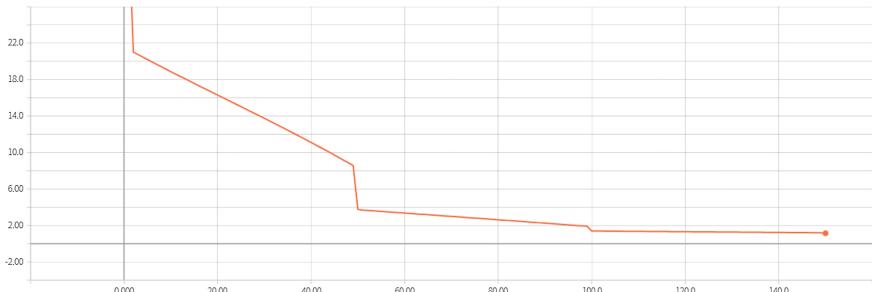


Fig. 3. Loss curve for each epoch of training.

DenseNet is faster than PoseNet based on VGG, with fewer parameters, and has a shorter time to estimates the pose of a picture.

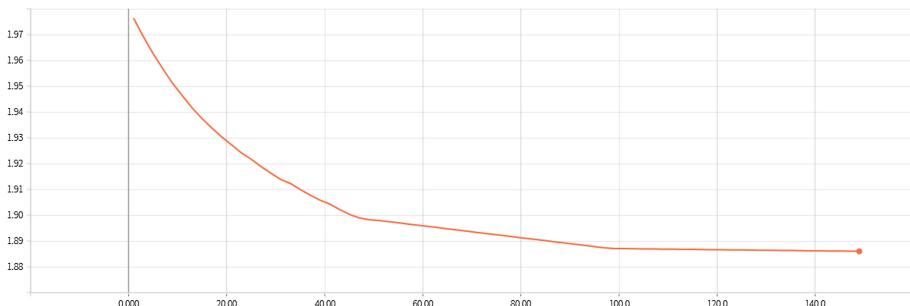


Fig. 4. Verification loss curve of pose x every two epoch.

The loss verification curve of the pose q is drawn in Fig. 5 shows. It is very gentle, and finally converges to around 2.55, which is 0.4 lower than the result of PoseNet. It



Fig. 5. Verification loss curve of pose q every two epoch.

can be seen that the regression effect of DenseNet in the pose q is better than that of PoseNet.

Overall, compared to PoseNet, DenseNet's training speed is increased, the amount of parameters is reduced, the time for calculating each picture is shortened, the accuracy of the pose is hardly improved, and the accuracy of the pose is increased by 0.4° .

5 Conclusions

We present a pose estimation network based on DenseNet to modify softmax to a fully connected layer to output pose information $p = [x, q]$. It improves the accuracy and training speed of PoseNet, but not high enough in practical applications. In the future, our work will be focused on multi-network pose estimation, that is, two or more main networks are used to extract high-dimensional feature information, and in the final output layer, the outputs of multiple networks are spliced to improve the accuracy of the network in the pose.

Acknowledgments. This work was supported by National Natural Science Foundation of China (grant numbers 61520106010, 61741302).

References

1. Kendall A, Grimes M, Cipolla R (2015) PoseNet: a convolutional network for real-time 6-DOF camera relocalization. In: CONFERENCE 2015, ICCV, Santiago, Chile
2. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: CONFERENCE 2015, CVPR, Boston, USA
3. Qiao Q (2018) Research on visual ego-motion estimation with convolutional neural network. University of Science and Technology of China, Hefei (2018)
4. Valada A, Radwan N, Burgard W (2018) Deep auxiliary learning for visual localization and odometry. In: CONFERENCE 2018, ICRA, Brisbane, Australia
5. Mayer N, Ilg E, Häusser P, Fischer P (2015) A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In: CONFERENCE 2015, CVPR, Boston, USA
6. Dosovitskiy A, Fischer P, Ilg E, Hausser P, Hazırbaş C, Golkov V (2015) FlowNet: learning optical flow with convolutional networks. In: CONFERENCE 2015, ICCV, Santiago, Chile
7. Wu C (2015) Towards linear-time incremental structure from motion. In: CONFERENCE 2015, 3DV, Seattle WA, USA



Kinematics Numerical Inverse Solution Algorithm for Arm Type AFP Machine Based on D-H Modeling

Kanzhong Yao and Yanjuan Zhu^(✉)

School of Aerospace Engineering and Applied Mechanics, Tongji University,
Shanghai 200092, China
zhu_yj@tongji.edu.cn

Abstract. A solution method for inverse kinematics of automated fiber placement (AFP) trajectory is proposed. The D-H method is used to establish the kinematical model of the AFP head, based on this kinematical model, the mathematical model of the forward solution of the AFP trajectory is constructed; The matrix equation is simplified to general polynomial equation group by using the tangent half-angle substitution method, the correct solution is automatically selected for the iteration according to the joint function of the AFP machine, and an effective algorithm for the reverse solution of the trajectory is obtained. Numerical simulation of the algorithm is carried out. The verification results show that the algorithm is feasible and effective.

Keywords: AFP · Kinematical model · Inverse solution algorithm

1 Introduction

Automated fiber placement (AFP) technology is one of the key technologies in aircraft manufacturing, and it is widely used in the automatic molding manufacturing of composite parts. The kinematics solution of the placement head is the core of the AFP technology, and it is of great significance to study its solving algorithm.

In the research of the kinematics algorithm for AFP machine, the inverse kinematics solution of the automatic fiber laying robot was discussed in the reference [1], and an analytical method of inverse kinematics solution was proposed, the method theoretically proved the feasibility of the kinematics solution of the arm type AFP machine [1]; The inverse transform kinematics optimization algorithm proposed in reference [2] was based on guiding fiber cutting, adding joint properties of machine arm into the optimization and improving the placement precision of AFP technology [2]; In terms of measuring the feasibility of the algorithm, a research based on Riemannian manifold introduced the space volume element evaluation method as the evaluation method of algorithm validity into the field of AFP kinematics algorithm, which provided a reference for subsequent research [3]. In addition, in a study [4] in 2017, researchers proposed a “Semi-offline trajectory synchronized algorithm”, which

used two parallel robots to work together to avoid the trajectory defects in AFP process, greatly improved the placement quality of the original method, and the verification of the algorithm was completed based on the specific placement robot; Di Vito et al. simulated several existing kinematics algorithms. Their research report [5] pointed out that the four mainstream algorithms [6–8] did not fully consider the joint constraint problem; exceeding constraints during continuous placement may result in failure.

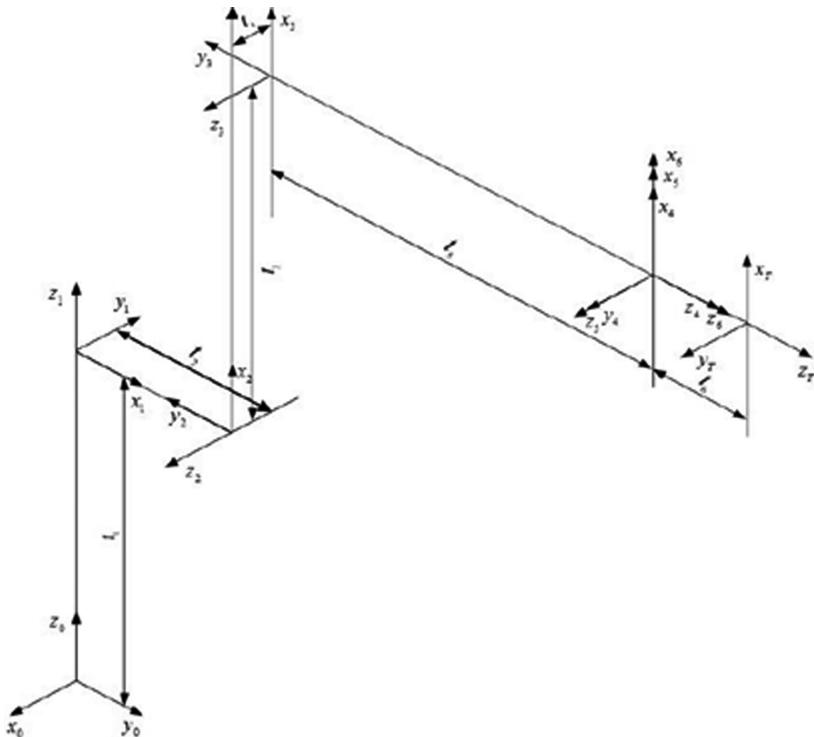
In the first part of the paper, a kinematical model and a forward-solving mathematical model are established for a typical six-degree-of-freedom robotic arm automatic laying machine. In the second part, the inverse kinematics model is proposed, using tangent half-angle substitution simplifies the matrix equation to general polynomial equation group, in the iterative process, the correct solution is automatically selected according to the joint function of the AFP machine, and a solution algorithm is proposed. In the third part, a placement model obtained in CATIA is imported into matlab environment to run the simulation. The verification results show that the algorithm is feasible and effective. The new method can be a helpful tool for subsequent researches.

2 Kinematical Modeling Based on D-H Method

The six-degree-of-freedom AFP arm type robot can be accurately modeled by D-H modeling. The D-H modeling steps are as follows:

1. **Establishment of the pedestal coordinate system:** establish a coordinate system (X_0, Y_0, Z_0) by right-hand rule on the pedestal, and determine the motion axis Z_1 of the next joint adjacent to the pedestal, define Z_0 parallel to Z_1 direction.
2. **Establishment of an independent coordinate system for the transition link Z_i :** align the axis Z_i with the axis of motion of the joint i, and the direction of X_i axis are along the common normal line of Z_i and Z_{i-1} , pointing away from the direction of the Z_{i-1} . When $a_i = 0$, $X_i = (Z_{i+1} \times Z_i) / \|Z_{i-1} \times Z_i\|$; Y_i axis is determined by the right-hand rule; the coordinate system origin O_i of the link i is located at the point of intersection of the common normal line of joints i, $i + 1$ and the axis of joint i; when Z_i is parallel to Z_{i+1} , the origin is taken at $d_{i+1} = 0$, where d_{i+1} is the offset of the link $i + 1$.
3. **Establishment of the end effector coordinate system O_T .**

The established AFP machine linkage coordinate system is showed in Fig. 1 [9]:

**Fig. 1.** Linkage coordinate system

The mechanical arm coordinate link parameters are determined in Table 1 [10]:

Table 1. Link parameters

| Link i | Variable θ_i | α_{i-1} | a_{i-1} | d_i |
|--------|---------------------|----------------|-----------|-------|
| 1 | θ_1 | 0° | 0 | t_1 |
| 2 | θ_2 | 90° | t_2 | 0 |
| 3 | θ_3 | 0° | t_3 | t_4 |
| 4 | θ_4 | 90° | 0 | t_5 |
| 5 | θ_5 | -90° | 0 | 0 |
| 6 | θ_6 | 90° | 0 | 0 |

Where α_i is the angle between two adjacent link axes, called Link Twist Angle; a_i is distance between common normal line of two links, called Link Length; d_i is the relative position of two adjacent links, called Link Distance; θ_i is the angle between two link normal, called Link Angle.

This linkage system is the kinematical coordinate system based on the D-H method of the AFP robot arm.

3 Mathematical Model Construction and Kinematical Solution

3.1 Positive Solution

Positive Solution: Given the link parameters of the AFP robot and the values of the joint variables to solve the posture of the robot end effector. In order to get enough parameters for the reverse solution and provide a reference for the feasibility of the algorithm, the positive solution is proposed based on D-H coordinate transformation.

The position and attitude of each joint point in the three-dimensional space are represented by the pose matrix T_{0x} dominated by the joint angle, and T_{06} represents the pose of the end effector of the AFP head. The pose transformation matrix between each joint point is $T_{(i-1)i}$, and the transformation matrix from link 1 to link 6 can be expressed as

$$T_{06} = T_{01}T_{12}T_{23}T_{34}T_{45}T_{56} \quad (1)$$

The coordinate transformation matrix is determined by the joint angle of the joint of the robot arm. According to the principle of homogeneous coordinate transformation, it can be obtained as

$$T_{(i-1)i} = \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) & 0 & a_{i-1} \\ \cos(a_{i-1})\sin(\theta_i) & \cos(a_{i-1})\cos(\theta_i) & -\sin(a_{i-1}) & -d_i \sin(a_{i-1}) \\ \sin(a_{i-1})\sin(\theta_i) & \sin(a_{i-1})\cos(\theta_i) & \cos(a_{i-1}) & d_i \cos(a_{i-1}) \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Simultaneous Eqs. (1) and (2), and bring the parameters of Table 1 into the equation, we can get the analytical form of pose matrix (4×4) of the end effector as follows:

$$T_{06} = \begin{bmatrix} n_x & a_x & o_x & p_x \\ n_y & a_y & o_y & p_y \\ n_z & a_z & o_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

Where, n_x, n_y, n_z are the component of x-axis positive direction unit vectors of the last joint coordinate system $x_6y_6z_6$ on the x, y and z axes of the pedestal coordinate system $x_0y_0z_0$.

a_x, a_y, a_z are the component of y-axis positive direction unit vectors of the last joint coordinate system $x_6y_6z_6$ on the x, y and z axes of the pedestal coordinate system $x_0y_0z_0$.

o_x, o_y, o_z are the component of z-axis positive direction unit vectors of the last joint coordinate system $x_6y_6z_6$ on the x, y and z axes of the pedestal coordinate system $x_0y_0z_0$.

p_x, p_y, p_z are the x, y and z coordinate value of the coordinate origin of the last joint coordinate system $x_6y_6z_6$ in the pedestal coordinate system $x_0y_0z_0$.

For any given set of $\theta_i(i = 1, 2, \dots, 6)$ initial values, the pose matrix of the end link can be uniquely determined.

3.2 Inverse Solution

The inverse solution algorithm is the core issue of the discussion, that is, the pose of the known end effector is inversely solved for the values of the various joint variables of the robot. The end effector pose matrix T_{06} described in Sect. 3.1 is obtained to solve a set of $\theta_i(i = 1, 2, \dots, 6)$. The solution steps are as follows:

Step One: by inversely transforming Eq. (1), following equation is obtained

$$(T_{01})^{-1}T_{06} = T_{12}T_{23}T_{34}T_{45}T_{56} \quad (4)$$

Due to space limitation, the specific expressions of the matrix elements are not listed here, let the two matrices' (2, 4), (1, 4), (3, 4) to be equal and the equations are obtained as follows

$$\begin{cases} -\sin(\theta_1)p_x + \cos(\theta_1)p_y = t_4 \\ t_2 + t_5(\cos(\theta_2)\sin(\theta_3) + \cos(\theta_3)\sin(\theta_2)) + t_3\cos(\theta_2) = \cos(\theta_1)p_x + \sin(\theta_1)p_y \\ t_3\sin(\theta_2) - t_5(\cos(\theta_2)\cos(\theta_3) - \sin(\theta_2)\sin(\theta_3)) = p_z - t_1 \end{cases} \quad (5)$$

Similarly, inverse transformation of Eq. (4) yields Eq. (6)

$$(T_{23.})^{-1}(T_{12})^{-1}(T_{01})^{-1}T_{06} = T_{34}T_{45}T_{56} \quad (6)$$

Let the left and right matrices' (1, 3) and (3, 3) elements to be equal and the following equations can be obtained:

$$\begin{cases} a_x(\cos(\theta_1)\cos(\theta_2)\cos(\theta_3) - \sin(\theta_2)\sin(\theta_3)) + a_z(\cos(\theta_2)\cos(\theta_3) - \sin(\theta_2)\cos(\theta_3)) \\ -a_y(\sin(\theta_1)\sin(\theta_2)\sin(\theta_3) - \sin(\theta_1)\cos(\theta_2)\cos(\theta_3)) - \sin(\theta_5)\cos(\theta_4) = 0 \\ a_x\sin(\theta_1) - a_y\cos(\theta_1) - \sin(\theta_4)\sin(\theta_5) = 0 \end{cases} \quad (7)$$

By transforming Eq. (6), the last inverse transformation equation can be written as

$$(T_{45})^{-1}(T_{34})^{-1}(T_{23.})^{-1}(T_{12})^{-1}(T_{01})^{-1}T_{06} = T_{56} \quad (8)$$

Let the left and right matrices' (2, 1) elements to be equal, and we get the following equation:

$$\sin(\theta_6)\cos(\theta_4) + \sin(\theta_4)\cos(\theta_5)\cos(\theta_6) + n_y\cos(\theta_1) - n_x\sin(\theta_1) = 0 \quad (9)$$

Step Two: using tangent half-angle substitution, Eqs. (5), (7), and (9) are transformed into polynomial equations. The substitution rules are as follows:

$$\begin{cases} \cos(\theta_i) = \frac{1-x_i^2}{1+x_i^2} \\ \sin(\theta_i) = \frac{2x_i}{1+x_i^2} \end{cases} \quad (10)$$

Substituting (10) into Eqs. (5), (7), and (9), the mathematical model of the inverse solution is obtained:

$$\begin{cases} f(x_1) = 0 \\ h(x_1, x_2, x_3) = 0 \\ k(x_1, x_2, x_3, x_4) = 0 \\ j(x_4, x_5) = 0 \\ g(x_1, x_2, x_3, x_4, x_5, x_6) = 0 \end{cases} \quad (11)$$

Step Three: the $x_i(i = 1, 2, \dots, 6)$ obtained by solving the Eq. (11) is substituted into the Eq. (10) to obtain the desired joint angle $\theta_i(i = 1, 2, \dots, 6)$, notably direct solution may result in two sets of joint angles $\theta_i(i = 1, 2, \dots, 6)$. In actual solution process, a initial value needs to be set for x_1 , and the error value is discarded before solving x_{2-6} .

4 Numerical Solution

4.1 Algorithm Flowchart

The flowchart of the algorithm is showed in Fig. 2.

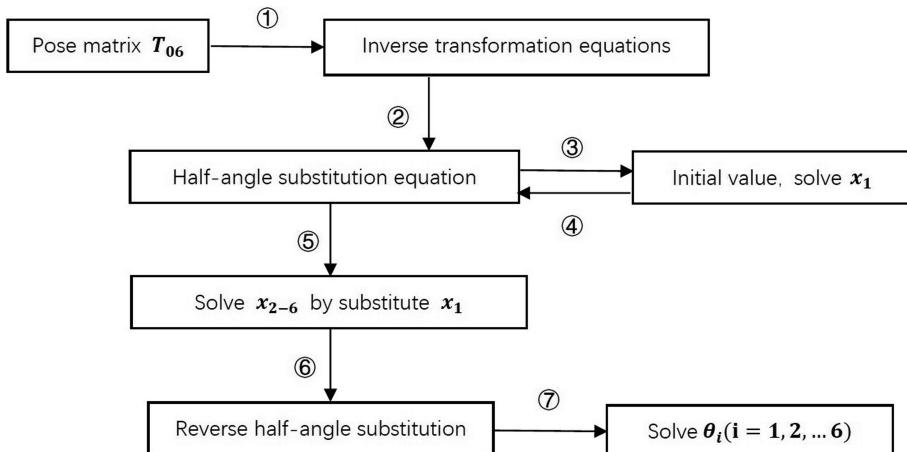


Fig. 2. Algorithm flowchart

In order to verify the feasibility of the algorithm, a target surface is established in catia, and a planned trajectory to be laid is obtained on the surface. Data points are extracted every 1 mm on the trajectory, and the data point pose matrix is used as algorithm input. The given trajectory is shown in Fig. 3:

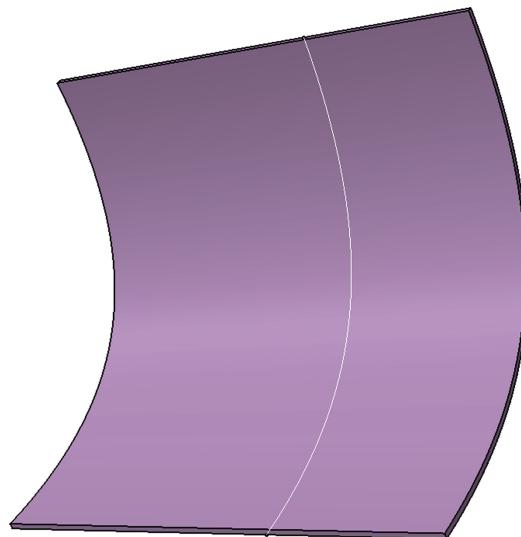


Fig. 3. The given trajectory

4.2 Result Output

It can be seen from Fig. 4 that although it is below the different coordinate system, curves showed in Fig. 4(a) and (b) are the same curve. In matlab environment, the algorithm feasibility verification succeeds.

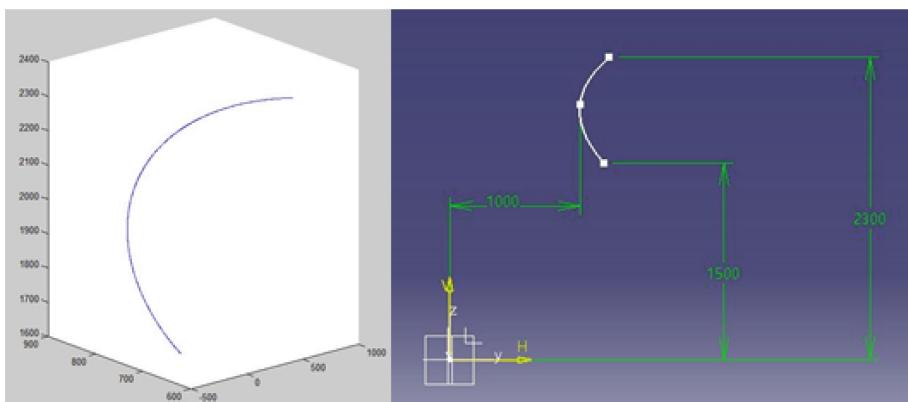


Fig. 4. Comparison of simulation results with theoretical values

5 Conclusion

The mechanical model of the AFP robot of the composite material is constructed. The mathematical model based on the matrix equation is used to solve the positive and inverse kinematical solutions of the AFP robot arm, and the function of the solution platform is realized. Based on the model, an inverse kinematical algorithm for AFP trajectory is proposed. The method of tangent half-angle substitution is used to transform the matrix equation into a polynomial equation group, which simplifies the solution process. For the problem that multiple solutions may occur in the inverse solution, the AFP robot arm joint attributes are considered in the algorithm to automatically select the correct solution, effectively eliminating the interference solution. The effectiveness and feasibility of the algorithm have also been verified.

References

1. Gong C, You Y (2006) Analytic method of automatic tow placement robot inverse kinematics problem. *J Mech Eng Autom* 2006(01):1–3
2. Lopes CS, Gürdal Z, Camanho PP (2008) Variable-stiffness composite panels: buckling and first-ply failure improvements over straight-fiber laminates. *Comput Struct* 86:897–907
3. Ge H, Ge X, Zhao D (2013) Study of robotic manipulators' kinematics manipulability. *J Zhengzhou Univ (Eng Sci)* 34(02):121–124
4. Zhang X, Xie W, Hoa SV (2018) Semi-offline trajectory synchronized algorithm of the cooperative automated fiber placement system. *Robot Comput Integr Manuf* 51:53–62
5. Di Vito D, Natale C, Antonelli G (2017) A comparison of damped least squares algorithms for inverse kinematics of robot manipulators. *IFAC-PapersOnLine* 50:6869–6874
6. Baerlocher P (2001) Inverse kinematics techniques for the interactive posture control of articulated figures. *Ecole Polytechnique Federal De Lausanne*
7. Deo AS, Walker ID (1995) Overview of damped least-squares methods for inverse kinematics of robot manipulators. *J Intell Rob Syst* 14(1):43–68
8. Sugihara T (2011) Solvability-unconcerned inverse kinematics by the Levenberg–Marquardt method. *IEEE Trans Rob* 27(5):984–991
9. Singh S (2016) Singla E (2016) Modeling and simulation of basic robotic configurations using D-H parameters based adaptive modules. *Perspect Sci* 8(C):513–515
10. Wang N, Liu Y, Xiao J (2008) Fiber-placement path design for composite structures in pipiform. *J Comput-Aided Des Comput Graph* 20(02):228–233



General Correntropy Based DOA Estimation for Coherent Source with Impulsive Noise

Sheng Guan¹, Jin Chen¹⁽⁾, Kaikai Li¹, and Yuqi Zhao²

¹ College of Electronic and Communication Engineering,
Tianjin Normal University, Tianjin 300387, China

cjwoods@163.com

² State Grid BaiCheng Power Supply Company, Baicheng 137000, China

Abstract. This paper reviews and compares three different decorrelation methods for Direction-of-Arrival (DOA) estimation of coherent sources, these include Real-Value (RV), Spatial Smoothing (SS), and Modified Singular Value Decomposition (MSVD). By applying the Correntropy Based Correlation (CRCO) rebuild covariance matrix of array output so as to control the impulsive noise, which is modeled as complex symmetric α -stable distribution. three algorithms were defined, that are CRCO_RV, CRCO_SS, and CRCO_MSVD. The simulation experiments comparing the performance of these methods, highlighting their commonalities and differences.

Keywords: Complex isotropic symmetric alpha-stable distribution · Coherent signal · Correntropy

1 Introduction

In multipath environments, there are coherent signals that arriving at arrays, including co-frequency interference and multipath propagation signals caused by reflection and refraction of background objects, research on decorrelation algorithm is the bottleneck of high-resolution algorithm theory into practice, many decorrelated algorithms have been proposed to solve the coherency problem.

Singular value decomposition (SVD) was proposed for finding DOA of coherent signals in 1987 [1], to improve the resolution of the SVD algorithm, an improved SVD algorithm was presented, namely, MSVD [2]. A forward/backward spatial smoothing (SS) method was proposed [3], SS method is a classical and effective one to solve the coherent sources localization issue. Unitary ESPRIT algorithm was developed by reducing the computational complexity for engineering application [4], this real-value (unitary) has recently drawn considerable attention in the literature [5].

In this letter, we proposed several algorithms for finding DOAs when coherent signals in impulsive noise, we first construct the covariance matrix of array output by the correntropy based on correlation (CRCO), then decorrelation by the modified singular value decomposition (MSVD), forward/backward spatial smoothing (SS) method and real-value (RV) method. Finally, the MUSIC algorithm was applied to DOA estimation and analysis of their performances.

Notation: $\text{diag}(A)$ denotes a diagonal matrix whose A, $\text{flipud}(\cdot)$ denotes up and down operator, $[\cdot]^*$ is a conjugate operator, $[\cdot]^H$ is conjugate-transpose, $E[\cdot]$ stand for expectation operator.

2 Correntropy Based Correlation and Generation of Complex Isotropic SuS Random Variables

2.1 Correntropy Based Correlation

Consider the P narrowband far-field signals $s_1(t), \dots, s_p(t)$ with different DOA $\theta_1, \dots, \theta_p$ impinging on the uniform linear array (ULA) with M sensors under additive impulsive noise $n_1(t), \dots, n_M(t)$. The ULA output $x_1(t), \dots, x_M(t)$ can be modeled as

$$X(t) = AS(t) + N(t) \quad (1)$$

Where

$$\begin{aligned} X(t) &= [x_1(t), x_2(t), \dots, x_M(t)]^T \\ A &= [a(\theta_1), a(\theta_2), \dots, a(\theta_p)]^T \\ S(t) &= [s_1(t), s_2(t), \dots, s_p(t)]^T \\ N(t) &= [n_1(t), n_2(t), \dots, n_M(t)]^T \end{aligned}$$

$a(\theta_p)$, $p = 1, 2, \dots, P$ are the steering vector which interspace half of the wavelength, it can be written as

$$a(\theta_p) = [1, e^{-j\pi \sin(\theta_p)}, \dots, e^{-j(M-1)\pi \sin(\theta_p)}]^T \quad (2)$$

The correntropy based correlation covariation matrix R_{CRCO} is defined as [6]

$$R_{CRCO} = E[\exp(-\frac{(X - uY)^2}{2\sigma^2})XY] \quad (3)$$

Where $\mu \in (0, 1)$ is a suppression parameter that makes sure the correntropy based autocorrelation to be finite, σ is kernel size that controls the bandwidth of the CIM “mix norm” [7], $\sigma = 1.4\hat{\delta}_s$, $\mu = 0.5$ according to [6], $\hat{\delta}_s$ is the estimated variance of the noise-free signal $s(t)$.

2.2 Generation of Complex Isotropic SuS Random Variables

Assume X_1 and X_2 are jointly SuS random variables and complex random variable $X = X_1 + jX_2$ with characteristic function [8, 9]

$$\begin{aligned}
\phi(w) &= E\{\exp[j * \text{real}(wX^*)]\} \\
&= E\{\exp[j(w_1X_1 + w_2X_2)]\} \\
&= \exp\left[-\int_{S_2} |w_1x_1 + w_2x_2|^\alpha d\Gamma_{x_1,x_2}(x_1, x_2)\right]
\end{aligned} \tag{4}$$

Where $\text{real}[\cdot]$ is the real part operator, $w = w_1 + jw_2$, and Γ_{x_1,x_2} is a spectral measure of the random variable X , and it is symmetric on the unit sphere S_2 . A complex random variable $X = X_1 + jX_2$ is isotropic *SaS* if Γ_{x_1,x_2} has a uniform spectral measure. This class of distribution is best defined by its characteristic function

$$\varphi(w) = \exp(-\gamma|w_1x_1 + w_2x_2|^\alpha) \tag{5}$$

Where γ is the dispersion parameter. A method for the generation of complex isotropic *SaS* random variables was given in letter [8] (Fig. 1).

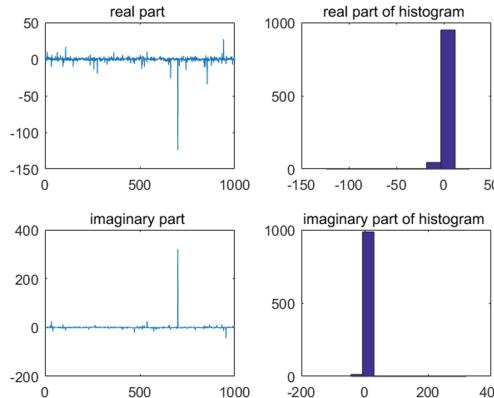


Fig. 1. Complex isotropic *SaS* random variables in time-domain $\alpha = 1.4$.

If there exist two zero-mean Gaussian variables $G1$, $G2$, and real *SaS* random variables $A \in \{0.5\alpha, 1, \cos^2(0.25\alpha\pi)\}$, 0.5α is the characteristic exponent, 1 is skewness, $\cos^2(0.25\alpha\pi)$ is dispersion. Therefore, the complex isotropic *SaS* random $X = X_1 + jX_2$ variables can be written as

$$X = A^{0.5}(G1 + jG2) \tag{6}$$

3 Decorrelation Algorithm Based on Correntropy

3.1 CRCO_RV Algorithm

Real-value (Unitary) algorithm is put the complex covariance matrix of the array output into a real-valued matrix by introducing a transformation matrix. Firstly, forward/backward spatial smoothing was applied to data matrix R_{CRCO} , therefor, coherent sources can be decorrelation by using this method.

$$R_{fb} = \frac{1}{2}(R_{CRCO} + J_M R_{CRCO}^* J_M) \quad (7)$$

Where J_M is the exchange matrix with ones on its antidiagonal and zeros elsewhere.

$$Q = \begin{cases} \frac{1}{\sqrt{2}} \begin{bmatrix} I_{\frac{M}{2}} & I_{\frac{M}{2}} \\ J_{\frac{M}{2}} & -J_{\frac{M}{2}} \end{bmatrix} & M = \text{even} \\ \frac{1}{\sqrt{2}} \begin{bmatrix} I_{\frac{M-1}{2}} & 0 & I_{\frac{M-1}{2}} \\ 0 & \sqrt{2} & 0 \\ J_{\frac{M-1}{2}} & 0 & -J_{\frac{M-1}{2}} \end{bmatrix} & M = \text{odd} \end{cases} \quad (8)$$

Where $I_{M/2}$ denote the $M/2 \times M/2$ identity matrix, note that Q is satisfied with $QQ^H = I$. and the real-valued symmetrical matrix can be achieved via a unitary transformation

$$R_{real} = Q_M^H R_{fb} Q_M \quad (9)$$

Finally, by conducting eigenvalue decomposition (EVD) on R_{real} , we can obtain the noise subspace U_n , the spectrum of conventional MUSIC algorithm [10] can be expressed as

$$P_\theta = \frac{1}{a(\theta)^H (QU_n)(QU_n)^H a(\theta)} \quad (10)$$

3.2 CRCO_SS Algorithm

The idea of the SS algorithm is to divide the ULA into overlapping forward and backward subarrays, the rank of covariance is restored to the number of sources so as to decorrelates. Figure 2 presents the diagram of SS algorithm, d is the number of elements subarray, L is the number of subarrays, $d = M - L + 1$.

The forward and backward averaged spatially smoothed covariance matrix R_f and R_b can be defined as

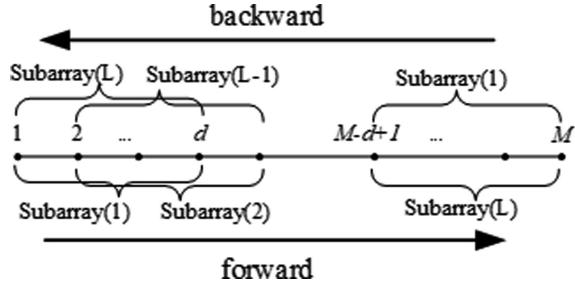


Fig. 2. Diagram of spatial smoothing

$$\begin{aligned}
 R_f &= \frac{1}{L} \sum_{l=1}^L R_l^f \\
 &= \frac{1}{L} \sum_{l=1}^L E[X_l^f(t) X_l^f(t)] \tag{11}
 \end{aligned}$$

$$\begin{aligned}
 x_l^f &= [x_l(t) \quad x_{l+1}(t) \quad \dots \quad x_{l+d-1}(t)] \\
 &= A_l D s(t) + n_l(t) \quad 1 \leq l \leq L \\
 &= A_l \text{diag}[e^{j\pi \sin(\theta_1)}, e^{j\pi \sin(\theta_2)}, \dots, e^{j\pi \sin(\theta_p)}] s(t) + n_l(t) \tag{12}
 \end{aligned}$$

Where $x_l^f(t)$ is the output of subarray-l, A_l is $d \times 1$ steering vector.

$$\begin{aligned}
 R_b &= \frac{1}{L} \sum_{l=1}^L R_l^b \\
 &= \frac{1}{L} \sum_{l=1}^L E[X_l^b(t) X_l^b(t)] \tag{13}
 \end{aligned}$$

$$\begin{aligned}
 x_l^b &= [x_{M-d-l+1}^*(t) \quad x_{M-d-l+2}^*(t) \quad \dots \quad x_{M-l+2}^*(t)] \\
 &= \text{flipud}(A_l D s(t) + n_l(t))^* \quad 1 \leq l \leq L \\
 &= \text{flipud}(A_l \text{diag}[e^{j\pi \sin(\theta_1)}, e^{j\pi \sin(\theta_2)}, \dots, e^{j\pi \sin(\theta_p)}] s(t) + n_l(t))^* \tag{14}
 \end{aligned}$$

Then the forward/backward spatial smoothed covariance matrix is defined as

$$R^{fb} = \frac{R^f + R^b}{2} \tag{15}$$

Finally, by performing singular decomposition to R^{fb} and the noise subspace U_n can be got, estimate DOAs of the signal by MUSIC algorithm.

3.3 CRCO_MSVD Algorithm

To begin with, by performing modified singular value decomposition (SVD) on R_{CRCO} , we can obtain noise subspace $U_n = [e_1, \dots, e_M]$, the eigenvector corresponding to the largest eigenvalue of the covariance matrix R_{CRCO} is a linear combination of all sources' steering vector.

$$e_1 = AT = \sum_{p=1}^P t_p a(\theta_p) \quad (16)$$

Where $e_1 = [e_{11}, \dots, e_{1k}, \dots, e_{1M}]$ is eigenvector corresponding to the largest eigenvalue and $T = [t_1, \dots, t_M]$ denote the linear combination factor. Then, define the reconstrue matrix Y from e_1 as

$$Y = \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1k} \\ e_{12} & e_{13} & \dots & e_{1k+1} \\ \vdots & \vdots & & \vdots \\ e_{1m} & e_{1m+1} & \dots & e_{1M} \end{bmatrix} \quad (17)$$

Where $m > P$, $k > P$, $m + k - 1 = M$, this can be proved [1] that

$$Y = A_1 R_d A_2^T \quad (18)$$

Where $R_d = \text{diag}[s_1(t), s_2(t), \dots, s_k(t)]$, A_1, A_2 are $m \times P$ and $k \times P$ matrices of array steering vector, respectively. The singular value decomposition of Y can be given as

$$Y = U \Lambda V^H \quad (19)$$

Where Λ is the $m \times k$ singular value matrix, U and V are the left and right singular vector matrix, respectively. Then, an effective decorrelation which is modified singular value decomposition (MSVD) algorithm can be applied [2]. Defined the following matrix

$$Y_0 = YY^H = U \Lambda V^H V \Lambda^H U^H = U \Sigma U \quad (20)$$

Where $\Sigma = \Lambda \Lambda^H$ is the eigenvalue matrix of Y_0 . The eigenvector matrix of Y and Y_0 are same, hence, the noise subspace of Y is equal to Y_0 . Just like the spatial smoothing processing, the technique can be applied for Y_0

$$Y_1 = \frac{1}{2} (Y_0 + J_m Y_0^* J_m) \quad (21)$$

Finally, Estimate DOAs of the signal by MUSIC algorithm.

4 Simulation

To assess CRCO_RV algorithm, CRCO_SS algorithm, and CRCO_MSVD algorithm, two performance criteria are introduced to assess the performance of the above algorithms. Firstly, Root-Mean-Square-Error (RMSE) was defined as

$$RMSE = \sqrt{\frac{1}{PC} \sum_{p=1}^P \sum_{c=1}^C (\hat{\theta}_{pc} - \theta_p)^2} \quad (22)$$

Where P denotes the number of independent Monte-Carlo runs, and $\hat{\theta}_{pc}^{\Lambda}$ is the estimated value of θ_p in the p-th run. The two signals are considered to be resolvable if the following resolution criterion

$$P(\theta_m) > \frac{1}{2} \{P(\theta_1) + P(\theta_2)\} \quad (23)$$

Where $P(\cdot)$ is MUSIC spectrum, θ_1 and θ_2 are the two arrival angles for the two signals, θ_m is the mid-range between them. A new SNR measure, generalized signal-to-noise ratio (GSNR) is defined as

$$GSNR = 10 \log_{10} \frac{\sigma_s^2}{\gamma} \quad (24)$$

Where σ_s is the variance of the signal, γ is the dispersion parameter of the noise.

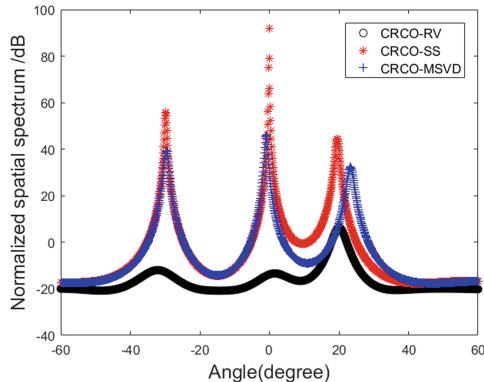


Fig. 3. Spatial spectrum ($\alpha = 1.1$)

Example 1. Spatial Spectrum

Spatial spectrum simulation is presented to demonstrate the performance of the proposed algorithms directly. Two coherent signals QAM impinges on the ULA with 8

elements from the direction of $\theta_1 = -30^\circ$ and $\theta_2 = 0^\circ$, also an uncorrelated signals QAM arriving at the array from $\theta_2 = 20^\circ$, the number of snapshots and GSNR is set respectively equal to 256 and 8 dB. 3 subarrays with 6 elements, simulation results are shown in Fig. 3. From Fig. 3 we can see that the performance of CRCO_RV algorithm fails to estimate the DOA. Hence, we only analysis the CRCO_SS algorithm and CRCO_MSVD algorithm in the follow simulations.

Example 2. Characteristic Exponent α

Two coherent sources QAM impinge on the ULA with 8 elements from $\theta_1 = 0^\circ$ and $\theta_2 = 10^\circ$, GSNR = 10 dB, snapshots = 1024, 3 subarrays with 6 elements. 100 Monte-Carlo runs are performed. Figure 4 shows that the performance of the SS method will be degradation along with the increase of α .

Example 3 Angular Separation

Two coherent sources imping on the ULA with 8 elements from $\theta_1 = 0^\circ$ and $\theta_2 = \theta_1 + \Delta\theta$ under the condition of $\alpha = 1.5$ and GSNR = 10 dB are considered, snapshots = 1024, $\Delta\theta \in [1^\circ, 10^\circ]$ step value is 1° . 3 subarrays with 6 elements. 100 Monte-Carlo experiments are conducted Fig. 5 gives the experiment result.

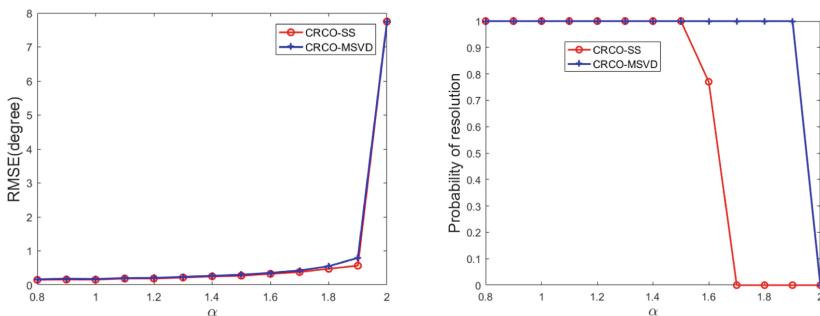


Fig. 4. RMSE and Probability of resolution versus characteristic exponent α

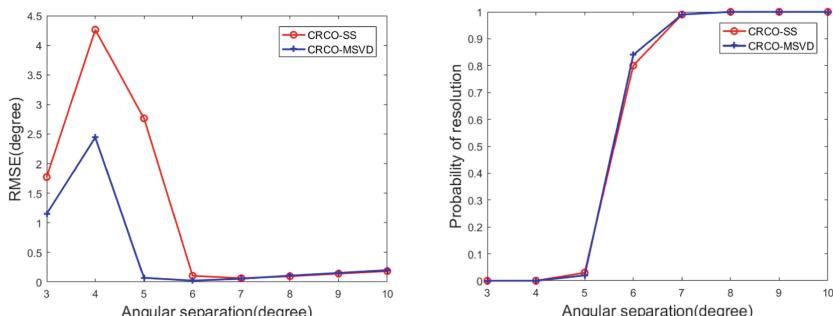


Fig. 5. RMSE and Probability of resolution versus angular separation

5 Conclusion

The characteristics of this paper can be summarized as:

- (1) By rebuild the covariance matrix of array output based on correntropy, we achieve the suppression of outliers.
- (2) By using the idea of spatial smoothing algorithm reconstruct matrix, we achieve the objective of decorrelation.
- (3) Simulation results demonstrate the effectiveness of CRCO_MSVD and CRCO_SS.

Acknowledgments. This work has been supported by the National Natural Science Foundation of China (Grant No: 61701344), Tianjin Edge Technology and Applied Basic Research Project (14JCYB15800), Tianjin Normal University Application Development Foundation (52XK1601), Tianjin Normal University Doctoral Foundation (52XB1603, 52XB1713) and Tianjin Higher Education Creative Team Funds Program in China.

References

1. Gao W, Bao Z (1987) Data-based matrix decomposition technique for high-resolution array processing of coherent signals. *IEEE Electron Lett* 23(6):643–645
2. Zhang J, Huang DS (2013) Estimation DOAs of the coherent sources based on SVD. In: 2010 2nd International Conference on Signal Processing Systems, vol 3, pp 174–177
3. Pillai SU, Kwon BH (1989) Forward/backward smoothing techniques for coherent signal identification. *IEEE Trans Acoust Speech Sig Process* 37(1):8–15
4. Haardt M, Nossek JA (1995) Unitary ESPRIT: how to obtain increased estimation accuracy with a reduced computational burden. *IEEE Trans Sig Process* 43:1232–1242
5. Pesavento M, Gershman AB, Haardt M (2000) Unitary root-MUSIC with a real-valued eigendecomposition: a theoretical and experimental performance study. *IEEE Trans Sig Process* 48:1306–1314
6. Zhang J, Qiu T, Song A, Tang H (2014) A novel correntropy based DOA estimation algorithm in impulsive noise environments. *Sig Process* 104:346–357
7. Santamaría I, Pokharel PP, Principe JC (2006) Generalized correlation function: definition, properties and application to blind equalization. *IEEE Trans Sig Process* 54(6):2187–2197
8. Tsakalides P, Nikias CL (1996) Robust covariation-based MUSIC (ROC-MUSIC) algorithm for bearing estimation in impulsive noise environments. *IEEE Trans Sig Process* 44(7):1623–1633
9. Ollila E, Eriksson J, Koivunen V (2011) Complex elliptically symmetric random variables—generation, characterization, and circularity tests. *IEEE Trans Sig Process* 59(1):58–69
10. Schmidt RO (1986) Multiple emitter location and signal parameter estimation. *IEEE Trans Antennas Propag* 34(3):276–280



Design and Research of 100KV DC Power Supply Based on SiC

Feng Ji^(✉) and John Xu

Faculty of Science and Engineering, University of Nottingham Ningbo China,
Ningbo 315000, China
Feng.Ji@nottingham.edu.cn

Abstract. In this paper, a 100kV DC power supply is designed by using SiC as power device material and DSP as control core. The power supply mainly consists of rectifier, voltage stabilizer, DC chopper, low pass filter, inverter, high frequency voltage converter, voltage doubling rectifier, feedback measurement, and protection subsystem. The advantages of SiC are introduced in this paper. The structure and principle of main circuits are analyzed in detail. The ZVS soft switching technology is adopted to reduce energy loss. A simulation is carried out by using MATLAB. The experimental results show that the design is feasible.

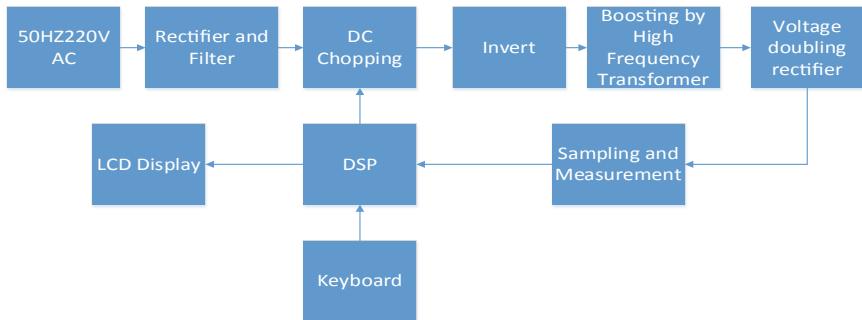
Keywords: SiC · PWM · Rectifier · ZVS · Inverter · DSP

1 Introduction

High voltage power grid is composed of a large number of high voltage equipment. The high voltage electrical equipment need to be tested for withstand voltage and insulation resistance before operation. In the process of testing, it is inevitable to use high voltage DC power supply. At the same time, DC power supply has been widely used in X-ray machine, CT machine, metal smelting, electrostatic precipitation and other fields. The traditional HVDC power supply has some shortcomings, such as large volume, high ripple coefficient, low precision, high energy consumption and low intelligence [1–3]. In this paper, a digital HVDC power supply with SiC power device and DSP from TI company is proposed. Its maximum output voltage is 100kV. The high voltage DC power supply has many advantages, such as small size, small output ripple coefficient, easily controllable and automatic protection. The overall design is shown as Fig. 1.

2 Analysis of SiC Power Devices

The traditional high voltage DC power supply uses Si-based power components. With the development of science and technology, it is difficult to furtherly improve the working frequency and efficiency because of the limit of physical properties of Si-based material. As a new generation of power electronic device material, SiC has the advantages of low on-resistance, low switching loss, high switching frequency and

**Fig. 1.** Overall design block diagram

good high temperature resistance. Table 1 shows the details of performance comparison between Si and SiC [4–7].

Table 1. Properties comparison of Si and SiC.

| Property | Si | SiC |
|--|--------------------|--------------------|
| Bandgap Energy, E_g (eV) | 1.1 | 3.23 |
| Breakdown Field, E_{crit} (MV/cm) | 0.3 | 2 |
| Thermal Conductivity, λ (W/cm * K) | 1.5 | 4.9 |
| Saturation Drift Velocity, v_{sat} (10^7 cm/s) | 1 | 2 |
| Electron Mobility, μ_n (cm ² /V * s) | 1400 | 1000 |
| Hold Mobility, μ_p (cm ² /V * s) | 471 | 115 |
| Baliga's Figure of Merit (BFoM) | 1 | 560 |
| Intrinsic Carrier Concentration, n_i (cm ⁻³) | 1×10^{10} | 5×10^{-9} |

Because of the advantages of SiC, SiC MOSFET and SiC Diode are planned to be used as the power components.

3 Rectifier Circuit with Filter and Voltage Stabilization

As shown in Fig. 2, 220 V and 50 Hz AC power source can be dealt with by the bridge rectifier circuit after step-down transforming. With the subsequent filter and stabilization, a smooth DC output is obtained. The principle of dynamic regulation of stable DC output is as follows:

When R_L remains unchanged, the input voltage $U_i \nearrow \rightarrow U \nearrow \rightarrow U_z \nearrow \rightarrow I_z \nearrow$
 $I_R \nearrow = I_z - I_L \rightarrow U_R \nearrow \rightarrow U_o = kU_i \nearrow - U_R \nearrow$

So the output voltage U_o remains stable.

When U_i remains unchanged, $R_L \nearrow \rightarrow I_L \searrow \rightarrow I_R \nearrow = I_z + I_L \rightarrow U_R \searrow \rightarrow U_z \nearrow = kU_i - U_R \rightarrow U_o \nearrow \rightarrow I_L \nearrow$

So it remains the stability of output.

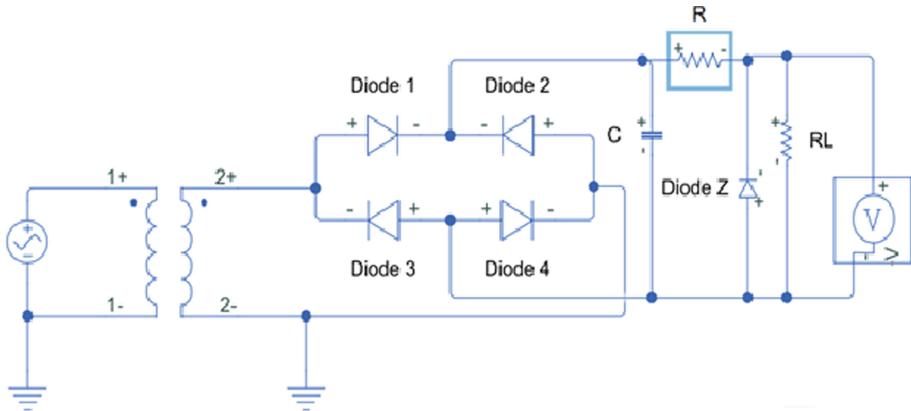


Fig. 2. Rectifier circuit

4 Buck DC Chopper Circuit

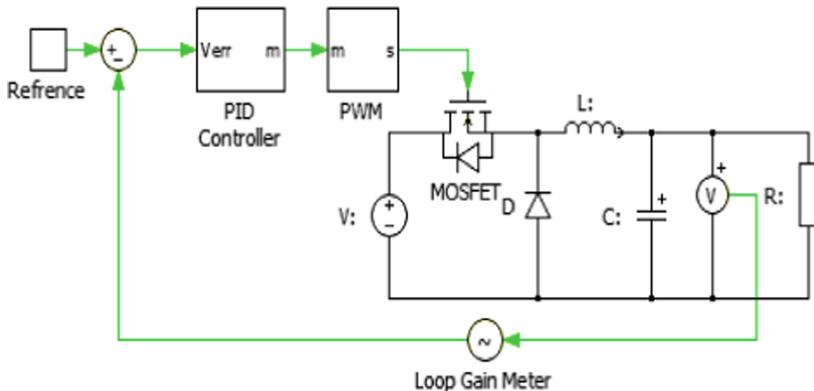


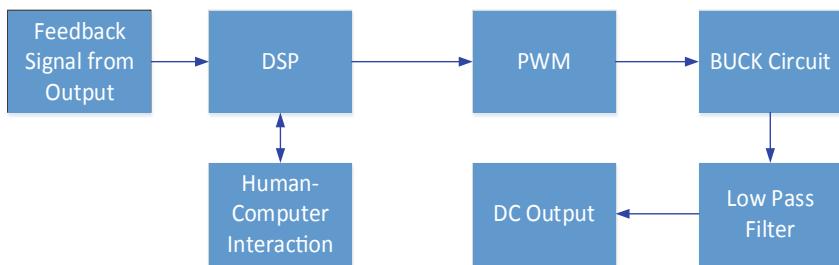
Fig. 3. Buck circuit

D is duty cycle. The circuit is buck chopper [8, 9] (Fig. 3).

$$U_o = D * U_i = U_i * T_{on}/T \quad (1)$$

Closed-loop PID control is proposed to achieve stable output. The PID formula is:

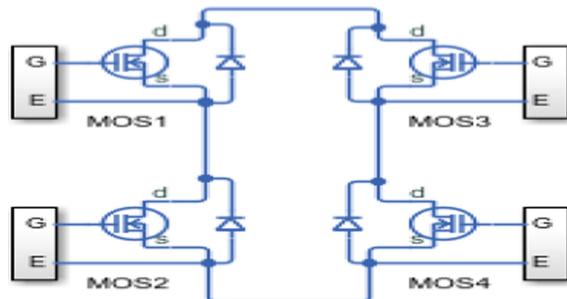
$$U(t) = K_P \left[E(t)dt + \frac{1}{T_1} \int E(t)dt + T_2 \frac{dE(t)}{dt} \right] \quad (2)$$

**Fig. 4.** DC block diagram

DSP receives the feedback signal from output and adjusts the duty cycle of PWM to achieve the purpose of voltage regulation (Fig. 4).

5 Full Bridge Inverter Circuit Based on ZVS Soft Switching

Full bridge inverter circuit will be used in the part. The full bridge circuit is shown as Fig. 5.

**Fig. 5.** Full bridge circuit

In the circuit, four SiC MOSFET on the bridge arm are driven by UG3524. The driving circuit is shown as below (Fig. 6).

Because there is no feedback from output, fixed PWM driver is directly generated by hardware. In order to reduce the loss of MOSFET during switching, ZVS resonant soft switching technology is proposed. There is a dead time between the on and off of MOSFET on the same bridge arm to prevent short circuit during the switching period. In series resonant circuit, the formula is:

$$L_e = L_r + nL \quad (3)$$

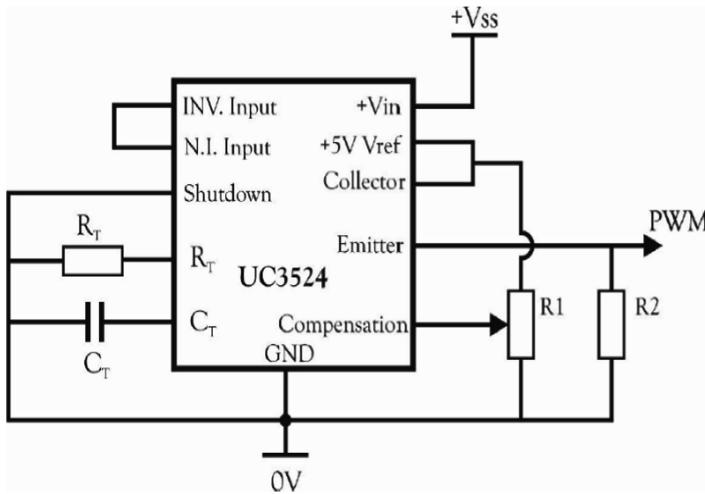


Fig. 6. UG3524 drive circuit

Among them, L_r is serial resonant inductance, L is primary inductance of transformer, L_e is the whole inductance. C is the parasitic capacitance of switch tube and distributed capacitance of transformer.

When $W * L_e = 1/(W * C)$, resonance occurs. The output voltage is square wave and the current is sinusoidal fundamental wave. In the vicinity of natural zero-crossing point of current, ZVS switching is realized with the parallel diode D and capacitor C.

The voltage after inversion is boosted by high frequency transformer and voltage doubling rectifier circuit. Then the output voltage can reach 100kV (Fig. 7).

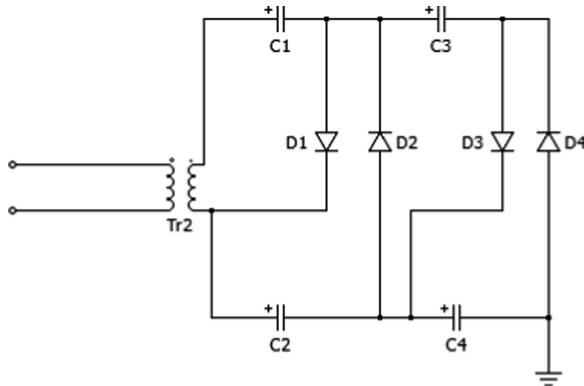


Fig. 7. Four times voltage circuit

The main reason for not just using high frequency transformer to boost voltage is that its insulation level is limit and it is difficult to reach 100kV. Therefore, the boost

voltage can be achieved together with 4 times voltage rectifier circuit. The formula is shown as the following:

$$U_{\text{out}} = U_{c2} + U_{c4} = 2U_{c1} + 2U_{c3} = 2U_{\text{in}} + 2U_{\text{in}} = 4U_{\text{in}} \quad (4)$$

6 The Whole System Simulation

The whole system will be turned off during overcurrent or overvoltage. And the overall output will be 0 V. The protection control equation is as follows:

$$\text{PWM output permission} = (\text{current} < \text{protection setting}) \& \& (\text{voltage} < \text{protection setting}) \quad (5)$$

The system level simulation is carried out by using MATLAB. The waveform is shown as Fig. 8.

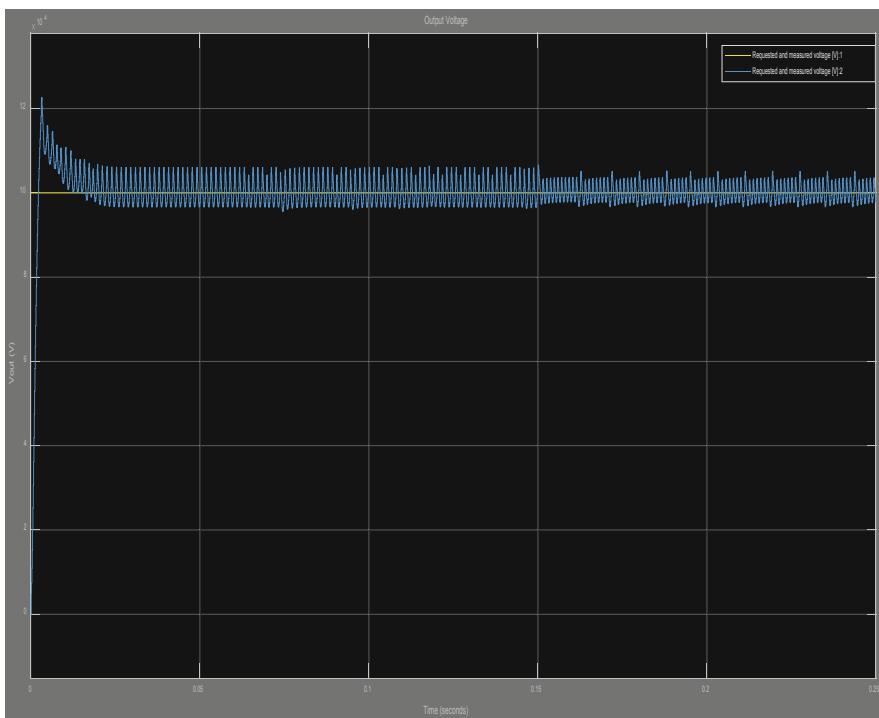


Fig. 8. Output waveform of simulation

The simulation result shows that the output of 100kV DC power instrument can work well and the design is feasible.

References

1. Bahmani MA, Thiringer T, Kharezy M (2016) Design methodology and optimization of a medium-frequency transformer for high-power DC-DC applications. *IEEE Trans Ind Appl* 52:4225–4233
2. Lutz J, Schlangenotto H, Scheuermann U, De Doncker R (2018) Semiconductor Power Devices - Physics, Characteristics, Reliability
3. Jovcic D, Hajian M, Zhang H, Asplund G (2012) Power flow control in DC transmission grids using mechanical and semiconductor based DC/DC devices. In: 10th IET international conference on AC and DC power transmission (ACDC 2012), Birmingham, pp 1–6
4. Baliga BJ (2008) Fundamentals of Power Semiconductor Devices, 1st edn. Springer, Heidelberg
5. Weitzel CE, Palmour JW, Carter CH, Moore K, Nordquist KK, Allen S, Thero C, Bhatnagar M (1996) Silicon carbide high-power devices. *IEEE Trans Electron Devices* 43:1732–1741
6. Baliga BJ (2006) Silicon carbide power devices. World Scientific, Singapore
7. Baliga BJ (2017) Gallium nitride and silicon carbide power devices. World Scientific, Singapore
8. Fernanda LAM, Vladimir CDG, Manuel TJ, Héctor BB, Erwin ZT, Felix QM Comparative analysis of an DC-DC interleaved converter of high power and control techniques. In: 2015 IEEE workshop on power electronics and power quality applications (PEPQA), Bogota, pp 1–5
9. Sousa TJC, Monteiro V, Pinto JG, Meléndez AAN, Afonso JL (2017) Single-phase shunt active power filter with UPS operation using a bidirectional Dc-Dc converter for energy storage interface. In: IECON 2017 - 43rd annual conference of the IEEE industrial electronics society, Beijing, pp 4972–4977



Design and Simulation Analysis of NA-PMSM Based on ANSYS

Yuyang Yuan^(✉), Bulai Wang, Jiale Zhong, and Xiangsheng Liu

Shanghai Institute of Technology, Shanghai 201418, China
yuysunrise@163.com

Abstract. In this paper, the design research of 11 kW a Non-uniform Air-gap Permanent Magnet Synchronous Motor(NA-PMSM). Firstly, RMxprt equivalent magnetic circuit method in ANSYS was used for the preliminary electromagnetic design of the motor, and then the two-dimensional transient field model of the motor was built by using the finite element method to further analyze the no-load reverse electric potential and load torque of the motor. At the same time, the air-gap magnetic density waveform and torque-power angle curve were tested to verify the feasibility of the electromagnetic design scheme and achieve the expected effect.

Keywords: NA-PMSM · Equivalent magnetic circuit method · RMxprt · The finite element method

1 Introduction

With the gradual extraction and exhaustion of fossil fuels and the increasingly serious environmental pollution problem, among which a large proportion of pollution is fuel cars, in order to improve the environment and sustainable development, the world must develop and utilize new energy sources. Electric cars have become more and more important in recent years. Obviously, the motor is its “heart”, the performance of the motor plays a decisive role in electric vehicles. Compared with Switched Reluctance Motor, Induction Motor and Brushless DC Motor, Permanent Magnet Synchronous Motor has the advantages of high power density, high overload capacity and high efficiency [1], and is widely used in the power drive system of electric vehicles.

RMxprt is a professional motor design module embedded in the finite element software ANSYS. Its basic principle is based on the design method of equivalent magnetic circuit, and the integrated simulation models of various motors can quickly realize the motor model and parameter calculation, and the accuracy basically meets the requirements of engineering design [2]. RMxprt completed the preliminary design of the motor, and then completed the model of the actual motor in Maxwell2D, constructed the transient field model, carried out more accurate finite element analysis and calculation, obtained the main data curve of the motor, checked the motor performance, and provided more accurate data support for further optimization design of the motor [3].

2 Method

In order to realize the social popularization of electric vehicles and accelerate the development of new energy as soon as possible, an electric vehicle enterprise proposed a permanent magnet synchronous motor, whose performance indexes are shown in Table 1 below:

Table 1. Technical index table

| Parameter | Number or type |
|-------------------|----------------|
| Rated output | 11 kW |
| Rated speed | 1500 rpm |
| Insulation class | 1F |
| Temperature rise | 175 K |
| Manner of working | S1 |
| Protection grade | IP54 |

2.1 Motor Size Design

When designing a motor, many dimensions must be determined, but the main and decisive factor is the main size of the motor. The main size of the motor is the armature core diameter D and length L_{ef} . For DC motor, armature diameter refers to the rotor diameter; For induction and synchronous motors of general construction, the specified subdiameter. The main size of the motor is as follows:

$$\frac{D^2 l_{ef} n}{p'} = \frac{6.1}{\alpha'_p K_{Nm} K_{dp} A B_\delta} \quad (1)$$

On the type: p' —calculated motor power, n —motor speed;

K_{Nm} —the waveform coefficient of air gap magnetic field is 1.1 when the air gap magnetic field is sinusoidal;

B_δ —the air gap flux density; K_{dp} —winding coefficient of the motor. The difference between the value and the fundamental wave winding coefficient K_{dp1} is very small; α'_p —calculate the polar arc coefficient; A—line load [4].

On the basis of the above calculation of motor parameters, and after the comparison and adjustment of multiple schemes, the inner diameter of the stator is determined to be 152 mm, the outer diameter of the stator is 260 mm, and the core length is 90 mm.

2.2 Motor Rotor Design

At present, most of the researches are carried out on the basis of the magnetic pole structure of the rotor with uniform air gap. The fundamental reason for the large core

loss and high electromagnetic noise of ordinary permanent magnet synchronous motor is the complex harmonic of air gap magnetic field. Non-uniform air gap structure can well weaken the harmonic of air gap magnetic field, so as to reduce loss and noise and improve motor performance [5].

The built-in permanent magnet has the advantages of strong rotor structure, strong anti-demagnetization ability, good weak magnetic property and wide high-efficiency operating range, etc. At the same time, the magnetoresistance torque can be obtained by using the different magnetoresistance of d and q axes. Here, the permanent magnet rotor structure with built-in radial “one” type is selected [6]. In the engineering design, we can estimate the size of permanent magnet according to experience, and take the magnetic property of permanent magnet, insulation grade of motor and price into consideration. In this paper, NdFeB is finally selected. Its specific parameters are as follows: Maximum magnetic energy product: 310 kJ/m³; HCJ: 1880KA/m; Remanence B_r : 1.22T; Maximum operating temperature: 180 °C; Density: 7.5 g/cm³. The designed rotor diagram is shown in Fig. 1. Single side air gap length: 0.5 mm; Silicon steel plate number: DW315-50; Magnetic steel size: 63 mm × 6 mm; Magnetic steel grades: N38UH.

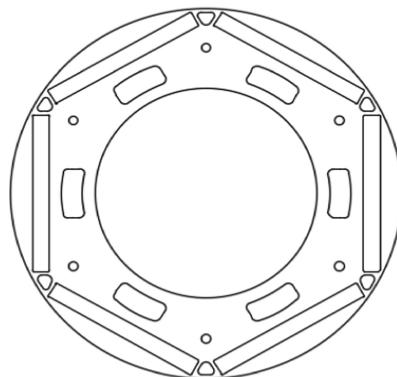


Fig. 1. Rotor structure

The stator and rotor core materials should be selected in accordance with the principle of strong magnetic conductivity, weak electrical conductivity and effective reduction of hysteresis loss and eddy current loss of the magnetic field in the core. Here, DW315-50, a cold-rolled non-oriented silicon steel sheet with high performance, is adopted. Its B-H curve is shown in Fig. 2.

It can be seen from the figure that this material has the maximum utilization rate at the critical saturation point of 1.6 T.

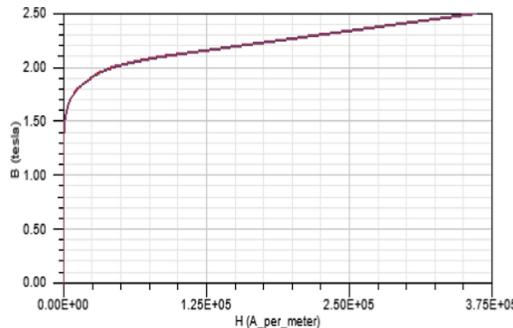


Fig. 2. B–H curve of DW315-50

2.3 Motor Stator Design

Pear-shaped groove can effectively reduce core loss and pulsating loss, and its groove area utilization rate is large, long service life and not prone to bending deformation, so this paper chooses pear-shaped groove. The pear-shaped groove is shown in Fig. 3, Standard sizes are not shown here.

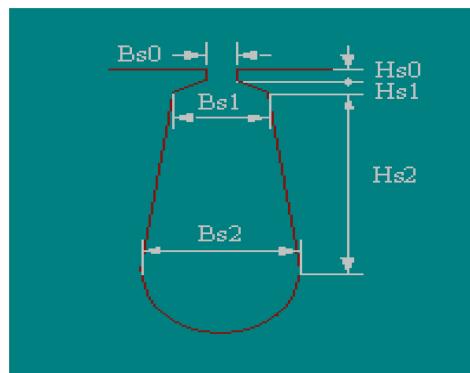


Fig. 3. Schematic diagram of pear groove

Torque ripple and electromotive force harmonic of permanent magnet synchronous motor should also be considered in the design to minimize torque ripple. In this paper, the motor is selected as 6-pole and 36-slot, double-layer short-pitch winding and y-shaped connection [7].

The wire diameter and enameled wire insulation are the reference conditions for the selection of stator windings. The rated current is estimated to be about 17 A by the motor design. Here, it is appropriate to choose multiple parallel connections and multiple heel and winding, because the current value is larger and exceeds the load of a single wire. Figure 4 is the stator connection structure diagram formed in RMxprt.

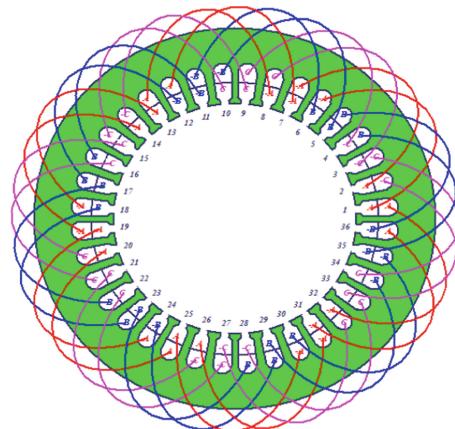


Fig. 4. Stator connection structure diagram

3 Prototype and Test Results

3.1 Maxwell Model Construction

In order to make the designed motor size accurate, this paper first accurately drew the motor model in AutoCAD, and then imported it to Maxwell for modeling and simulation of each part of the motor before processing. Because this paper is a six-pole motor, and the corresponding part of each pole is symmetrical, in order to save simulation time, split division is used to leave only the first quadrant part. The motor model after the pretreatment is shown in Fig. 5:

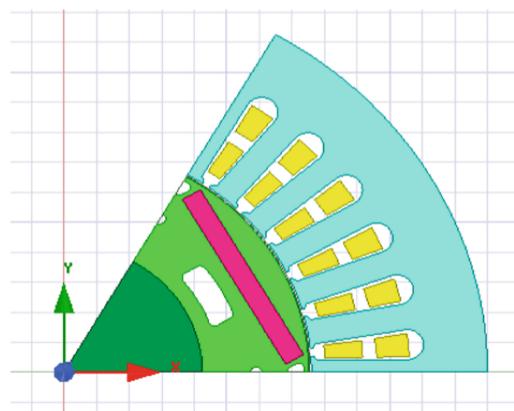


Fig. 5. Motor model in Maxwell

3.2 Test Results

The magnetic density cloud diagram and magnetic force diagram of the motor can be obtained by finite element analysis as shown in Fig. 6. It can be seen from the figure that the direction of the permanent magnet magnetic field line is basically parallel to the magnetization direction of the permanent magnet, and the magnetic field line at the end of the permanent magnet deviates strongly from the magnetization direction. At the same time, it can be seen that the more brightly colored the position magnetic field is, the stronger it is, indicating that the magnetic field effect established by the motor is better.

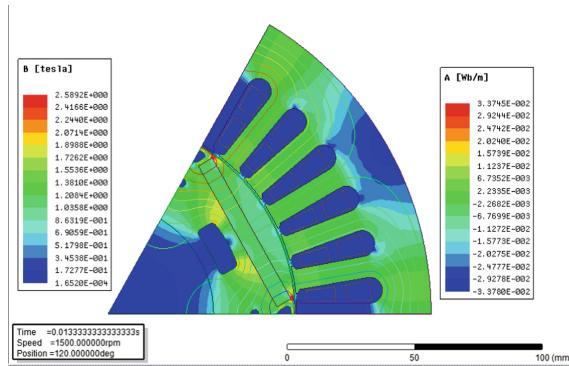


Fig. 6. Magnetic clouds figure and magnetic figure

The rated torque of the motor can be obtained by calculating the torque formula below.

$$T_N = \frac{9.549 \times P_N \times 10^3}{n_N} = 70.026 \text{ (N)} \quad (2)$$

The rated torque is obtained by applying rated current to the stator winding. The simulation results are shown in Fig. 7.

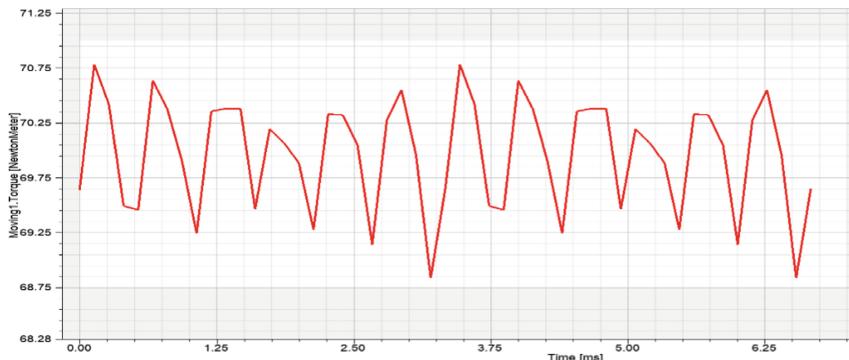


Fig. 7. Torque diagram

As can be seen from the simulation figure, the maximum torque is 70.75 N, and the minimum torque is 68.85 N. Due to the special cogging torque problem of permanent magnet motor, the torque curve fluctuates a little, but it is within the normal range. Generally speaking, the result is ideal.

Figure 8 shows the waveform of air-gap magnetic density of motor load obtained by finite element analysis of motor. The top of the waveform of air-gap magnetic density is stable without distortion, indicating that the motor can run smoothly.

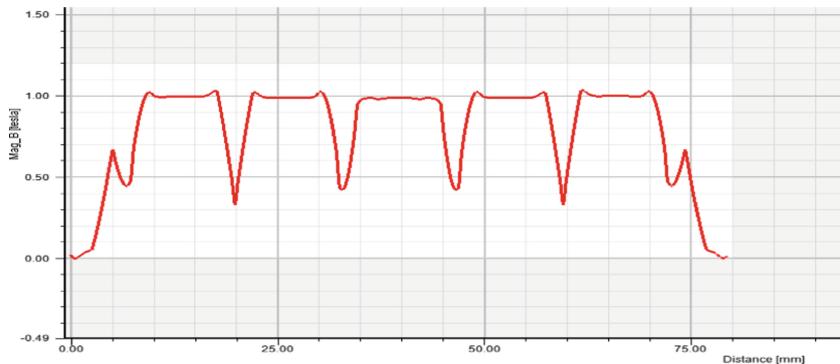


Fig. 8. Air gap magnetic density diagram

The three-phase no-load counter-potential curve of the motor is shown in Fig. 9. It can be seen from the figure that it changes sinusoidal, and the period and amplitude also meet the design requirements, which proves the rationality of motor winding design from the side.

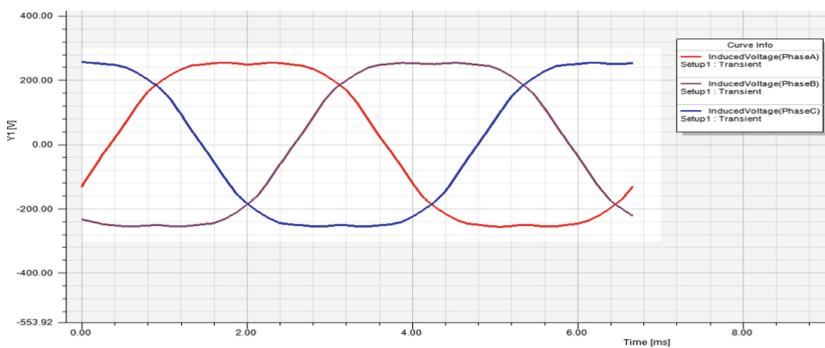


Fig. 9. No load back potential curve

The torque-power Angle curve of the motor is plotted with the Toolkit of ANSYS, as shown in Fig. 10. This curve represents the torque of the motor at different power factor angles, indicating that there is a corresponding torque corresponding to each power Angle, and the operating range is relatively wide.

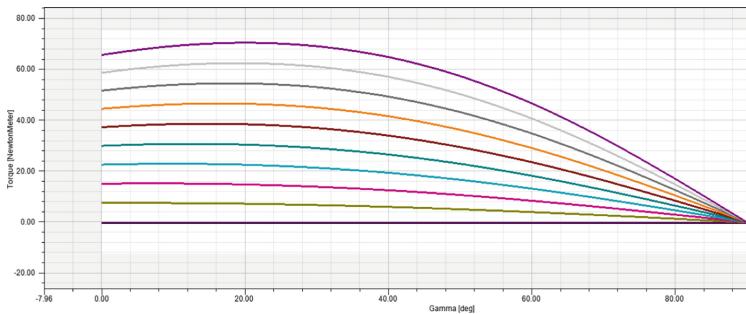


Fig. 10. Torque and gamma curves

4 Conclusion

Based on the ANSYS software to design a paragraph 11 kw of permanent magnet synchronous motor, using RMxprt module to complete the preliminary design of the motor, and then import the Maxwell do further processing and simulation, its performance can be analyzed by the simulation results, the basic consistent with design requirements at the same time, it is broad run efficiently interval also satisfy the design index, achieved the purpose of energy conservation and efficient, so this design is feasible.

Acknowledgments. The outcomes of this thesis and teacher wang and senior lab, the teacher elder sister also have their own efforts are inseparable, thank teacher wang in the guidance and teachings, elder sister at the same time thanks to the help and encouragement, is that they give encouragement when I am in trouble, let me hold on and succeed, thank them again, my teacher, brother, the teacher elder sister.

References

1. Carraro E, Bianchi N (2014) Design and comparison of interior permanent magnet synchronous motors with non-uniform airgap and conventional rotor for electric vehicle applications. *IET Electric Power Appl* 8(6):240–249
2. Wang G, Huang S, Jian Gao (2011) Analysis of built-in permanent magnet synchronous motor based on Ansoft. *Micromotor* 44(2):70–73
3. Wang B, Ji X (2018) Research on electromagnetic design of ultra-high efficiency permanent magnet synchronous motor. *Electric Transm* 48(6):73–74
4. Chen S (2000) Motor design. China Machine Press, Xi'an
5. Chen L, Wu C (2018) Effect of load characteristics on demagnetization of asynchronous starting permanent magnet synchronous motor with non-uniform air gap structure. *J Electr Eng* 13(13):2
6. Wang G, Huang S (2011) Design of interior permanent synchronous motor based on Ansoft software. *Micromotors* 44(2):70–73
7. Hsu J, Ayers C, Coomer CL (2004) Report on toyota/prius motor design and manufacturing assessment. Oak Ridge National Laboratory



A Novel Hybrid Model of Wind Speed Forecasting Based on EWT, BiLSTM, SVR Optimized by BOA in Inner Mongolia, China

Renqiang Ding, Wuneng Zhou^(✉), and Hangyang Cheng

School of Information Science and Technology, Donghua University,
Shanghai 200051, China
zhouwuneng@163.com

Abstract. As an important renewable energy, wind power plays an important role in promoting environmental protection and economic improvement. Accurate wind speed prediction is of great significance for wind farms to utilize wind energy and equipment management and maintenance. A novel hybrid model of wind speed forecasting based on EWT, BiLSTM, SVR optimized by BOA in Inner Mongolia China is proposed in this paper. The hybrid EWT-BiLSTM-SVR model is constructed as follows: (a) The EWT is employed to decompose the original wind speed series into two sub-layers; (b) The BiLSTM networks are utilized to predict each sub-layer; (c) The Bayesian optimization algorithm is used to optimize the hyperparameters of SVR, and the SVR model is adopted to predict the error series; (d) The IEWT is executed to reconstruct the predicted sub-layers and error series. The 10 min wind speed data from a wind farm in Inner Mongolia, China is used in this paper as case studies. Compared with the other six models, the results show that the hybrid EWT-BiLSTM-SVR model can improve the accuracy of wind speed forecasting and has better performance.

Keywords: Wind speed forecasting · BiLSTM (Bi-directional Long Short-term Memory) network · SVR (Support Vector Regression) · EWT (Empirical Wavelet Transform) · BOA (Bayesian Optimization Algorithm) · Hybrid model

1 Introduction

Since the beginning of industrialization, petrochemical energy, such as coal and oil, has become the material basis for human society to survive. However, traditional petrochemical energy storage is limited and non-renewable. With the continuous reduction of petrochemical energy, the energy crisis is becoming more and more prominent. At the same time, the massive use of petrochemical energy has caused damage to the ecological environment, and the emission of large amounts of greenhouse gases has caused global warming. The development and utilization of renewable energy has become the consensus of all countries in the world. In China, the share of non-fossil fuel as the primary consumptions will be increased to 20% in 2030 [1]. As a kind of green and clean renewable energy, wind energy has extremely important practical significance for solving energy crisis and environmental pollution, and has gradually become the focus

of attention of all countries in the world. According to the statistics of the Global Wind Energy Association (GWEC), the total installed wind energy capacity of the world has reached 432.9 GW by the end of 2015, with a cumulative increase of over 17% [2].

With the continuous development of wind power, it also brings severe challenges to the security and stability of power system [3, 4]. Therefore, a stable and accurate wind speed prediction method is needed to solve this problem [5, 6]. After continuous research, many forecasting models have been put forward to improve the accuracy of wind speed prediction. These proposed models can be classified into physical models, statistical models, artificial intelligence models and hybrid models [7].

The input data of the physical models is usually a variety of physical descriptions, including meteorological information and wind farm topographic features, surface roughness, obstacles and other geographic information [8, 9]. The advantage of the physical models is that it considers complex physical quantities such as meteorology and geography, and has less demand for historical data of wind farms, and is suitable for medium and long-term wind speed prediction. The drawback is that it needs to establish an accurate geophysical model with high computational complexity. The statistical models, without the aid of numerical weather prediction (NWP), a better prediction result can be obtained using only historical data of wind farms [10]. But it lack the ability to deal with the nonlinear patterns, at the same time, the limitations of data limit the generalization of the statistical models.

With the rapid development of computer technology, artificial intelligence models have emerged. They commonly consist of an input layer, one or more hidden layers, and an output layer. Each layer has some artificial neurons which are connected with the neurons of the previous layer with a connection function. The results show that the different structures of the network lead to different wind speed forecasting performances [11, 12]. The artificial intelligence models can solve the problem that the statistical model can not deal with the nonlinear situation. It has important use value for dealing with wind speed data with a large number of nonlinear features, and can improve the accuracy of wind speed prediction.

Owing to the extremely unsteady characteristic of the wind speed, the individual artificial intelligence models depend on a large amount of historical wind speed data to build models, and sometimes stuck into local minimum, failed to achieve satisfactory performance. Therefore, the hybrid models were proposed [13–15]. The hybrid models can be divided into two main directions. The first mainstream focuses on the processing of the input signals. Through different signal processing methods, such as empirical mode decomposition (EMD), wavelet transform (WT) and singular spectrum analysis (SSA), to achieve denoising and frequency division processing of the input signal. Naik et al. proposed a hybrid model by combining the EMD and the KRR (kernel ridge regression). The EMD was executed to decompose the original wind speed and wind power series into several components [16]. The experimental results indicated that the EMD was beneficial to remove the mutual effects among different components. Another mainstream focuses on combining the predictions of different single prediction models and optimizing the weights to achieve optimal predictions.

Inspired by the success of deep learning methods have gotten in the field of computer vision and speech signal processing, the deep learning methods were

involved in the field of wind speed and wind power forecasting recently [17]. Deep learning is a widely used big data analysis technology. Compared with the ordinary neural network, it has a stronger ability to learn complex functions. The gradient dispersion caused by random initialization can be overcome by layer-by-layer learning and fine-tuning, and the high-level abstract concept can be extracted and expressed, effectively avoiding falling into local optimum [18, 19].

In the study, a novel hybrid EWT-BiLSTM-SVR model is proposed to improve the wind speed forecasting accuracy. The hybrid model is constructed as follows:(a) The EWT (empirical wavelet transform) is employed to decompose the original wind speed series into two sub-layers; (b) The BiLSTM (bi-directional long short-term memory) network is utilized to predict each sub-layer; (c) The Bayesian optimization function is used to optimize the hyperparameter of SVR (support vector regression), and the SVR model is adopted to predict the error series; (d) The IEWT is executed to reconstruct the predicted sub-layers and error series. A novel EWT method is used to preprocess the wind speed data, and the wind speed data is decomposed into two parts as the input of the predictor. The deformed BiLSTM model, which is a popular deep learning method in sequence analysis, is used as the predictor to give full play to the advantages of the depth learning model. The SVR model is combined and used as a predictor of error sequence. Meanwhile, the Bayesian optimization function is adopted to optimize the hyperparameter, making the predictor in the optimal state. Finally, the wind speed data is reconstructed by the IEWT (inverse empirical wavelet transform) method to obtain the optimal wind speed prediction result. The experimental results show that the proposed method can improve the accuracy of wind speed prediction.

2 The Proposed Hybrid EWT-BiLSTM-SVR Model

As a novel signal analysis and processing method, EWT can divide the wind speed data sequence into data sequences with different frequencies. As an excellent deep learning time series prediction method, BiLSTM can explore and exploit the hidden information of dynamic time series efficiently. Based on the principle of minimum construction risk, SVR has strong generalization ability and can avoid highly dependence on wind speed data. The hybrid EWT-BiLSTM-SVR model proposed in this paper combines three excellent methods and uses the Bayesian optimization algorithm to optimize the hyperparameters of the SVR model. The detailed method steps are as follows:

- (1) Use EWT algorithm to decompose the original wind speed data series into different frequency data sequences. The case in this paper is to decompose the original wind speed data into two different frequency data sequences.
- (2) Set the data sequence obtained by step (1) as Training Set and let it be the input to R-BiLSTMs. At the same time, the Training Set is divided into two parts: Training Set 1 and Training Set 2, which are inputs to L-BiLSTMs and SVR predictors respectively, as shown in Fig. 1.
- (3) Training Set 1 is used as the input of the L-BiLSTMs predictor. The L-BiLSTMs predictor is to predict the wind speed data with the same length of Training Set 2, and let it be Predicted Set 1, as another input to the SVR predictor.

- (4) Training set 2 and Predicted Set 1 are the two inputs of the SVR predictor, and the subtraction process is first performed, and the obtained value is used as the training data of the SVR predictor. And at the same time, Bayesian optimization algorithm is applied to the hyperparameter optimization of SVR, and SVR predictor is used as an error predictor to predict the wind speed error sequence, as shown in Fig. 2.
- (5) Reconstruct the wind speed data predicted by the R-BiLSTMs predictor and the error sequence predicted by the SVR predictor by the IEWT algorithm to obtain the final wind speed prediction result, as shown in Fig. 3.

The specific flowchart of the hybrid EWT-BiLSTM-SVR model proposed in this paper is shown in Fig. 4.

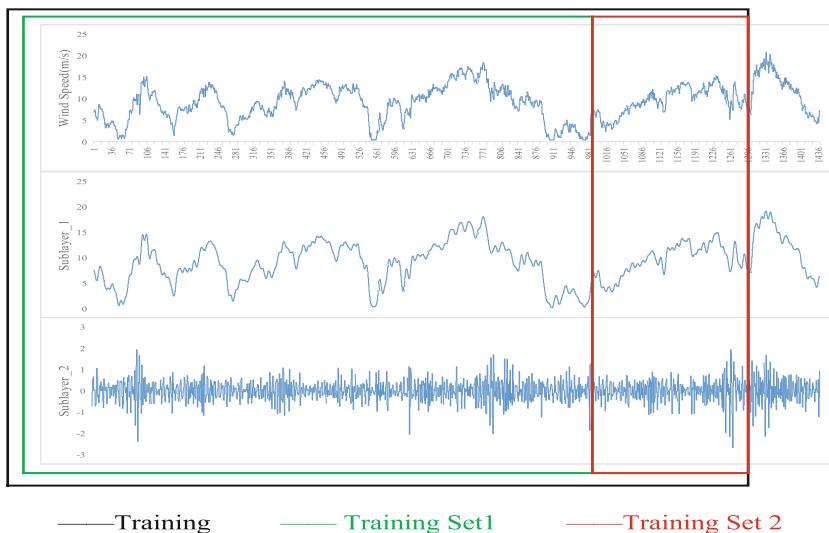


Fig. 1. Wind speed data decomposition and division

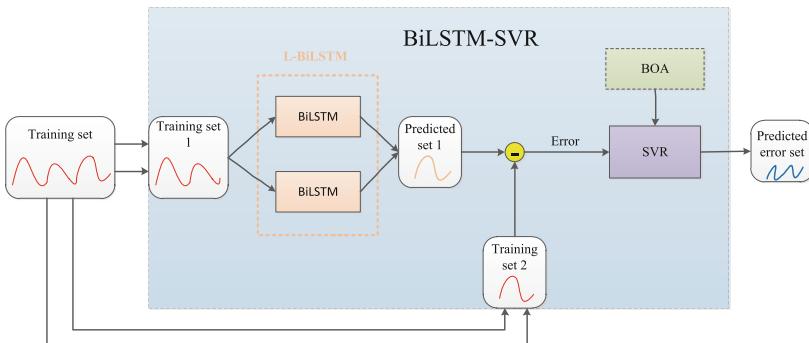


Fig. 2. The structure of error series prediction model

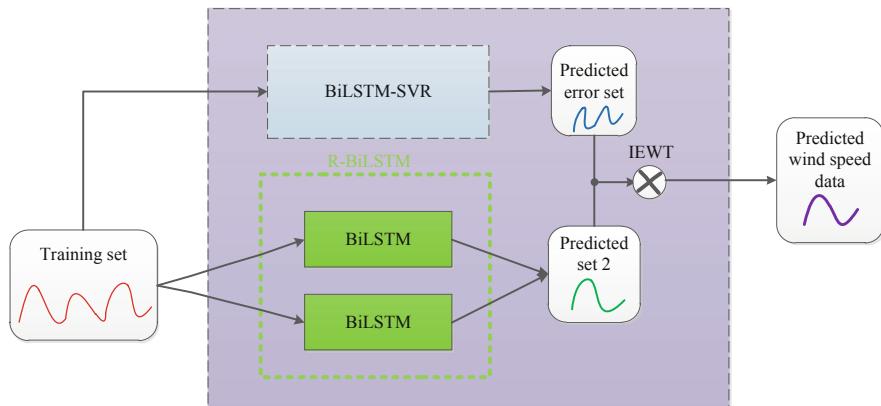


Fig. 3. The structure of wind speed series reconstruction model

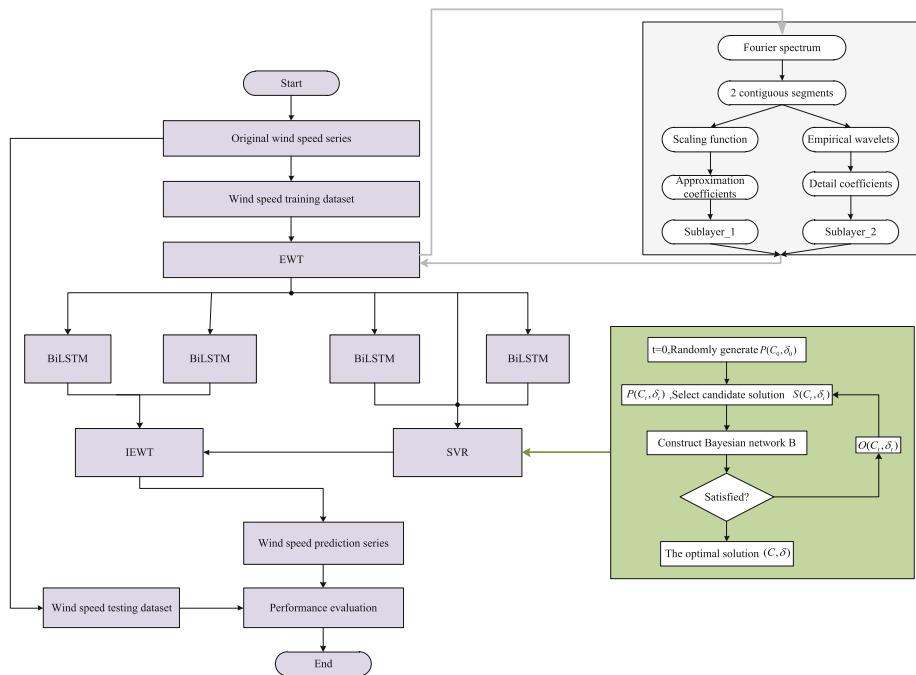


Fig. 4. The flowchart of the proposed hybrid EWT-BiLSTM-SVR model

3 Experiments and Conclusions

3.1 Performance Evaluation Indexes and Comparative Experiments

Four performance evaluation indexes are used in the experiment to evaluate the prediction accuracy of the wind speed prediction model, they are correlation coefficient

(R), mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE), which can be expressed as:

$$R = \frac{\sum_{i=1}^N (\hat{h}(i) - \bar{\hat{h}})(h(i) - \bar{h})}{\sqrt{\sum_{i=1}^N (\hat{h}(i) - \bar{\hat{h}})^2} * \sqrt{\sum_{i=1}^N (h(i) - \bar{h})^2}} \quad (1)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{h}(i) - h(i)| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{h}(i) - h(i))^2} \quad (3)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|\hat{h}(i) - h(i)|}{h(i)} \times 100\% \quad (4)$$

where N represents the total number of wind sequence data, $h(i)$ represents the true wind speed value, $\hat{h}(i)$ represents the average value of predicted wind speed, \bar{h} represents the average value of true wind speed, and $\bar{\hat{h}}$ represents the average value of predicted wind speed.

3.2 Results and Discussion

In order to verify the validity of the hybrid EWT-BiLSTM-SVR model proposed in this paper, the KNN (k-nearest neighbor) model, ARIMA model, SVR model, LSTM model, BiLSTM model and mixed EWT-BiLSTM model are applied to the same wind speed data as the comparison experiment. And at the same time, two sets of wind speed data from Inner Mongolia, China are used in the experiments.

The wind speed results predicted by the hybrid EWT-BiLSTM-SVR model are shown in the Figs. 5 and 6. At the same time, Tables 1 and 2 show the comparison between the proposed hybrid EWT-BiLSTM-SVR model with the other six models.

Table 1. Performance evaluation indexes and comparison (1)

| | MAE (m/s) | RMSE (m/s) | MAPE (%) | R |
|----------------|---------------|---------------|---------------|---------------|
| ARIMA | 0.5166/263.8% | 0.6608/247.6% | 8.7116/321.1% | 0.9127/92.0% |
| KNN | 0.5694/356.3% | 0.7132/267.2% | 9.0828/334.8% | 0.904/91.2% |
| SVR | 0.5132/262.1% | 0.6537/244.9% | 8.2885/305.5% | 0.9135/92.1% |
| LSTM | 0.5183/264.7% | 0.6614/247.8% | 8.9673/330.5% | 0.9107/91.8% |
| BiLSTM | 0.2529/129.2% | 0.3296/123.5% | 4.0715150.1% | 0.9653/97.3% |
| EWT-BiLSTM | 0.2055/105.0% | 0.2753/103.1% | 3.2776/120.8% | 0.9910/99.9% |
| EWT-BiLSTM-SVR | 0.1958/100.0% | 0.2669/100.0% | 2.7132/100.0% | 0.9917/100.0% |

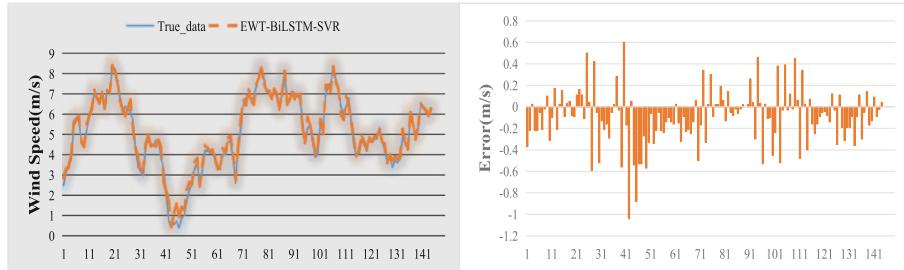


Fig. 5. The original wind speed data and wind speed forecasting results with errors (1)

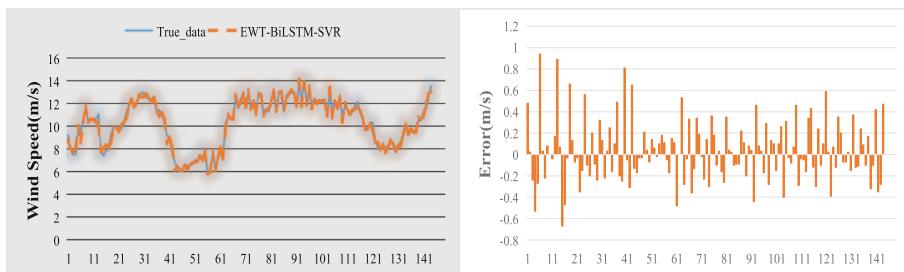


Fig. 6. The original wind speed data and wind speed forecasting results with errors (2)

Table 2. Performance evaluation indexes and comparison (2)

| | MAE (m/s) | RMSE (m/s) | MAPE (%) | R |
|----------------|---------------|---------------|---------------|---------------|
| ARIMA | 0.7192/337.5% | 0.9548/339.9% | 5.3069/370.6% | 0.8811/88.9% |
| KNN | 0.8005/375.6% | 1.0273/365.7% | 6.4713/451.9% | 0.8622/87.0% |
| SVR | 0.7261/340.7% | 0.9635/343.0% | 5.4870/383.2% | 0.8875/89.5% |
| LSTM | 0.7987/374.8% | 1.0786/384.0% | 5.7995/405.0% | 0.8718/88.0% |
| BiLSTM | 0.3229/151.5% | 0.4157/148.0% | 2.3900/166.9% | 0.9584/96.7% |
| EWT-BiLSTM | 0.2141/100.5% | 0.2877/102.4% | 1.5126/105.6% | 0.9909/100.0% |
| EWT-BiLSTM-SVR | 0.2131/100.0% | 0.2809/100.0% | 1.4320/100.0% | 0.9912/100.0% |

From Tables 1 and 2, we can clearly see that the hybrid EWT-BiLSTM-SVR model presented in this paper shows excellent prediction results for different sets of wind speed data. In the first set of wind speed data experiment, $MAE = 0.1958(m/s)$, $RMSE = 0.2669(m/s)$, $MAPE = 2.7132(\%)$, $R = 0.9917$. In the second set of wind speed data experiment, $MAE = 0.2131(m/s)$, $RMSE = 0.2809(m/s)$, $MAPE = 15126 (\%)$, $R = 0.9912$. It can be seen from the above two tables that the performance of different models will be different for different wind speed data, and the model proposed in this paper has better stability. Compared with the traditional statistical model and machine learning model, the hybrid EWT-BiLSTM-SVR model proposed in this paper has a significant improvement in performance, which shows that the prediction performance of the proposed model is better.

3.3 Conclusions

Accurate and stable wind speed forecasting is a complex and difficult task, but it has extremely important practical significance for wind power generation. This paper has proposed a hybrid EWT-BiLSTM-SVR model for wind speed forecasting, which combines EWT, BiLSTM network, SVR, IEWT and Bayesian optimization algorithm. EWT is utilized to decompose the original wind speed data, BiLSTM network and SVR is utilized to predict the decomposed data, including wind speed numerical prediction and error numerical prediction, Bayesian optimization algorithm is utilized to optimize the hyperparameters of SVR, and IEWT is utilized to reconstruct the wind speed prediction data. Two sets of wind speed data are used to compare the proposed hybrid EWT-BiLSTM-SVR model with six other wind speed prediction models: ARIMA model, KNN model, SVR model, LSTM model, BiLSTM model and mixed EWT-BiLSTM model. The experimental results show that the hybrid EWT-BiLSTM-SVR model proposed in this paper is superior to other single prediction models in prediction performance and has better performance in prediction accuracy.

References

1. Liu Q, Lei Q, Xu H (2018) China's energy revolution strategy into 2030. *Resour Conserv Recycling* 128:78–89
2. The Global Wind Energy Council (2016). *Global Wind Report 2015: Annual market update*
3. Lu Z, Li H, Qiao Y (2016) Power system flexibility planning and challenges considering high proportion of renewable energy. *Autom Electric Power Syst* 40(13):147–158
4. Lopes VS, Borges CLT (2015) Impact of the combined integration of wind generation and small hydropower plants on the system reliability. *IEEE Trans Sustain Energy* 6(3):1169–1177
5. Zhu X, Gention MG (2012) Short-term wind speed forecasting for power system operations. *Int Stat Rev* 80(1):2–23
6. Chen J, Zeng G, Zhou W (2018) Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. *Energy Convers Manag* 165:681–695
7. Wang B, Sun Y, Wu W et al (2015) A real-time wind power control method for power systems with high wind power penetration and its implementation. *Autom Electric Power Syst* 39(21):23–29
8. Mahoney WP, Parks K, Wiener G et al (2012) A wind power forecasting system to optimize grid integration. *IEEE Trans Sustain Energy* 3(4):670–682
9. Landberg Lars, Giebel Gregor et al (2003) Short-term prediction - an overview. *Wind Energy* 6(3):273–280
10. Giebel G, Brownsword R, Kariniotakis G, et al (2011) The state-of-the-art in short-term prediction of wind power: a literature overview. *ANEMOS.plus*
11. Cadena E, Rivera W (2009) Short term wind speed forecasting in La Venta, Oaxaca, Mexico, using artificial neural networks. *Renew Energy: Int J* 34:274–278
12. Ramasamy P, Chandel SS, Yadav AK (2015) Wind speed prediction in the mountainous region of India using an artificial neural network model. *Renew Energy: Int J* 80:338–347
13. Catalao JPS, Pousinho HMI, Mendes VMF (2011) Hybrid wavelet-PSO-ANFIS approach for short-term wind power forecasting in Portugal. *IEEE Trans Sustain Energy* 2:50–59

14. Barbounis TG, Theocharis JB (2006) Locally recurrent neural networks for long-term wind speed and power prediction. *Neurocomputing* 69:466–496
15. Dong Q, Sun Y, Li P (2017) A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: a case study of wind farms in China. *Renew Energy* 102:241–257
16. Naik J, Satapathy P, Dash PK (2018) Short-term wind speed and wind power prediction using hybrid empirical mode decomposition and kernel ridge regression. *Appl Soft Comput* 70:1167–1188
17. Längkvist M, Karlsson L, Loutfi A (2014) A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recogn Lett* 42:11–24
18. Bengio Y (2009) Learning deep architectures for AI. *Found Trends Mach Learn* 2:1–127
19. Erhan D, Bengio Y, Courville A et al (2010) Why does unsupervised pre-training help deep learning? *J Mach Learn Res* 11:625–660



A Novel Bearing Intelligent Fault Diagnosis Method Based on Modified VMD and 1-D CNN

Xin Chen^{1,2} and Kaixiang Peng^{1,2(✉)}

¹ 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, People's Republic of China
kaixiang@ustb.edu.cn

² Institute of Artificial Intelligence, University of Science and Technology Beijing, Beijing 100083, People's Republic of China

Abstract. Bearings are one of the most critical components of rotating machinery. Timely and accurate bearing fault detection and diagnosis are very important for the reliability and safe operation of industrial systems. Due to the non-stationary, non-linear, and high-dimensional characteristics of the rolling bearing vibration signal, the signal pre-processing and fault feature extraction are particularly important, because they can directly affect the efficiency and accuracy of fault diagnosis. This paper proposes an improved intelligent diagnosis method combining the modified Variational Mode Decomposition (MVMD) and the 1-D Convolutional Neural Network (1-D CNN). This method first uses the traditional Empirical Mode Decomposition (EMD) to adaptively extract the Intrinsic Mode Function (IMF), and then proposes a sensitive IMFs evaluation index to choose sensitive IMFs and then reconstruct the vibration signal. The VMD iterates over the reconstructed signal and searches for the optimal solution of the variational model, determines the frequency center and bandwidth of each component, and realizes the frequency domain splitting of the reconstructed signal and the separation of the modes. Finally, the processed data is divided into train set and test set, and feature extraction training is performed on the training set through 1-D CNN. The trained CNN network uses the test set to evaluate the accuracy of fault diagnosis. The experimental results show that the method can not only accurately decompose the bearing vibration signal and mine the available fault features, but also classify different fault types.

Keywords: Empirical mode decomposition · Modified variational mode decomposition · Rolling bearings · Fault diagnosis · Convolutional neural network

1 Introduction

Rolling bearings are key components of most mechanical transmissions and are extremely fragile. If the rolling bearing fails, it will cause serious damage to the entire mechanical equipment. Therefore, bearing fault diagnosis has important significance and is attracting more and more attention. Feature extraction based on vibration signals

has been proven to be an effective method for solving rotational machinery fault diagnosis problems because the vibration signals collected from the sensors contain rich fault-related information [1, 2].

As a data processing tool for describing non-stationary signals, Wavelet Trans-form (WT) [3] has been widely used in mechanical equipment condition monitoring and fault diagnosis because it has the characteristics of multi-resolution analysis, which is useful for identifying weak fault features from noisy signal. However, WT is still limited by the specified binary subdivision scheme, and sometimes the transient bearing fault vibration features cannot be recognized. In order to avoid the problem of binary subdivision, Huang et al. [4] proposed an adaptive time-frequency analysis algorithm namely Empirical Mode Decomposition (EMD). In [5], a suitable indicator was introduced to improve the traditional EMD and Hilbert Transform (HT), which can automatically select sensitive IMF for fault diagnosis. Du and Yang [6] replaced the EMD-based average with a traditional envelope average to diagnose bearing faults. Although the EMD method has been widely adopted, since the method belongs to recursive mode decomposition, exists modal mixing and end effect, and the decomposition error is large. In 2014, Dragomiretskiy et al. [7] proposed a new method of signal processing, “Variational Mode Decomposition (VMD)”, which determines the frequency center and bandwidth of each component by iteratively searching for the optimal solution of the variational model. The frequency domain splitting of the signal and the effective separation of the components can be adaptively implemented. Zhang et al. [8] established a bearing fault signal calculation model with different position defects through failure mechanism analysis, and extracted the main mode of the fault signal through VMD, and successfully used it for the composite fault diagnosis of rolling bearings. In [9], an improved VMD method based on kurtosis is proposed, and the envelope spectrum of the selected transient pulse is analyzed to diagnose the faulty bearing. All of the above methods prove that EMD and VMD are effective processing methods for complex signals.

When the mechanical system is very complex, manual selection of features or optimization methods based on manual selection requires a considerable amount of mechanical expertise and profound mathematical knowledge, which has great limitations. It is very important for bearing vibration fault diagnosis to automatically extract the fault characteristics. Therefore, intelligent diagnostic methods based on Back Propagation (BP), Artificial Neural Network (ANN) or Deep Neural Network (DNN) are gaining more and more attention [10], the learning process usually includes three main steps: data preprocessing, feature extraction and selection, and fault identification [11, 12]. Convolutional Neural Network (CNN) is a supervised deep learning algorithm widely used in the field of pattern recognition [13]. CNN is different from other deep learning algorithms, it has the three prominent features: local perception field, weight sharing and pooling. Not only does it reduce the complexity of the network, but also reduces the risk of overfitting, making it possible to build a deep learning framework that processes massive amounts of data. Shao et al. [14] first used automatic encoder to compress the da-ta, and then used Gaussian visible unit to construct a new deep convolution belief network learning fault feature. Wen et al. [15] converted the signal into a two-dimensional image and performed fault feature extraction and fault diagnosis through a new CNN network based on LeNet-5, obtained

good results. However, the traditional 2-D CNN method still has a major problem in bearing fault diagnosis: the time series represented by the one-dimensional matrix is used as the input of the network, and does not match the convolutional network of 2-D structure, causes slow training speed and poor classification results. Based on this problem, 1-D Convolution Neural Network (1-D CNN) was proposed. Levent et al. [16] directly used the original time domain vibration data as input, integrated fault feature extraction and fault classification into the CNN network, and obtained better results than the 2-D network.

Based on the above research, this paper will use a modified VMD (MVMD) method based on EMD and VMD to preprocess the original vibration signal of the bearing and decompose the various modes. Then, the 1-D CNN is used to extract the deep features of the preprocessed data, and the test data is analyzed for the accuracy of fault diagnosis.

2 Proposed MVMD-1DCNN Method

2.1 The Principle of MVMD

Variational Mode Decomposition (VMD) is a new adaptive and quasi-orthogonal signal processing method based on Wiener filtering, one-dimensional Hilbert Transform and heterodyne demodulation. Different from EMD, the core idea of the VMD method is to assume that most of each mode is tightly around a certain center frequency, and then transform the problem of solving the bandwidth of mode into a constrained optimization problem [7].

The bearing vibration signal contains a lot of noise and it is difficult to select a reasonable mode number for the VMD. Therefore, this paper firstly uses the EMD method to adaptively decompose the original vibration signal, separate the vibration fault feature from the noise, and then propose a new sensitive IMF evaluation index, determine the number of sensitive IMF and reconstruct the noise reduction signal. Finally, the reconstructed signal is again modally decomposed by VMD. The specific implementation process of the MVMD is shown in Fig. 1:

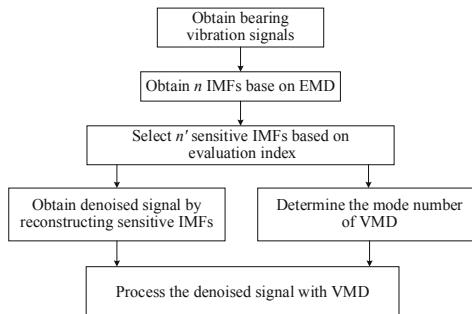


Fig. 1. Flow chart of the MVMD

First, the n IMFs are obtained by the EMD, then, an index combining the skewness, the correlation between the original signal and the IMF is used to select the sensitive IMF. The specific calculation process is as follows:

- (1) Calculate the correlation coefficient $r_{c_i,x}$ between the i th IMF and the original signal $x(t)$:

$$r_{c_i,x} = \frac{1}{n} \left\| \frac{\sum_{t=1}^n (x(t) - \bar{x})(c_i(t) - \bar{c}_i)}{\sigma_{c_i} \sigma_x} \right\| \quad (1)$$

where $c_i(t)$ is the i th IMF; σ_{c_i} and σ_x are the standard deviations of $c_i(t)$ and $x(t)$, respectively, $\sigma_{c_i}^2 = \sum_{t=1}^n (c_i(t) - \bar{c}_i)^2$, $\sigma_x^2 = \sum_{t=1}^n (x(t) - \bar{x})^2$; \bar{c}_i and \bar{x} are the average values of $c_i(t)$ and $x(t)$, respectively

- (2) Calculate the skewness of each IMF:

$$b_{c_i} = \frac{\frac{1}{N} \sum_{i=1}^N (c_i(t) - \bar{c}_i)^3}{\left(\frac{1}{N-1} \sum_{i=1}^N (c_i(t) - \bar{c}_i)^2 \right)^{\frac{3}{2}}} \quad (2)$$

where N is the length of $c_i(t)$.

- (3) Design IMF Sensitive Evaluation Index (SEI):

$$I_{SEI}(i) = \frac{r_{c_i}}{\sum r_{c_i}} + \frac{b_{c_i}}{\sum b_{c_i}} \quad (3)$$

- (4) Sort n IMFs by $I_{SEI}(i)$, Select the first n' IMFs with the largest SEI as sensitive components and use them to reconstruct a denoised signal $x_{re}(t)$:

$$x_{re}(t) = \sum_{k=1}^{n'} c_k(t) \quad (4)$$

The process of reconstructing the denoised signal is actually a process of detecting whether the sensitive IMF is fault-relate information.

2.2 1-D Convolution Network Training Process

The CNN processing steps can be summarized as follows: Firstly, get the IMF data from the previous section and divide it into train set and test set; secondly, determine the number of convolutional and pooling layers of the CNN, determine the number of filters (the number of feature maps) and the pooling mode of each pooling layer, then

input the train set into the network to train MAX_STEP times; thirdly, The training process adopts the form of “batch”, that is, each step captures a subset input network from the training set for training, and then adds a fully connected output layer, fine-tunes network parameters according to the health label of the sample, completes CNN training; Finally, test samples are used to fault diagnosis and healthy classification.

The structure of the 1-D CNN used in this paper includes 3 convolutional layers, 3 pooling layers, and 2 fully connected layers, The number of feature maps of the three convolutional layers is 16, 32, and 64, respectively; The pooling method is max pooling, and the number of neurons in the two fully connected layers is 64 and 7, respectively.

2.3 Framework of Proposed MVMD-1DCNN Method

In order to separate the fault features and reduce the influence of noise, make the features easier to be extract and classified, this paper proposes a MVMD method to denoise and preprocess the bearing data and decompose the modes of the signal. Then, combined with the one-dimensional characteristics of the bearing vibration time domain data, the faulty feature extraction and diagnosis classification of the processed samples are performed using the 1D CNN mentioned above. The overall framework of this paper is shown in Fig. 2. The specific steps are summarized as follows: (1) Obtain Bearing fault data through the Case Western Reserve University Bearing Data Center. (2) The original time domain data is mo-dally decomposed using the traditional EMD method, and n original IMFs are obtained. (3) A new IMF sensitive evaluation index is constructed to select the n' IMFs that are most correlated with the original signal and reconstruct the bearing signal data without low frequency noise. (4) The number of modes of the VMD is determined to be n', and the reconstructed signal is subjected to VMD decomposition to obtain a new set of IMFs. (5) The IMF data in (4) is compressed and re-shaped, and the train set and test set are divided. (6) The convolutional layer, the pooling layer, and the fully connected layer construct a 1-D CNN, which is used for train set learning, and then the learned deep features are input to the fully connected layer for fault classification. (7) Test set is used to verify the performance of the proposed method.

3 Experimental Validation

3.1 Data Processing

The vibration data used in this paper is from the Case Western Reserve University Rolling Bearing Data Center [17]. This paper selects 7 different types of fault data to verify the accuracy of fault identification. The bearing vibration data is recorded on the drive end by the acceleration sensor, and the sampling frequency is 12 kHz. Each sample contains 1000 sample points.

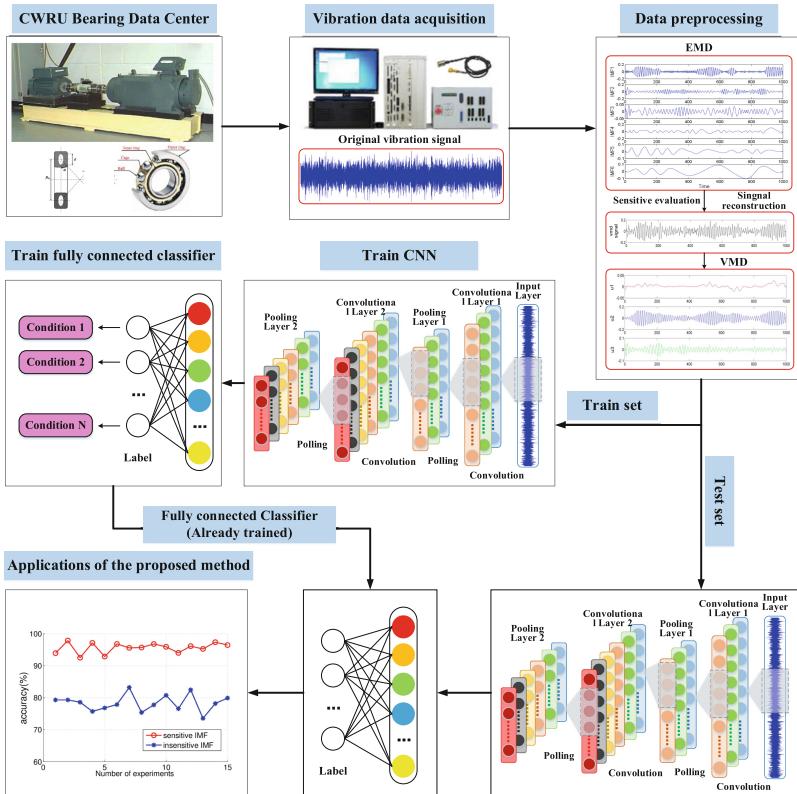


Fig. 2. Framework of proposed method

The vibration signal collected by the bearing is first decomposed into several different IMFs by EMD and the six IMFs are shown in Fig. 3(a). The Fig. 3(b) shows the processed data according to sensitive IMFs; the Fig. 3(c) shows the processed data according to non-sensitive IMFs. It can be seen from the figure that the difference between (b) and (c) is obvious. Since (b) contains most of the high-frequency bearing vibration signal components, the data details are more abundant, the periodicity is also very obvious.

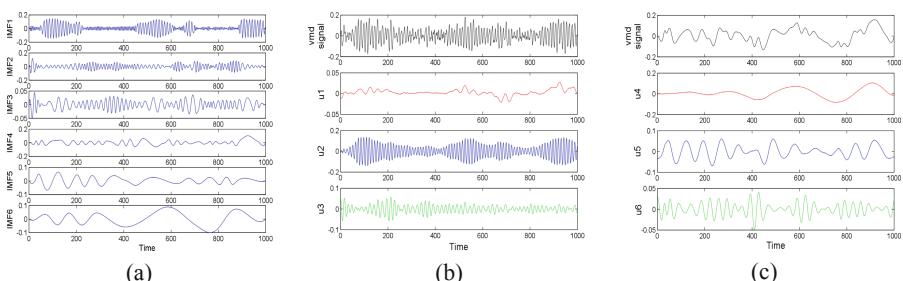


Fig. 3. Processed data

3.2 Experimental Result

After data processing, three sets of comparative experiments were designed to observe the relationship between the number of training samples and the accuracy of fault diagnosis. The raw vibration data of the bearing is directly sent to the CNN for training without any processing, and the final fault classification result is shown in Fig. 4(a); (b) and (c) represent the results of inputs sensitive IMFs and non-sensitive IMFs, respectively. It can be seen that the fault identification accuracy of the three test data increases as the number of samples increases. However, tests based on sensitive IMFs are clearly superior to the other two data in accuracy results.

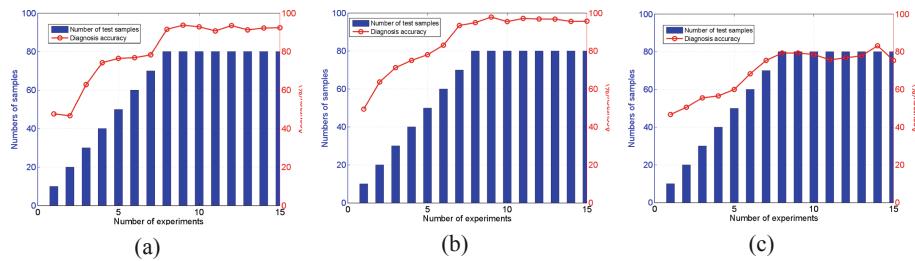


Fig. 4. Number of training samples and identification accuracy relationship

Then, continue to use 80 training samples for 15 times to verify the stability and accuracy of the proposed method. The results are shown in Table 1:

Table 1. Diagnostic accuracy of 15 experiments with 3 data

| Data | Minimum accuracy (%) | Maximum accuracy (%) | Average accuracy (%) | Standard deviation |
|-------------------|----------------------|----------------------|----------------------|--------------------|
| Original data | 90.76 | 93.75 | 92.10 | 0.0102 |
| Sensitive IMF | 94.93 | 97.86 | 96.24 | 0.0086 |
| Non-sensitive IMF | 73.53 | 83.21 | 78.34 | 0.0261 |

The above results show that sensitive IMFs contain more signal features and less noise than non-sensitive IMFs. Its fault features are easier to extract than the original signal, which resulting in better experimental results for sensitive IMFs.

In order to verify the superiority of the 1-D CNN proposed for fault diagnosis of bearing vibration data, a single hidden layer Back Propagation (BP) network and multiple hidden layers Autoencoder (AE) networks were selected for comparison. The experimental results are shown in Table 2 below:

Figure 5 below shows the training error curves of the three methods in the first test. After 1000 iterations, the training error of proposed method is 0.0001943, the BP network's is 0.00271686, and the AE network's is 0.13278988. Compared with the proposed method, the BP network and the AE network have larger training errors, the

network training cannot be completed within a limited number of steps, resulting in a lower accuracy of the final fault diagnosis.

Table 2. Diagnostic accuracy of 15 experiments with 3 networks

| Network | Minimum accuracy (%) | Maximum accuracy (%) | Average accuracy (%) | Standard deviation |
|------------|----------------------|----------------------|----------------------|--------------------|
| 1-D CNN | 94.93 | 97.86 | 96.24 | 0.0086 |
| BP network | 25.21 | 70.15 | 42.56 | 0.1297 |
| AE network | 54.52 | 65.35 | 59.67 | 0.0320 |

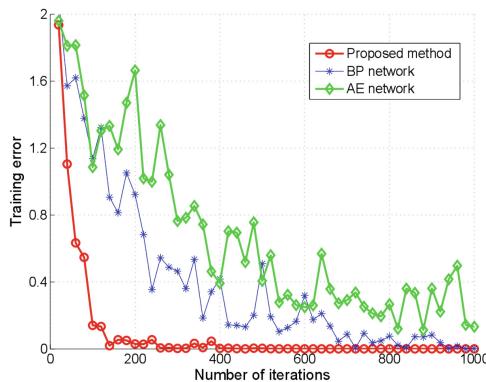


Fig. 5. Training error curves of three networks

4 Conclusion

This paper analyses the nonlinear and non-stationary characteristics of bearing vibration signals, combines the optimal characteristics of traditional EMD and VMD and considers the adaptability of 1-D CNN to one-dimensional vibration signals, an improved intelligent diagnosis method for bearing faults of MVMD combined with 1-D CNN (MVMD-1DCNN) is proposed. The experimental results show the effectiveness of the proposed method.

References

- Li YF, Liang XH, Zuo MJ (2017) Diagonal slice spectrum assisted optimal scale morphological filter for rolling element bearing fault diagnosis. *Mech Syst Sig Process* 85:146–161
- Chen J, Pan J, Li Z et al (2016) Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals. *Renew Energy* 89:80–92

3. Huo ZQ, Zhang Y, Francq P et al (2017) Incipient fault diagnosis of roller bearing using optimized wavelet transform based multi-speed vibration signatures. *IEEE Access* 5:19442–19456
4. Huang NE, Shen Z, Long SR et al (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc Roy Soc A* 454 (1971):903–995
5. Ricci R, Pennacchi P (2011) Diagnostics of gear faults based on EMD and automatic selection of intrinsic mode functions. *Mech Syst Sig Process* 25(3):821–838
6. Du QH, Yang SN (2007) Application of the EMD method in the vibration analysis of ball bearings. *Mech Syst Sig Process* 21(6):2634–2644
7. Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. *IEEE Trans Sig Process* 62(3):531–544
8. Zhang M, Jiang ZN, Feng K (2017) Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump. *Mech Syst Sig Process* 93 (1):460–493
9. Jiang F, Zhu ZC, Li W (2018) An improved VMD with empirical mode decomposition and its application in incipient fault detection of rolling bearing. *IEEE Access* 6:44483–44493
10. Jia F, Lei YG, Lin J et al (2016) Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech Syst Sig Process* 72:303–315
11. Jia F, Lei YG, Guo L et al (2018) A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines. *Neurocomputing* 272:619–628
12. Lei YG, Jia F, Lin J et al (2016) An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Trans Ind Electron* 63(5):3137–3147
13. LeCun Y, Bottou L, Bengio Y et al (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
14. Shao HD, Jiang HK, Zhang HZ et al (2017) Electric locomotive bearing fault diagnosis using novel convolutional deep belief network. *IEEE Trans Ind Electron* 65(3):2727–2736
15. Wen L, Li XY, Gao L et al (2017) A new convolutional neural network-based data-driven fault diagnosis method. *IEEE Trans Ind Electron* 65(7), 5990–5998
16. Levent E, Turker I, Serkan K (2018) A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier. *J Sig Process Syst* 91(2):179–189
17. Loparo KA (2003) Bearings vibration data set, Case Western Reserve University. <http://www.eecs.cwru.edu/laboratory/bearing/download.htm>



The Design and Implementation of Simple CoreXY Structure Writing Robot

Rongchuan Yu¹, Yahang Qin¹, Jiansheng Peng^{2(✉)}, Tan Guo¹,
and Xiaoping Tang¹

¹ Guangxi Science and Technology Normal University, Laibin, China

² He Chi University, Hechi, China

sheng120410@163.com

Abstract. The control system of simple writing robot was designed based on the CoreXY structure. In order to realize the writing function, built the system hardware framework, the two methods of Chinese character strokes and picture characters were designed. The control interface of the system was designed by using MATLAB GUIDE toolbox. Improved genetic algorithm was introduced to improve the apparent jitter problem existing in the system, and finally complete the writing task. Experiments show that the CoreXY structure of the writing robot can complete the writing of Chinese characters and pictures depicting, which verified the design of CoreXY structure writing robot is feasible.

Keywords: CoreXY structure · Writing robot · Improved genetic algorithm

1 Introduction

Writing robot is the main research object of robot teaching, which is of great significance to the development of robot teaching field. More and more training institutions began to educate young people in the robot with the continuous development of the robot education industry [1, 2]. At present, the research on writing robot mainly based on the existing font image or through the external input of the standard image of the image recognition and extraction and to get the start point through the decomposition of font image or the vectorization of the picture. The design of the way to write fonts directly through the external “WordPad” or a combination of a variety of input had not been found [3, 4]. This topic is the realization of CoreXY structure writing robot. The writing robot can be used as a teaching aid to enhance students’ awareness of the robot and improve students’ interest in robotics.

2 System Overall Design

2.1 CoreXY Structural Design

CoreXY structure is a basic components which provides a number of computerized production tools (Cartesian movement) with a simple and adaptable format. CoreXY

main problem is how to effectively achieve the X-axis and Y-axis linkage. Its structure is shown in Fig. 1. The X-axis motor is fixed and the Y-axis is driven by the X-axis. When the X and Y axes of the motor at the same time movement, the slider can be moved to any point within the allowable range of the hardware frame. In this design, CoreXY structure writing robot is based on this principle to achieve. The motion equation of the CoreXY structure is given by formula (1).

$$\Delta A = \Delta X, \Delta B = -\Delta Y \quad (1)$$

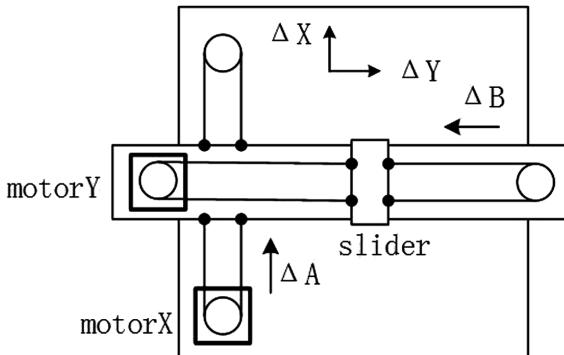


Fig. 1. CoreXY structure principle diagram.

2.2 CoreXY Structure Frame Structure

In this subject, to build the whole structure of CoreXY adopts 10 mm diameter guide and aluminum alloy bracket as the skeleton, according to CoreXY structure. The overall structure is relatively simple, reducing the weight of moving parts, can be appropriate to improve the motor acceleration and exercise convenience.

As shown in Fig. 2, the main components of the writing robot include: two permanent magnet motors, two motor controllers, an angle steering gear. The X-axis motor is fixed on the frame of the CoreXY structure. The Y-axis motor is fixed in the X-axis direction by some fixed parts and moves as the X-axis moves. The maximum stroke of the X-axis is 23 cm, the maximum stroke of the Y-axis is 21 cm, the diameter of a point is 1 mm, and the size of a Chinese character is 25 mm × 25 mm.

With an angle steering gear to control the movements of writing pen and the angle of the steering gear is directly controlled by the I/O port of the XMC4500 chip. It is only necessary to control the level time of the I/O to complete pen up and pen down.

Since a controller provides only one motor interface, two controllers are required. The motion controller is capable of two-axis closed-loop control. Timer can achieve variable pulse output, drive the servo signal has three kinds, namely the direction + pulse signal, a difference of 90° of dual pulse signal and single pulse signal. The

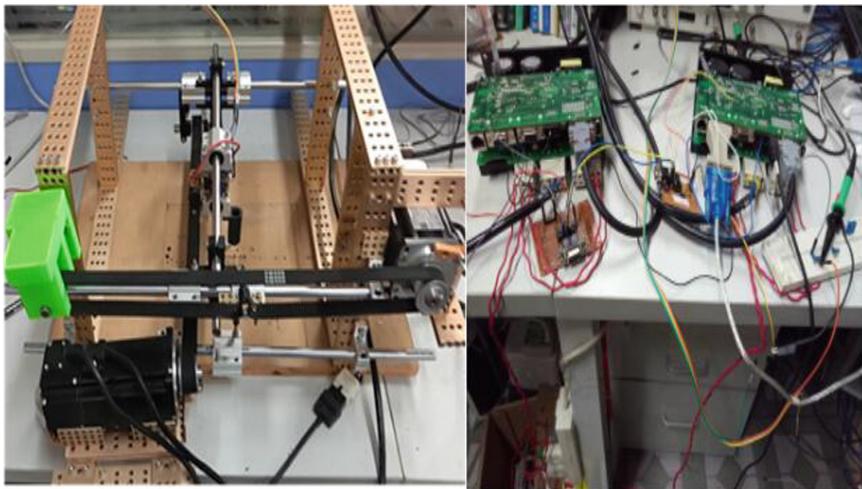


Fig. 2. Hardware framework diagram.

motion controller can realize the above three pulse drive modes. Single-axis signal has a servo drive signal, servo enable signal, brake signal, zero signal [5, 6]. System hardware interface diagram shown in Fig. 3.

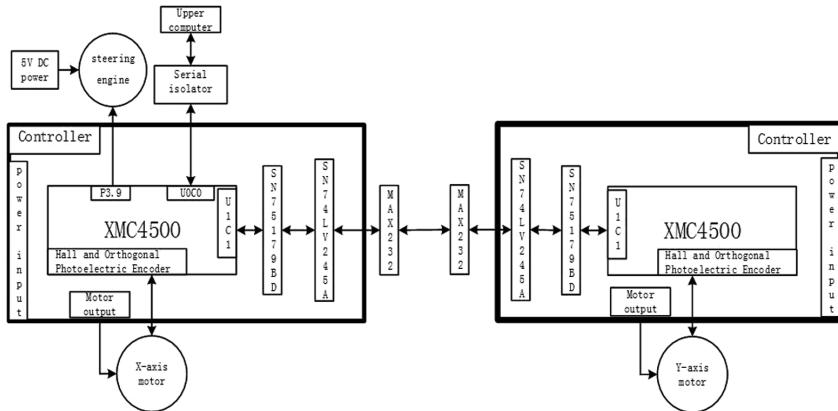


Fig. 3. Schematic diagram of the system hardware interface.

The motion controller has perfect control function. RS232 for the host computer and the lower computer communication, control the X-axis movement and control Y-axis movement of the Infineon XMC4500 microcontroller through the UART to RS422 circuit between the two controllers to achieve data transmission, and respectively According to the received instructions to achieve the motor running and stopped.

2.3 Writing Program Design

- Chinese Character Strokes Design:** In this design, with 36×36 matrix to store a Chinese character. Such as “天”, as shown in Fig. 4, defined from left to right is the positive direction of the X axis, from top to bottom is the positive direction of the Y axis. First split it into four strokes “horizontal - horizontal - write - Na”, as shown in Fig. 5, and then store the four strokes in order. Once you have finished writing a stroke, store once. Each stroke is processed after all the strokes are stored. Extract the coordinates of the starting point, midpoint, and end point of each stroke. The operation of the strokes are handled in the host computer, including the identification of strokes and the extraction of coordinate points (0 represents a writing, 1 on behalf of no handwriting). For the identification of strokes, writing in the tablet, will record the coordinates of the pen down (starting point) and pen up (end point). If the coordinates of the Y-axis direction remain constant or change within a certain range (the defined range is positive and negative 2) and the X-axis coordinates have been increasing, then the stroke is “horizontal”. If the coordinates of the X-axis direction remain constant or within a certain range of small changes, and the Y-axis coordinates have been increasing, then the stroke is “vertical”. If the coordinates of the Y-axis direction are increasing, and the coordinates of the X-axis are also increasing, then this stroke is “Na”. The operation of extracting coordinate of Strokes is relatively simple, in all coordinate points of each stroke, through the comparison to find the suitable coordinate points. After finding the coordinates of the three points, then need to let X, Y axis to realize synchronization. Such as a stroke, its X-axis coordinates are 6, Y-axis coordinates are 20, each move a coordinate on the need for a pulse, X-axis will need 6 pulses, Y-axis needs 20 pulses. Then X, Y-axis motor from the starting point at the same time, how to achieve the next coordinate point at the same time needs to be resolved. Design ideas: By comparing the relative distance between the X and Y axes of the two points (starting point and midpoint, midpoint and end point). The rounding and removal operations will be implemented between relative distance of the numerical value and the relative distance of the small value, and then solve the X-axis and Y-axis movement of the corresponding number of pulses. A stroke requires the Y axis to move n pulses, requiring the X axis to move m pulses ($m < n$), Such as the formula (2) and the formula (3).

$$n/m = a \quad (2)$$

$$n \% m = b \quad (3)$$

Formula (2) is n to m rounding, and formula (3) is n to removal. To ensure the accuracy of the moving coordinates, the Y-axis moving pulse and the X-axis moving pulse are divided into two parts. The first part of the Y axis moves $(a - 1)$ pulses each time, the first a pulse when the X-axis moves, Y axis total moves $[a \times (m - 1)]$ pulses. The second part of the Y axis needs to move $(a + b - 1)$ pulses, the X axis moves a pulse when the Y axis moves the first $(a + b)$ pulses.

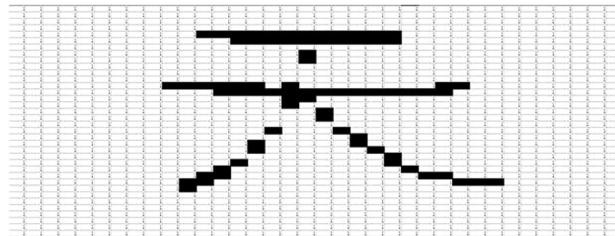


Fig. 4. The matrix of Chinese “天”

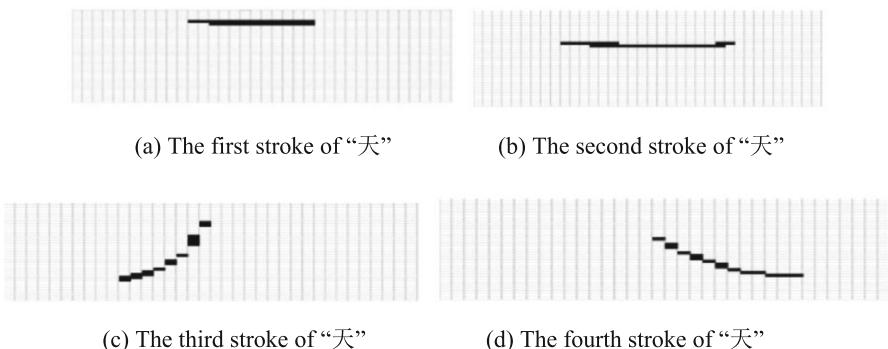


Fig. 5. The decomposition of Chinese “天”.

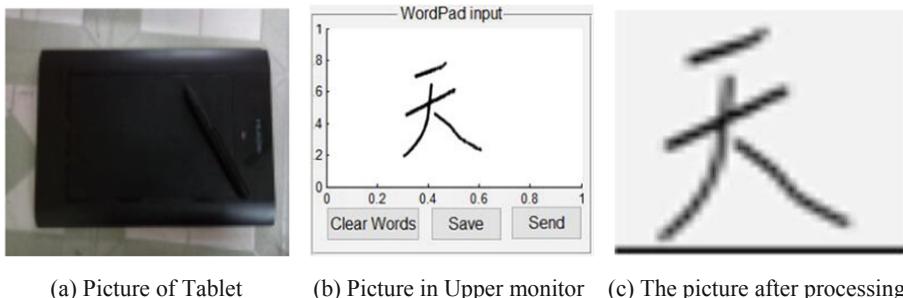
- (2) **Text Picture Design:** In this method, first of all, change the Chinese character into picture format, and then binarize the picture, so as to obtain the entire picture of the binary matrix [7]. And then send the matrix to the controller to control the motor to complete the picture depiction by line-by-line scan. This way to support the Chinese character input, picture input and tablet input, it is more flexible, and the user's operation is simple, just click ‘Send’ when you have finished typing. For example, the user entered a word “天”, as shown in Fig. 6(a), the host computer program will automatically generate a picture, and then through the program from the original image to cut out the “天”, as shown in Fig. 6(b), finally binarize, the resulting picture shown in Fig. 6(c), more Chinese characters input processing flow is similar.



(a) Original picture (b) Picture after interception (c) Picture after binarization

Fig. 6. User input.

Users can also connect the WordPad to the computer, you can see what you write or draw on the WordPad in the host computer in real time. As shown in Fig. 7(a) is a WordPad, Connect to the computer via a data cable. Figure 7(b) is the input through WordPad. Save it in image format, in order to facilitate the storage and transmission of data, you need to compress the saved image. As shown in Fig. 7(c) is the image after processing. The image size is processed as a 36×36 dot matrix, which is beneficial for the controller to receive a unified size of the data and handle.



(a) Picture of Tablet

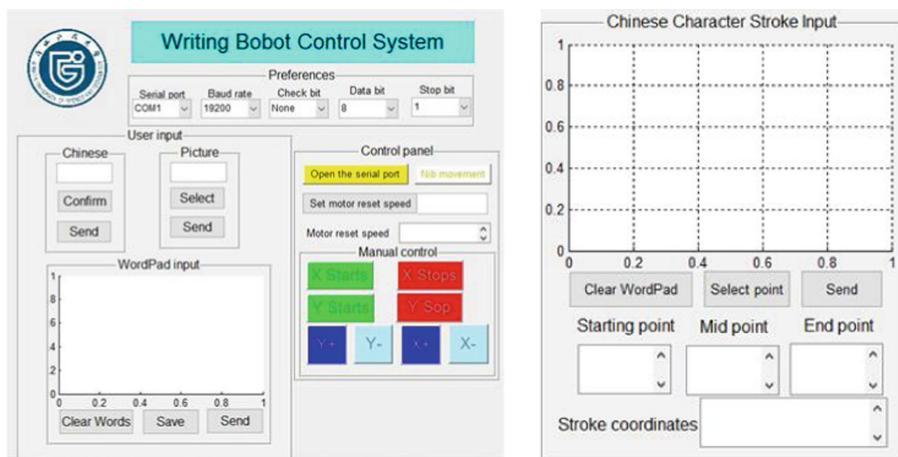
(b) Picture in Upper monitor

(c) The picture after processing

Fig. 7. Tablet input diagram.

2.4 Host Computer Platform Design

The host computer is shown in Fig. 8, Fig. 8(a) diagram contains parameter settings, control panel and user input module, mainly to achieve the input of Chinese characters or pictures. Figure 8(b) is a sub-bit machine, mainly to achieve the input of Chinese characters strokes, identify strokes, remove coordinates of each stroke in the 36×36 lattice, send the writing information of Chinese characters to the controller.



(a) The main upper monitor

(b) The vice upper monitor

Fig. 8. Upper monitor.

3 Improvement and Implementation of Control Algorithm for Writing Robot

In this design is the use of permanent magnet synchronous motor, in the course of running found that the motor will produce significant jitter. After testing, it is necessary to change the parameters of the PI regulator, and the jitter problem will be improved. It is time-consuming to change the parameters of the PI regulator, and the effect is not obvious. The proportional parameter (K_p) and the integral parameter (K_i) is difficult to determine, so the genetic algorithm is introduced to optimize the parameters of the PI regulator [8–10]. In order to improve the accuracy of parameter optimization, based on the traditional genetic algorithm and improved it, hoping to get more accurate parameters of PI regulator.

This algorithm optimizes the PI controller control parameters and determines the parameters to improve the genetic algorithm to solve the optimal parameters. Determine whether the parameters meet the optimal conditions, if the optimal condition is satisfied, the flow of the algorithm is ended and the current optimal result is output. If it is not satisfied, the evaluation of the fitness will be repeated and the genetic operation will be repeated, that is, the optimization process of the parameters is repeated until the optimal condition is satisfied.

The improvement of genetic algorithm mainly includes two aspects: initial population and mutation probability. Initializing the population is to determine the operating space of the algorithm, and the mutation probability is to optimize the operation of the individual within the space.

3.1 Initialize the Population Improvement

In the initial population, the general practice is randomly generated, this approach will lead to access to the objective function of the information is too low. In order to avoid this problem, we first use the Monte Carlo sequence to generate a set of random numbers, and then use the knowledge of matrix to extend this set of random numbers into matrices that meet our dimension.

The population size set in the algorithm is 30, then a set of random numbers A generated by the Monte Carlo sequence is a 30×1 matrix. To achieve the expansion of the dimension, the A matrix is first transposed, and then the matrix A is multiplied by Transpose matrix A' to get a 30×30 matrix B as shown in Eq. (4), B matrix is the initial population.

$$B = A * A' \quad (4)$$

Because the individual produced by the Monte Carlo sequence has a property of variability, this reduces the randomness of the population and increases the uniformity to obtain the information of the objective function. This approach can take care of randomness and uniformity. Specific steps are shown in Fig. 9.

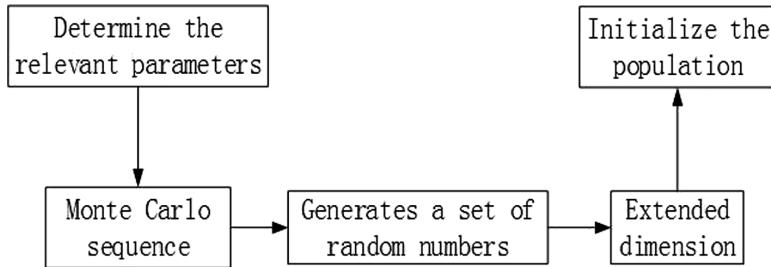


Fig. 9. Improved initialization population.

For the PI controller, the first estimate of the PI range of two parameters, after a certain degree of mathematical processing, the genetic algorithm to achieve the entire feasible parameters within the scope of the search. Before and after the improvement of the algorithm to take 10 sets of simulation data as a contrast, and then select the best group to observe the step response curve, In the traditional genetic algorithm, $K_p = 1.6602$, $K_i = 3.2617$, Initialize the improved genetic algorithm in $K_p = 1.5625$, $K_i = 2.6758$, initialization improved contrast is shown in Fig. 10.

In Fig. 10, the dashed line is the curve before the improvement, and the solid line is the improved curve. It can be seen that after the improvement of the initial population, the optimized result is achieved on the overshoot, but it is not obvious, need to continue to improve the genetic algorithm.

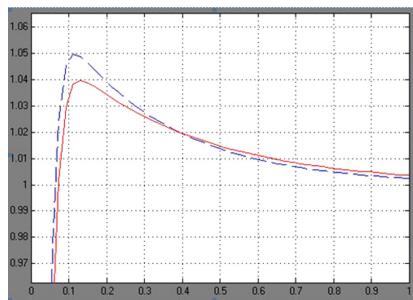


Fig. 10. Initialization improved contrast.

3.2 Improvement of Variation Operations

In the traditional genetic algorithm, it is common to select a fixed mutation probability to mutate all the individuals of the population, which will be eliminated by the individuals with higher fitness and individuals with low fitness. In this paper, we use the method of segmentation to select the variation probability, divide all the individuals into three parts according to the fitness degree, and then select a mutation probability

for each part to carry on the mutation operation. This approach will largely retain the good individuals in the population and can produce better individuals. The variance probability segmentation formula is given by (5).

$$p_m = \begin{cases} p_{m2} + \frac{(p_{m1}-p_{m2})(Bestfi-f_i)}{Bestfi-f_{32}} & f_i \geq f_{32} \\ p_{m3} + \frac{(p_{m2}-p_{m3})(Bestfi-f_i)}{Bestfi-f_{31}} & f_{32} \geq f_i \geq f_{31} \\ p_{m1} & f_{31} \geq f_i \end{cases} \quad (5)$$

Where $p_{m1} = 0.6$, $p_{m2} = 0.05$, $p_{m3} = 0.001$, $Bestfi$ is the fitness of the best fit in the population, f_i is the fitness of the individual to be mutated, f_{31} and f_{32} are fitness Value from high to low in the 1/3, 2/3 of the individual fitness value.

From Eq. (5), it can be seen that the probability of individual variability is low, and it is easy to mutate to produce new individuals, and the higher the fitness is not easy to be mutated, the probability of variation is low, so the better individuals can be better preserved.

The improved genetic algorithm is used to obtain the parameters of the PI regulator $KP = 1.5625$, $Ki = 1.1621$, and the traditional genetic algorithm PI regulator parameters $KP = 1.6602$, $Ki = 3.2617$, the simulation comparison of two algorithms are shown in Fig. 11.

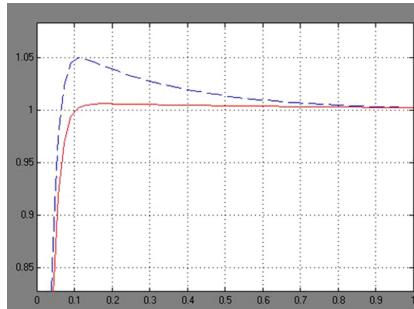
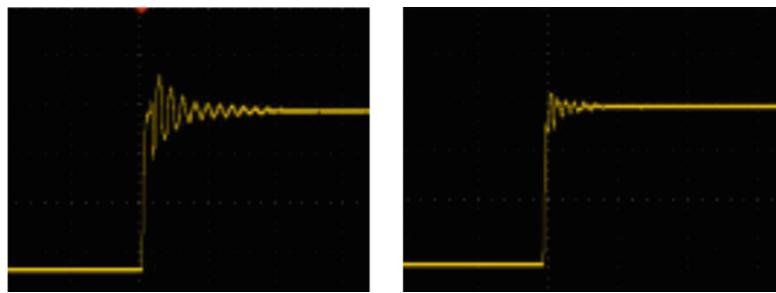


Fig. 11. Simulation comparison of two algorithms.

It can be seen from Fig. 11 that the unit step response of the genetic algorithm with improved mutation probability is superior to the traditional genetic algorithm in the overshoot, and it is earlier than the traditional genetic algorithm in steady state time.

The effect of the motor control effect before and after the introduction of the improved genetic algorithm is tested. The influence of the output waveform analysis algorithm on the motor is observed by the oscilloscope. In Fig. 12(a) shows the output waveform before the PI controller is optimized. Figure 12 (b) is an output waveform optimized for the PI controller.



(a)The waveform before improving

(b) The waveform after improving

Fig. 12. Output waveform comparison.

It can be seen from the output waveform that the output waveform of the PI controller optimized by the improved genetic algorithm on the overshoot and response time has a better effect than the output waveform before optimization, the overshoot is reduced, the response time is accelerated, this is similar to the result of the simulation. After the improved motor operation is also relatively smooth.

Combined with the PC in the manual control module to test the PI parameter before and after the improvement of the system output, Respectively, from the no algorithm, coupled with the traditional genetic algorithm and improve the genetic algorithm three aspects to observation system output, Table 1 shows the comparison of output results.

Table 1. Comparison of output results

| Algorithm | Task | Result |
|-------------------------------|-------------|--------|
| None | Draw a line | |
| Traditional genetic algorithm | Draw a line | |
| Improved genetic algorithm | Draw a line | |

According to Table 1 can be obtained, when no algorithm is added, the system output results in the beginning of the line segment jitter is obvious, and there is a slight pause phenomenon; After adding the traditional genetic algorithm, the result of the system output is only at the beginning of an upward shift phenomenon, in the follow-up process has a slight jitter; After adding the improved genetic algorithm, the output of the system has obvious improvement in jitter, but the line thickness is obviously inconsistent. From the above results can be drawn, the design of the algorithm on the system output has some improvement, but need to continue to improve.

4 Experiments

The experiment of the writing robot system mainly from two aspects, on the one hand test whether the system in accordance with the stroke order to write Chinese characters, on the other hand is the picture content can be written or painted out.

- (1) In the sub-bit machine to write some simple Chinese characters, after the extraction of Chinese characters, sent to the next bit machine to complete the writing of Chinese characters. The experimental results of Chinese character stroke design are shown in Fig. 13.

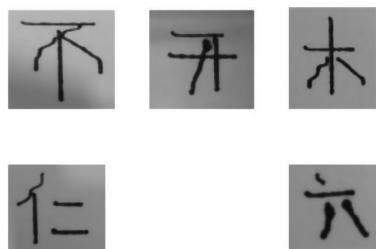


Fig. 13. Experiment results.

Through the experiment, you can verify that the writing robot can basically complete the writing of four strokes of Chinese characters, from the effect point of view, the main problems: First, the hardware structure of the problem, this will lead to Chinese characters appear skew, the second is identified in the next bit Time error occurs, and the next bit in the handling of X, Y-axis motor using the method. In the follow-up study, mainly for the mechanical structure and processing algorithms to optimize, and ultimately improve the entire system.

- (2) In the main host computer input method input Chinese characters, you can also use the mouse to write Chinese characters or draw pictures, the input content processed into binary lattice sent to the next bit machine to complete the final writing or picture drawing. The experimental results of the text picture mode are shown in Figs. 14, 15 and 16.

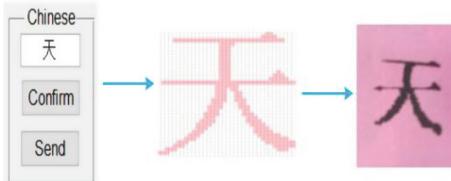


Fig. 14. Chinese analog output and actual output.

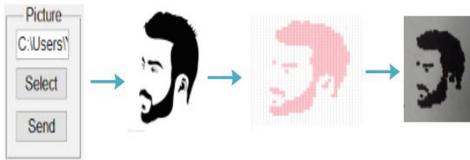


Fig. 15. Picture analog output and actual output.

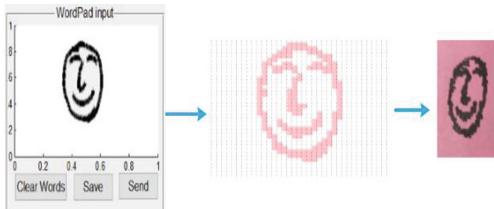


Fig. 16. Handwriting board analog output and actual output.

Figure 14 can be seen that the basic output of Chinese characters can be completed, but the effect is not better the result of analog output, from the actual output results, “day” word a bit oblique, which is due to the belt caused by the scalability.

Figure 15 shows the picture input mode of the analog output and the actual output. As can be seen from Fig. 15, the output similarity of the head is still relatively high, because the design of the picture compression is too large, resulting in a lot of details are eliminated. As Fig. 16 shows the tablet output and the actual output.

Figure 16 is a picture drawn with a tablet, the output has been deformed, the problem is not enough picture compression accuracy caused.

Through the above experimental results can be seen that the writing robot can basically complete the design of the task, Chinese characters strokes to write a Chinese character takes a little time-by-step sweep mode, because of the reasons for the algorithm, the effect is not progressive scan.

5 Conclusions

This topic design CoreXY structure writing robot, build a hardware platform, to achieve the X axis and Y axis linkage purposes.

Based on MatlabGui design of the host function to achieve a single text input, picture input, tablet input, and can achieve the recognition of Chinese characters and stroke coordinates of the extraction.

The improved genetic algorithm is introduced to optimize the PI parameters of the motor control. The simulation results show that the optimization of the genetic algorithm is improved. In the actual operation of the robot, the results obtained by adding the improved genetic algorithm are better than those obtained without the algorithm and the traditional genetic algorithm.

Experiments show that the CoreXY structure writing robot can realize the writing of Chinese characters and the description of pictures, so the design scheme is feasible.

Acknowledgements. This paper is supported in part by the Key Disciplines for Operational Research and Cybernetics of the Education Department of Guangxi Province, and in part by the basic ability Promotion Project of Young and Middle-aged Teachers in Guangxi Colleges and Universities in 2018 under grant 2018KY0703, and in part by Guangxi Science and Technology Teachers College Research Project in 2018 under grant GXKS2018YB012, and under grant 2018GKSYJGZ02.

References

1. Huang R, Liu D, Xu J, Chen N, Fan L, Zeng H (2017) Educational robot development status and trend. *Modern Educ Technol* 27(1):13–14
2. Jin Y, Cheng F, Zheng C (2009) Smart lettering robot mechanism and hierarchical control design. *J China Jiliang Univ* 20(3):249–251
3. Zeng H, Huang Y, Chao F, Zhou C (2016) Survey of robotic calligraphy research. *CAAI Trans Intell Syst* 11(1):15–19
4. Zhang X, Kong Q, Ma X (2010) Study on control system of four-degree-of-freedom writing robot. *Mech Eng Autom* (5):141–143
5. Meng Q (2015) Research on motion control system based on XMC4500. Harbin Institute of Technology, Heilongjiang
6. Li D (2015) The implementation technology research of permanent magnet synchronous motor servo system based on XMC4500. Harbin Institute of Technology, Heilongjiang
7. Wang J (2016) Research on the Number Writing of Mechanical Arm Based on Arduino and Digit Recognition. Beijing University of Technology
8. Hu X, Shen X, Xiang W, Tang L (2015) Study of speed control system on ship diesel engine for iterative-account based on genetic algorithm. *Ship Eng* 37:150
9. Caponio A, Cascella L (2007) A fast adaptive memetic algorithm for online and offline control design of PMSM drives. *IEEE Trans Syst Man Cybern* 37(1):28–41
10. Shen Y, Hu J, Meng L, Ma S (2016) Application of genetic algorithm in optimization of PI control parameters of motor. *Technol Innov Appl* 28:48–50



The Research and Design of Intelligent Monitoring and Management Platform for River and Lake Water Environment

Jiabin Yu, Zhe Shen^(✉), Xiaoyi Wang, Jiping Xu, Li Wang,
Qian Sun, Zhiyao Zhao, and Huiyan Zhang

School of Computer and Information Engineering,
Beijing Technology and Business University, Beijing 100048, China
1124219701@qq.com

Abstract. The water environment quality of rivers and lakes in China is deteriorating and the eutrophication of water body is aggravating. Aiming at the problem of water environmental pollution in these rivers and lakes, an intelligent monitoring and management platform for water environment in rivers and lakes is proposed. Firstly, the overall system architecture of the water environment Internet of Things (IOT) management platform is designed, and then the hardware system of water environment monitoring and the software system of water environment management are designed and implemented. Then, the hardware equipment of water environment monitoring is specified, and the multi-dimensional monitoring scheme is put forward. Then, the software of water environment management is described from Web and APP client respectively. The platform can provide strong technical support for monitoring and prevention of water pollution in rivers and lakes.

Keywords: Water environment · Monitoring · Algal blooms · APP client

1 Introduction

Rivers and lakes are the foundation of human industry, agriculture, tourism and sustainable socio-economic development. Therefore, the water quality of rivers and lakes is very important for humans. According to the survey, about 40% of the rivers and lakes in the world suffer from different levels of eutrophication. The biggest problem currently facing European lake waters is eutrophication. 80% of the 96 lakes in the statistics show eutrophication to varying degrees. In the past 20 years, the eutrophication of lakes in China has developed at a relatively fast rate, and it is one of the countries with the most serious and widespread distribution of algal blooms in the world [1].

The survey shows that the proportion of eutrophic lakes in China has increased from 41% in the late 1970s to 61% in the late 1980s, and rose to 77% in the late 1990s. Among the 26 key lakes controlled by the state, the water quality is generally poor, which is lower than the Class V standard of the Surface Water Environmental Quality Standard (GB3838-2002). The nitrogen and phosphorus pollution is high, and a

considerable number of lakes have algal blooms (such as Chaohu, Taihu, Hongze Lake, etc.) [2, 3].

Based on the above background, the related ideas and construction schemes for the design and research of the river lake water environment intelligent monitoring and management platform are proposed. Firstly, the overall planning and architecture design of the water environment intelligent detection and management platform is carried out, and the water environment intelligent monitoring system is analyzed and designed. Management system; research and development of monitoring equipment for different types of water environment; on the basis of research and development testing equipment, build water environment management system from the Web side and APP client respectively, develop relevant software to achieve integrated management. The realization of the design of the river lake water environment intelligent monitoring and management platform can greatly improve the river and lake water environment pollution prevention and control capabilities, and improve the water environment management and governance level.

2 Water Environment Intelligent Monitoring and Management Platform Planning

2.1 Overall System Architecture Design

In order to clarify the organizational structure relationship between the intelligent monitoring equipment of water environment and management software, and to meet the general requirements of the Internet of Things system [4]. The Internet of Things application platform uses Visual Studio 2017 as a development tool, combined with the geographic information development platform Map for Windows, through the comprehensive use of GIS technology, database technology, embedded technology, high-level language programming technology, COM component technology, etc., to achieve 4G-based water quality Information remote wireless transmission, surface image display, color map display, historical data query, report printing, regional water eutrophication evaluation, algae bloom prediction and early decision-making and governance decision-making functions. Considering that the user is relatively fixed, the data consistency and integrity requirements are high, and the data analysis function is required, the C/S architecture mode is adopted in the software design, which effectively reduces the network traffic and the server operation amount; Through the dynamic webpage programming technology, the water environment monitoring information and analysis results are released to the Internet and various mobile terminals in real time, and finally realize the water environment monitoring, water quality comprehensive evaluation, algae bloom prediction and governance decision and information integration of the rivers and lakes management. The framework of the river lake water environment intelligent monitoring and management platform is shown in Fig. 1.

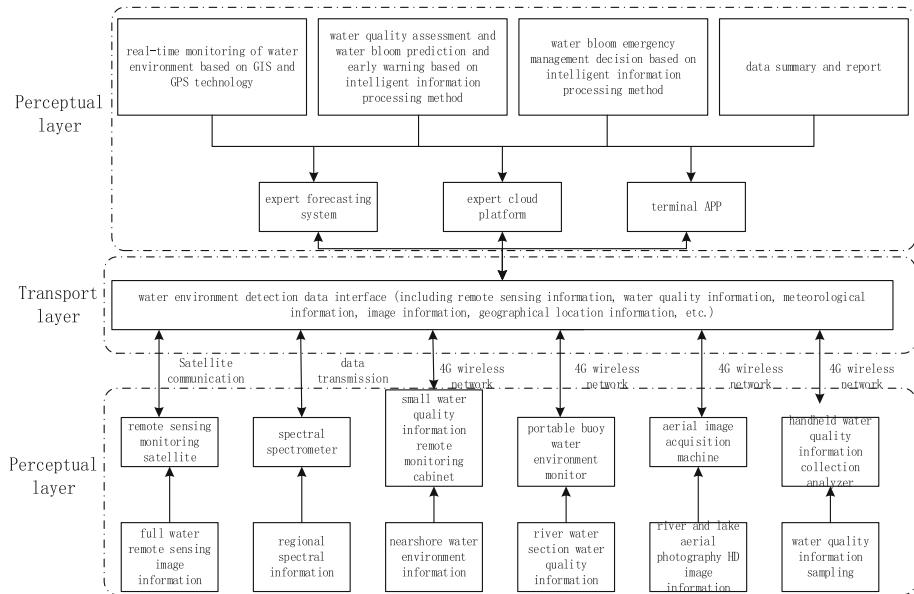


Fig. 1. Platform architecture design

The overall architecture of the river and lake water environment intelligent monitoring and management of the Internet of Things consists of the sensing layer, the transport layer and the application layer. The sensing layer is the foundation of the Internet of Things system. It is composed of various water environment monitoring devices and can collect multi-dimensional data of different types and different regions. The transport layer connects the sensing layer and the application layer, and implements on-demand transmission of different types of data through the network infrastructure and the self-developed data interface. The application layer is user-oriented and consists of subsystems with different functions. It can be embedded with intelligent algorithms to achieve water environment monitoring, comprehensive water quality assessment, algae bloom prediction and governance decision, integrated information management and data generation. Features such as summarization and reporting.

2.2 Design of Water Environment Information Monitoring System

The water environment monitoring equipment has undergone an evolving process: the initial water environment monitoring is carried out through the analysis of water sample chemical experiments, and the monitoring process is complicated by the environment [5]; The traditional water environment monitoring method mainly monitors the chlorophyll concentration, temperature, dissolved oxygen, pH value, nitrogen and phosphorus parameters in the water by using a water quality parameter sensor in a specific water area [6]. However, the traditional water environment detection method is highly susceptible to the geographical type, and there are great defects in the comprehensive analysis of water quality in the whole region. In order to overcome the

shortcomings of traditional water environment detection methods and realize comprehensive analysis from point radiation to water quality in the whole region, the river and lake water environment intelligent monitoring and management platform has built a multi-dimensional water environment information monitoring system, and with satellite remote sensing monitoring and aerial photography. The collection and detection methods and the means of spectrometer measurement are combined to obtain the water quality status of the whole region, and then the corresponding water quality parameter information is obtained in real time from the monitoring stations in different regions, so as to realize the comprehensive analysis from the point radiation to the water quality in the whole region. The architecture of the water environment information monitoring system is shown in Fig. 2.

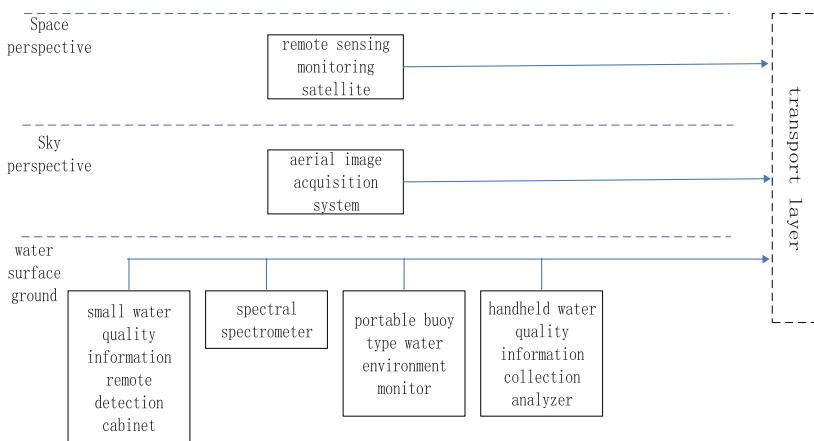


Fig. 2. Water environment information monitoring system

The multi-dimensional water quality monitoring hardware system mainly includes remote sensing monitoring satellite, aerial sleep image acquisition instrument, small water quality information remote monitoring cabinet, ground spectrum spectrometer, portable buoy water environment monitor, handheld water quality information acquisition analyzer, combined with wireless communication technology. Real-time acquisition and remote transmission of water environment information by remote sensing technology, global positioning technology, geographic information technology, and embedded technology.

2.3 Water Environment Management Software System Design

The water environment management software system can realize functions such as water quality information monitoring, information management, water quality analysis, and information push, and realize intelligent management functions through user interaction. When designing the water environment management software system, using a hierarchical management method, combined with a modular design concept,

packaging modules with different functions can not only improve the operating speed of the software system, but also reduce the upgrade and maintenance work the amount [7]. The system structure is shown in Fig. 3.

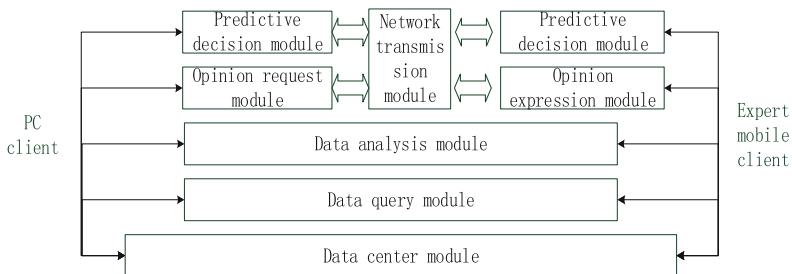


Fig. 3. Management software system

In the water environment management software system, the functions of each main module are as follows:

- (1) Data center module. The processing core of the data can realize the storage, addition, deletion, modification, etc. of the first-line information such as water quality information and expert opinions, and at the same time realize the management of transit information such as data analysis, prediction, and decision results;
 - (2) Data analysis module and data query module. Built on the data center module, it is a direct data interface between the WEB client and the expert mobile client. It can realize real-time or historical data query and horizontal and vertical comparison of data for managers and experts respectively.
 - (3) Opinion request/expression module. The two realize the information interaction through the network transmission module, and the real-time mutual transmission of information between the manager and the expert can be realized, the intermediate part of the decision can be shortened, and the work efficiency can be improved;
 - (4) Prediction decision module. The core function module of the software system, embedded in the algae bloom prediction, intelligent decision-making algorithm for water bloom management, can comprehensively process water quality information data and expert opinions, and give the analysis results.

3 Lake Reservoir Water Environment Management System

Under the framework of multi-dimensional information monitoring system of lake water environment, combined with GIS technology, database technology, embedded technology, high-level language programming technology, combined with geographic information development platform Map for Windows, embedded algae bloom

prediction and emergency decision-making. The model realizes real-time collection and remote transmission of water environment information, water quality patch map display, water surface image display, intelligent information processing, expert decision-making and analysis, report printing, and information release.

3.1 Administrator-Oriented Web Client

In view of the large amount of water quality information data and various algorithms embedded in the Web client, the system platform needs to meet the requirements of large data volume and fast calculation speed. In the development process of the Web client, Visual Studio 2017 is used as the programming environment, and the C# programming language is adopted.

(1) System startup and main interface

When the system starts, the login interface will appear after the system is initialized. Because the system has certain professionalism, the user registration function is not opened, and the user name is obtained through allocation. The main interface of the system includes the water quality status monitoring area, function button bar, user information area, auxiliary function area and other parts.

(2) Evaluation of eutrophication of water bodies

The assessment of water eutrophication refers to the determination of the nutritional status of water bodies through a series of indicators related to the nutritional status of water bodies and the correlation between them. The real-time water quality status of different sites can be viewed, and the evaluation index of the degree of eutrophication of the water body is given by the embedded algorithm. At the same time, historical data can be queried and compared to visually check the changes in the water quality status of the site.

(3) Data analysis

Click the “Data Analysis” icon on the main interface to select the site, parameter, and data time range, and you can view the different changes of the corresponding parameters of the selected site in the selected time range. The module can also make multiple iterations and draw the curve to show the results more intuitively.

(4) Algae bloom prediction

Algae bloom prediction [8] is the core function of the system. Under the condition of analyzing historical data, the real-time monitoring of water quality information obtained by embedded intelligent algorithm analysis predicts the growth trend of algal blooms. Take appropriate measures to manage.

(5) Intelligent decision

After collecting the sites and expert opinions of the algal bloom management decision, click the “Start Decision” button. The system automatically selects the current state of algae bloom management according to the embedded algorithm. The display result consists of the ideal program attribute weight, the program sorting histogram, etc. The user can also click the “Query” button to view the decision history of the current site.

3.2 Mobile APP Client for Experts in the Water Environment

The expert mobile phone APP [9] client of the river and lake water environment intelligent monitoring and management platform can realize the real-time request and transmission of expert opinions in the water field, reducing the time for collecting expert opinions, improving the collection efficiency of expert opinions, and the science of decision-making results. Increased protection of sex and effectiveness. The expert mobile client is based on Android 8.0 development, and the development platform uses Eclipse. The expert mobile APP client includes one aspects: core function interface. The following is a detailed introduction to the two major functional areas. The APP interface is shown in Fig. 9.



Fig. 4. Expert mobile phone client core function interface

The core function interface is mainly composed of map display, site data, forecasting and warning, expert decision and user information. The interface design is shown in Fig. 4.

- (1) Map display: Click on the map display under the APP to locate the rivers and lakes to be carried out. After the positioning, the water quality evaluation results of the relevant rivers and lakes will appear, weather conditions of the day and related factors affecting water quality. Through this interface experts can have a general understanding of the positioning of the rivers and lakes.
- (2) Site data: It mainly includes water quality information and functions related to viewing historical data of different water quality impact factors. The water quality information is mainly the values of chlorophyll, blue-green algae, dissolved oxygen, sulfate, nitrate, rainfall, ammonia nitrogen, PH value, water temperature,

water depth, temperature, air temperature, etc., which is more intuitive and accurate in reflecting the amount of water quality information; Below the interface is a view of the historical data of the water quality information, first select the corresponding water quality information (such as water temperature), and then select the historical data days (such as 7 days), the historical data changes of the relevant water quality information (water temperature) will be displayed below. line chart. Through the line chart expert, you can have an accurate understanding of the water quality information of the river and lake in the near future, so that you can make a reasonable decision-making plan.

- (3) Predictive and early warning: Predictive and early warning [10] is mainly to predict the relevant water quality indicators in the future, so as to better help experts make decisions. This interface mainly includes functions such as site selection, indicator selection, forecasting days, and comprehensive evaluation of water quality. The site selection is mainly to select the rivers and lakes that you want to predict; the indicator selection is mainly to select the water quality indicators that the selected rivers and lakes want to predict; the forecast days are to predict the indicators in the next few days. After selecting the corresponding data, a predicted line graph corresponding to the water quality index content will appear below, which can more intuitively display the forecast data in front of the experts. At the bottom of the interface, a comprehensive assessment of water quality in the next few days will appear.
- (4) Expert decision-making: The expert decision-making interface is mainly for experts to propose corresponding governance methods for rivers and lakes where problems occur. This interface mainly includes site selection, methods for managing different water quality information, and historical decision records for different water quality information. The choice of site pair is the same as that of the forecasting and warning interface. It will not be repeated here. The main method is to propose the treatment method for different water quality information. At the bottom of the interface is the historical decision-making method for different water quality information. The chlorophyll factor of rivers and lakes is treated by microbial method in 2018-08-03, which can provide a reference for experts to control related factors. After the expert makes a decision, clicking the submit button will transfer the decision information to the web in real time.

The main operation process and idea of the entire expert mobile APP client: The “map display” function first needs to select the site, then the location information of the site and the latest water quality data; the “site data” function enables simple water quality data analysis, experts You can select the corresponding site and parameters, and then you can check the changes of water quality parameters in the near-term time of the site. The “predictive warning” provides the experts with the query status of the water control decisions of different monitoring sites, and can view the algal blooms. Time and the final choice of governance methods provide reference for current decision-making; “expert decision-making” is the core function module expressed by expert opinions. Experts first select the site, and then select the degree of action of different parameter attributes for different alternatives as evaluation opinions. Then send expert comments to the web client.

4 Summary

It is proposed to build an intelligent monitoring and management platform for rivers and lakes. Firstly, the design of the overall system architecture of the water environment IOT management platform is completed, and then the system design of the water environment monitoring hardware system and the design and implementation of the water environment management software system are completed. The water environment management software system was designed from the Web side and the APP side respectively, emphasizing the powerful functions displayed by the mobile phone expert APP client in the river and lake water environment intelligent monitoring and management platform. The lake water environment monitoring and intelligent management of the Internet of Things platform provide theoretical and technical support for water environment prevention and control, which can greatly improve the river and lake water environmental pollution prevention and control capabilities, and improve water environment management and governance.

Acknowledgments. This work was supported by the National Natural Science Foundation of China under Grant (61703008) and the Project of High-level Teachers in Beijing Municipal Universities in the Period of 13th Five-year Plan (Grant: CIT&TCD201804014).

References

1. Chen C (2012) Modeling and system implementation of emergency management decision for lake and reservoir blooms. Beijing Technology and Business University, BeiJing
2. Bai Y (2014) Study on emergency management decision of algae blooms in rivers and lakes. Beijing Technology and Business University, BeiJing
3. Lu X, Xu F, Zhan W et al (2003) Research status and development trend of lake eutrophication model. *Adv Water Sci* 14(6):792–798
4. Shen S, Fan Q, Zong P et al (2009) Research on the architecture and related technologies of internet of things. *J Nanjing Univ Posts Telecommun: Nat Sci Edn* 29(6):1–11
5. Fei X (2005) Modern water quality monitoring and analysis technology. Chemical Industry Press, BeiJing
6. Xu J, Bai Y et al (2015) Internet of things platform for water environment monitoring and intelligent management of lakes and reservoirs. Papers Collection of the Annual Academic Meeting of the Chinese Society of Environmental Sciences
7. Shao F, Shi Y, Wang X et al (2014) Recognition of algae blooms in urban lakes and reservoirs based on complex network. *J Environ Sci* 8:2121–2125
8. Wang L, Liu Z, Wu C et al (2013) Prediction and factor analysis of water blooms based on multivariate time series analysis. *J Chem Eng* 12:4649–4655
9. Zhao B, Pan X (2018) APP architect practice guide. People's Posts and Telecommunications Publishing House, Beijing, pp 185–240
10. Kong F, Ma R, Gao J et al (2009) Theory and practice of prevention, prediction and early warning of cyanobacterial blooms in Taihu lake. *Lake Sci* 21(3):314–328



A Parametric Mesh Generation Method for Aero-Engine Compressor Blades

Min Jiang¹, Fengshun Lu², Long Qi²⁽⁾, and Yufei Pang²

¹ Department of Electrical and Mechanical Engineering, Mianyang Polytechnic, Mianyang 621000, China

² Computational Aerodynamics Institute, China Aerodynamics Research and Development Center, Mianyang 621000, China
qilong1226@163.com

Abstract. In order to solve the mesh generation problem for complex configurations in aerodynamic design and optimization of aero-engine compressor blades, a parametric mesh generation method is proposed. A detailed comparison is performed in the aspects of generation efficiency, mesh quality and simulation results between two meshes generated with Spider and NUMECA, given the same experimental conditions. The experimental results show that the proposed method can efficiently and automatically generate meshes with good quality.

Keywords: Compressor blades · Computational mesh · Topology structure · Parametric mesh generation

1 Introduction

Aero-engine performance determines the development of modern aviation industry. The aerodynamic stability design of compressor has been one of the research hotspots in the development of modern aero-engine [1]. Among the factors affecting the compressor performance, the blade geometry and aerodynamic performance play a key role in the overall performance of compressors [2, 3]. In order to accurately understand its performance and the relevant flow phenomenon, enormous experiments or computational fluid dynamics (CFD) simulations have to be carried out [4–6]. Considering the characteristics of low cost and short cycle, CFD simulations have been widely used in the compressor blade optimization design.

Mesh generation is a pre-processing step for the CFD analysis, which is usually a time-consuming and interaction-intensive process. Consequently, mesh generation has become an obstacle in the CFD simulations of compressor blades [7]. In order to address this issue, we propose a parametric mesh generation method for compressor blades, which can greatly improve the mesh generation efficiency and reduce the amount of involved manual operation while ensuring the mesh quality.

2 Method

There are many typical mesh topologies (e.g., O and C) for generating meshes. However, the traditional mesh topologies cannot satisfy the requirements for high mesh quality targeting complex configurations, especially for the areas around the compressor blades and trailing edges [8, 9]. In order to address this issue, a parametric mesh generation method is proposed in this paper. As shown in Fig. 1, the combined O-H topology is utilized and the mesh generation software Spider, which is developed by China Aerodynamics Research and Development Center, has been performed as our basic platform.

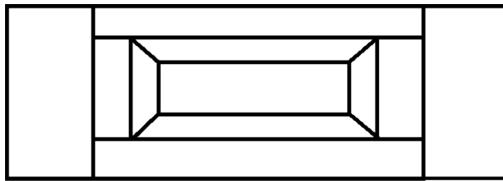


Fig. 1. O-H combined topology

The strategy for parametric mesh generation for compressor blades is as follows. First, the two-dimensional mesh topology is constructed according to the geometric characteristics. Second, the three-dimensional mesh for each blade is generated by stacking the two-dimensional mesh along the direction of blade height constricted by certain specific parameters [10]. Finally, the whole mesh is generated and optimized by adjusting a few control parameters.

As depicted in Fig. 2, we utilize the concept of local topology to decouple the complex relationships among connector, domain and block. Each local topology is aggregated with many points and connectors, and many of them can construct the global topology. Based on the steps shown in Fig. 2, the parametric mesh generation for blades can be performed automatically.

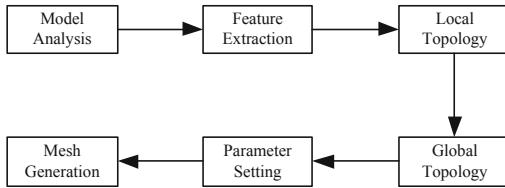


Fig. 2. The flowchart of parametric mesh generation for blades

3 Implementation

3.1 Determining Computational Domain

Multi-block structured meshes are widely utilized in the aircraft industry, since they can efficiently achieve accurate solutions, especially for viscous flows of complex

configurations. One of the difficulties for generating multi-block structured mesh lies in the determination of computational domain boundary and the partition of computational domain [11]. The layout of compressor blades is periodic and each blade row contains dozens of blades. In practical engineering applications, one or two blades are often used for numerical simulations. In this research, a single channel of one complete blade is selected as the computational domain and the middle of the flow passage is treated as the circumferential boundary. Theoretically, determining the computational domain in this way not only reduces the amount of the overall grid points, but also refines the region with relatively complex flow field. Consequently, the numerical simulation accuracy of the whole mesh can be improved.

In order to accurately simulate the working conditions of the compressor and obtain comprehensive flow field information, it is not enough only selecting the region between the blades as the computational domain of CFD simulations. Therefore, we extend the boundary of the inlet and outlet to a certain distance so as to control the flow difference between the inlet and outlet and the static pressure distribution on the outlet surface [12], as shown in Fig. 3.

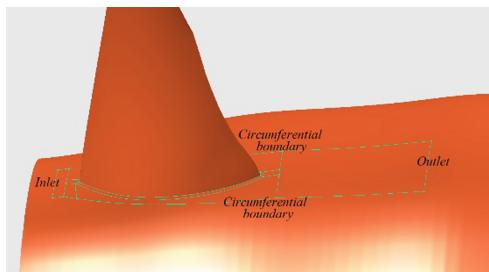


Fig. 3. The boundary of computational domain

3.2 Constructing Mesh Topology

3.2.1 Building Topology Principles

In the practical engineering application, the mesh generation process usually includes four steps. First, the geometric model is analyzed and processed if necessary. Second, the mesh topology is determined. Third, the mesh generation is carried out. Finally, the mesh quality is optimized using the smoothing algorithm [13–17]. The quality of topology determines the quality of generated meshes [18]. Therefore, the topology scheme should confront to the following principles: (1) being correct so that a usable mesh without any negative volume elements can be generated; (2) being simple in order to simplify the mesh generation process for complex configurations; (3) being changeable so that it can adapt to the transform of geometric shapes without modifying the topology.

3.2.2 Building Local Topology

As mentioned in the previous section, we utilize local topology to simplify the parametric mesh generation. The construction of two-dimensional mesh topology includes boundary layer topology, leading and rear edge topology, etc.

As shown in Fig. 4, constructing the topology for the near-blade surface consists of four steps: (1) four characteristic points and profile lines at the leading and rear edges of the blade [19] are obtained by mathematical modeling and relevant calculation; (2) the inner topological line is constructed by feature extraction technique; (3) through the propulsion technology, the profile lines are extended in the normal direction for a certain distance (e.g., the height of the viscous layer), and the lifting lines are obtained; (4) the boundary layer topology for rotor blade at the bottom part is constructed by projecting the lifting line onto the hub.

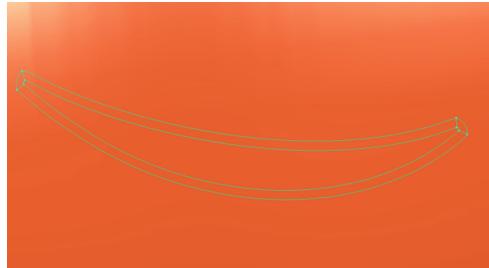


Fig. 4. Boundary layer topology

In addition to the boundary layer, the leading and rear edges also show characteristic flows. We hereby take the former edge as an example to illustrate the topology construction for the leading and rear edges. As shown in Fig. 5, the computational domain determines four boundary points, which are then utilized to construct auxiliary lines. Then the angle between lifting line and transition line is checked to meet the specified conditions. Finally the movable points on auxiliary lines are determined. The topology for the leading edge is generated by connecting the movable points and the characteristic points on the boundary layer. Note that the topology for the trailing edge can be obtained similarly.

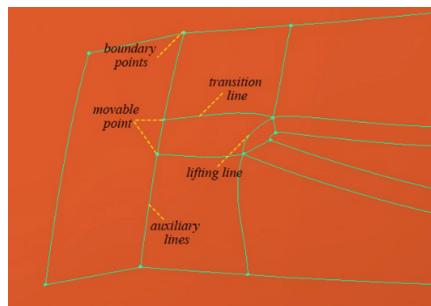


Fig. 5. Leading edge topology

3.2.3 Constructing Global Topology

After the construction of local topologies, the corresponding points of these topologies are connected to form two-dimensional topology slices. The global topology can be constructed through the continuous torsion and stacking of two-dimensional slices in the direction of leaf height, as shown in Fig. 6. In order to connect arbitrary two slices, it is necessary to distinguish the up- and down- side of each slice. Since each slice has the same data structure, a coupler line can be obtained by connecting the corresponding points in the two slices. After the connectors are created, the assembly of domain and block can be performed.

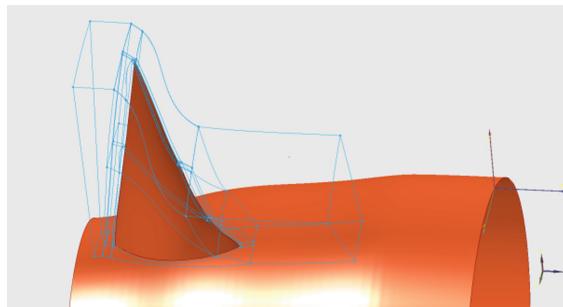


Fig. 6. The overall topology

3.3 Setting Mesh Parameters

The parametric mesh generation has great advantages in the aspect of adaptability. When the geometry changes, engineers only need to import the geometry file again and modify the relevant parameters before a brand-new mesh is automatically generated. The configurable parameters include topology type, mesh density, number of blades, position of blade row, height of the first layer in the boundary mesh, location of inlet and outlet, etc. The mesh density level is set globally (generally set to 3 or 5). The actual discrete points of each connector can be calculated with the following equation. Note that the mesh density can be adaptive according to the length of the connector density parameter:

$$Num = \max\left(\frac{preNum * ratio * len}{prelen}, 5\right) \quad (1)$$

Num is the amount of discrete points in the current connector, which is not less than 5. *preNum* is the number of points set in standard length, *ratio* is the density control level, *len* is the current length and *prelen* is the standard length.

4 Experimental Results and Analysis

In this paper, we utilize the single rotor of NASA Rotor67 compressor [20] to verify our proposed method. The meshes generated by NUMECA and our method are compared, and then simulated with the same computational parameters (turbulence model, spatial discrete format, etc.).

4.1 Mesh Quality Analysis

High-quality mesh is an important prerequisite to ensure accurate CFD simulations [21]. Figure 7 presents the two blade meshes, where we focus on the parts of hub, casing, blade and blade tip.

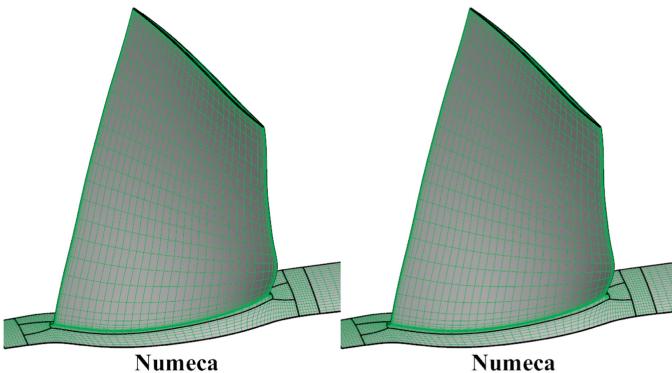


Fig. 7. The blade meshes

The mesh quality can be analyzed from the aspects of mesh density distribution, boundary layer density and mesh orthogonality [22]. Generally speaking, the two meshes are comparable in the above aspects. However, in order to ensure the quality of the grid, our proposed method adds auxiliary lines when generating the mesh, as can be seen in Fig. 7.

4.2 Numerical Results and Analysis

Given the same boundary conditions (total temperature for inlet is 288.15 K and total pressure is 101325 pa; back pressure outlet for is 101325 pa) and the same calculation parameters (turbulence model, spatial discrete format, etc.), the convergence speeds for the iterative solution of the two meshes are comparable, just as shown in Fig. 8.

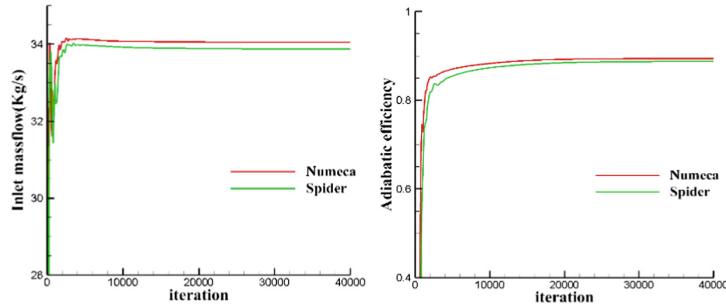


Fig. 8. Iterative convergence process

As listed in Table 1, the simulation results relevant to the two meshes are similar.

Table 1. Simulation results comparison

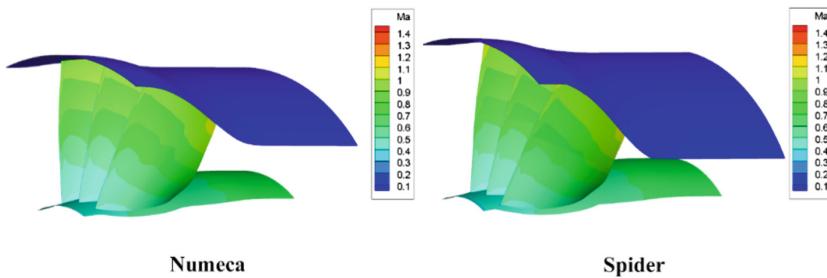
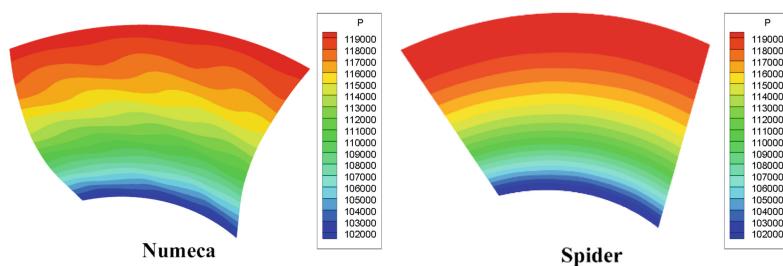
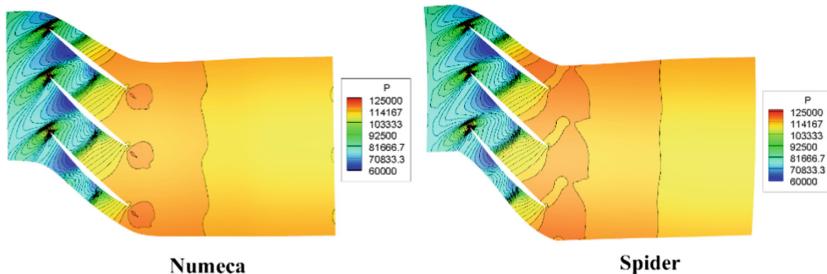
| | NUMECA | Spider | Relatively difference |
|--------------------------|--------------|--------------|-----------------------|
| Inlet mass flow(kg/s) | 0.34045E+02 | 0.33870E+02 | -0.51% |
| Outlet mass flow(kg/s) | 0.34220E+02 | 0.33930E+02 | -0.85% |
| Total pressure ratio | 0.15614E+01 | 0.15914E+01 | 1.92% |
| Static pressure ratio | 0.13774E+01 | 0.13859E+01 | 0.62% |
| Total temperature ratio | 0.111519E+01 | 0.11600E+01 | 0.7% |
| Static temperature ratio | 0.111116E+01 | 0.111153E+01 | 0.33% |
| Adiabatic efficiency | 0.89356E+00 | 0.88714E+00 | -0.72% |

Specially, the difference between inlet and outlet mass flow for our proposed method and NUMECA is 0.18% and 0.51% respectively, which indicates that the mesh generated with our method is superior to that of NUMECA.

Figure 9 shows the distribution of absolute Mach number on the outlet surface. It shows that the absolute Mach number at the casing of Spider is basically 0, meaning that the boundary layer mesh generated is reasonable.

Figure 10 shows the static pressure distribution on the outlet surface. Compared to the NUMECA mesh, the flow field of Spider mesh has been fully developed at the outlet, which accords well with the radial distribution.

We obtain three observations from Fig. 11. First, the static pressure distributions at 50% height of blades for the two meshes are similar. Second, the positions for detached shock wave at the leading edge and the positive shock wave are almost the same. Finally, the two meshes also exhibit analog static pressure distribution.

**Fig. 9.** The distribution of absolute Mach number**Fig. 10.** Static pressure distribution on the outlet surface**Fig. 11.** Static pressure cloud of the 50% upper leaf

5 Conclusion

In this paper, a parametric method has been proposed to perform automatic mesh generation for compressor blades. We perform a detailed comparison in the aspects of generation efficiency, mesh quality and simulation results between two meshes generated with Spider and NUMECA, given the same experimental conditions. The experimental results show that the proposed method can efficiently and automatically generate meshes with good quality.

Acknowledgments. This work was supported by the National Key Research and Development Plan of China under Grant No. 2016YFB0200704 and 2017YFB0202101. We also express our gratitude to Sumei Xiao and Dan Guo from SWUST for their technical supports.

References

1. Wang G (2012) Application of CFD simulation technology in aero-engine. *Aeron Manuf Technol* 400(4):98–99
2. Zhang J (2009) Optimization for turbo machine blade based on parallel genetic algorithm. Nanjing University of Aeronautics and Astronautics, The Graduate School
3. Gbadebo SA, Hynes TP, Cumpsty ANA (2004) Influence of Surface Roughness on Three-Dimensional Separation in Axial Compressors
4. Wang G (2009) Research on Aerodynamic Optimization Design of Fan/Compressor Blade using Parallel Genetic Algorithm. Nanjing University of Aeronautics and Astronautics, The Graduate School
5. Morini M, Pinelli M, Spina PR et al (2011) Numerical analysis of the effects of nonuniform surface roughness on compressor stage performance. *J Eng Gas Turbines Power* 133 (7):072402
6. Wang D (2016) Study on the dynamics of vibrations for the aero-engine compressor blades considered the fluid-structure interactions. Dissertation for the Doctoral Degree in Engineering
7. Cui J, Jiang X, You B (2003) Two-dimensional finite element meshing. *J Nanchang Univ (Eng Technol)* 25(2):150–157
8. Li H, Zhao Y, Shi Y et al (2013) Key technology of NC machining for aeroengine fan/compressor blade. *Aeronaut Manuf Technology*(16):34–37
9. Wei Y (2017) Study on the performance of the compressor blade of an certain aero-engine in service. *Sci Technol Eng* 17(5):119–125
10. Yang C, Liu L, Hu L et al (2006) Design of two-stage centrifugal compressor and analysis of internal flow field. In: Chinese society of engineering thermophysics academic conference on thermal-pneumatic dynamics, pp 369–373
11. Xu W (2008) Research on pre-processor of CFD grid about impeller. Nanjing University of Aeronautics and Astronautics, The Graduate School
12. Guo D (2014) Research on Automatic Generation of Grid computing for Aero-engine Compressor. Southwest University of Science and Technology
13. Ghaly WS, Mengistu TT (2003) Optimal geometric representation of turbomachinery cascades using nurbs. *Inverse Probl Eng* 11(5):359–373
14. Zhang Y, Bajaj C, Xu G (2009) Surface smoothing and quality improvement of quadrilateral/hexahedral meshes with geometric flow. *Communications in Numerical Methods in Engineering*, vol 25, no 1, pp 1–18
15. Ava S, Alireza J (2007) Airfoil shape parameterization for optimum Navier-Stokes design with genetic algorithm. *Aerospace Sci Technol* 11:443–450
16. Su Y, Lee KH, Kumar AS (2005) Automatic hexahedral mesh generation using a new grid based method with geometry and mesh transformation. *Comput Methods Appl Mech Eng* 194:4071–4096
17. Eiseman PR (2003) Grid generation for fluid mechanics computations. *Ann Rev Fluid Mech* 17:487–522
18. Zhang L, Wang W, Wang R (2011) Grid generation of different forms of turbocharger volute for flow simulation. *Agric Equip Veh Eng* 05:50–52

19. Yang H, Wang Yu, Shan C (2015) Survey of measurement method of aero-engine compressor blade edge. *Aviat Percis Manuf Technol* 51(6):43–45
20. Zhang P, Zuo H Tan Z et al (2012) Numerical analysis of FSI for NASA-Rotor67 blade. In: The fifth conference on vibration utilizing engineering in China & the fourth workshop on ultrasonic motor techniques in China, pp 724–730
21. Xu P, Xu B, Luo Y et al (2017) Based on hypermesh of high quality computational fluid dynamics mesh generation for reactor top structures. *Chin J Appl Mech* 34(2):281–286
22. Li H, Wu J, Liu J et al (2012) Finite element mesh generation and decision criteria of mesh quality. *China Mech Eng* 23(3):368–377



Centrifuge Fault Diagnosis Method Based on Random Forests Algorithm

Yanxi Li¹, Jing Bai^{2(✉)}, and Hezheng Li¹

¹ College of Electric Information Engineering, Beihua University,
Jilin 132001, China

² Science Park of Beihua University, Beihua University, Jilin 132001, China
jlbyj@163.com

Abstract. In this paper, centrifuge fault diagnosis method based on random forests algorithm is proposed, which can solve the problems of poor diagnostic effect, long time and low accuracy of traditional centrifuge fault diagnosis methods. According to the state parameters of centrifuge, Python is used to build a centrifuge model based on random forests to realize fault diagnosis. The results show that compared with other intelligent algorithms, centrifuge fault diagnosis method based on random forests algorithm has a good diagnostic accuracy and rapidity. Therefore, this method has broad development prospects in the field of centrifuge fault diagnosis.

Keywords: Fault diagnosis · Random forests algorithm · Python

1 Introduction

The centrifuge, which uses a centrifugal force generated by its own rotor to perform a series of operations and corresponding analysis on the required substances [1], is a widely used equipment in many industries, such as medical and industrial production.

The reliable and safe operations of the centrifuge are an important basis for industrial production. The reliability of the equipment is a new discipline proposed in recent years. Because the centrifuge may break down in the process of production and operation, which may cause serious damage to equipment and even staff, we need to establish a different fault diagnosis method [2]. How to judge and predict the operating state of the centrifuge more accurately and rapidly is essential.

Data mining is to find out the valuable knowledge hidden in the data from a large number of data and to mine the data through a variety of analysis, such as artificial intelligence, machine learning and so on [3]. It can solve the dilemma of multiple data and lack of knowledge, widely used in various industries and has significant effects [4].

In this paper, by programming in Python environment for data preprocessing and random forests model construction to diagnose whether the centrifuge fault and compared with other intelligent algorithms, the simulation results show that random forests algorithm provides a better solution for centrifuge fault diagnosis.

In the early research on equipment fault diagnosis, by relying on the experience of the staff, it is possible to judge whether the equipment is faulty. But this method can judge only some simple faults, or when the fault occurs to a certain degree of severity,

people can find it by observing the appearance, so the early diagnosis technology can not detect the faults at the beginning of the fault, let alone prevent the faults in advance [5]. So far, most of the researches on the diagnosis of centrifuge rely on the experience of experts, which will reduce the efficiency of maintenance and increase the uncertainty of diagnosis. Therefore, it is necessary to use artificial intelligence technology to diagnose the fault of the centrifuge. Random forests algorithm can overcome these problems, and has the advantages of fast training speed, no feature selection, and strong model generalization ability.

2 Centrifuge Fault Diagnosis Technology

The safety of centrifuge equipment is the basis for safe, reliable and stable operation in industrial production, which is of great significance for industrial production. Therefore, it is necessary to establish a real-time fault diagnosis on the centrifuge. Traditional diagnostic methods, such as manual diagnosis and instrumental diagnosis, have poor diagnostic effect, long time-consuming and low accuracy, while intelligent diagnostic methods have better diagnostic accuracy and rapidity [6]. Therefore, this paper adopts the intelligent diagnosis method, which is mainly base on the data mining algorithm. The algorithm is established to solve the fault diagnosis of the centrifuge. The main algorithms are the expert system, Bayesian network, artificial neural network (ANN) and so on [7].

3 Random Forests

3.1 Definition and Basic Principles of Random Forests

Random forests, a integrated and novel machine learning algorithm, which have the functions of supervision. The “randomness” of random forests are reflected in two aspects: the randomness of data sampling and feature selection, selecting 60% to 80% of all features at a time. The construction of “forests” is to put many decision trees together in parallel, mostly trained in a “bagging” method. The bagging method is bootstrap aggregation, which randomly sampling with replacement for the trained data. The establishment of classifiers, the overall final result can be obtained through the algorithm used, which is the average of multiple simultaneous training classifiers [8]. The equation is as follows:

$$f(x) = \frac{1}{M} \sum_{m=1}^M f_m(x), \quad (1)$$

where M is the total number of features, m is the number of features randomly selected from the M features.

Random forests are the use of a variety of different methods for the initial training data to be placed back randomly sampling, each time take P samples, and m times sampling to obtain m training sets. Then through training to get m decision trees, ultimately, from the obtained m decision trees to build a random forests model. For the classification problem, by the statistics of the results of multiple classifiers, the final category is more as the result of the classification. CART tree of random forests is a classification trees, which can achieve the target of classification. Gini index, a criterion for classification trees, which means the probability that the prediction result is wrong. Gini index can be written as follows:

$$Gini(p) = \sum_{k=1}^k p_k(1 - p_k) = 1 - \sum_{k=1}^k p_k^2, \quad (2)$$

where p_k is the probability that the sample is classified into the k category, $(1 - p_k)$ is the probability of the sample being misplaced.

Random forests voting algorithm can be expressed as follows:

$$C_p = \arg \max_c \left(\frac{1}{n_{tree}} \sum_{i=1}^{n_{tree}} I\left(\frac{n_{n_i,c}}{n_{h_i}}\right) \right), \quad (3)$$

where n_{tree} is the number of trees, n_c is the total number of categories. After voting, a confusion table CM is generated, which is a $n_c \times n_c$ table.

3.2 Random Forests Algorithm

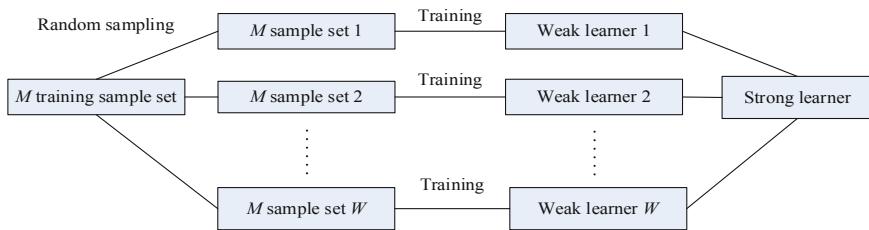
A decision trees start from the root node step by step to the leaf node, and all the data will fall to the leaf node, that is, the final result is achieved [9]. The root node is the first selected feature point, and the leaf node is the final decision result. The selection node is a segmentation feature, which is based on the value of the calculated entropy to find the root node and the non-leaf node.

The algorithm of random forests can be briefly described as:

Set the data set to $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, W is the number of iterations of the weak classifier, $f(x)$ is the output of the final strong classifier.

- (1) The training data set are randomly sampled m times to obtain m sample sets D_n . Let one of them is the nth time, then the nth decision trees model are constructed by the obtained sample set. By taking a part of the obtained samples, the best feature is found as the division of the decision tree according to the measure of entropy.
- (2) The final target is a classification task, the most or one of the categories is voted by the W classifiers, which is the final category.

The schematic diagram of random forests is shown in Fig. 1.

**Fig. 1.** Schematic diagram of random forests

4 Data Mining

The data collected in this paper are from a centrifuge with a room temperature of 20 °C, the voltage of 380 V, rated current of 120 A and operating speed range of 300 n/min–1300 n/min. This data mining is based on the change in operating current to determine whether the centrifuge is faulty. Therefore, data mining can be performed according to this working condition.

4.1 Reading and Preprocessing of Data

Data reading, is the first step in data mining, which is very important for data preprocessing and subsequent modeling. The pandas library in Anaconda is used to read and preprocess data, which encapsulates many functions and is convenient for data preprocessing.

The result obtained is shown in Fig. 2.

| | Time | Run | Spin | Pressure | Temperature | Horizontal | Vertical | Scurrent | Rcurrent | Front | Back | Feed | Status |
|---|------|-----|------|----------|-------------|------------|----------|----------|----------|-------|------|------|--------|
| 0 | 1 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | normal |
| 1 | 2 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | normal |
| 2 | 3 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | normal |
| 3 | 4 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | normal |
| 4 | 5 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | normal |

Fig. 2. Reading of data

The first line is the variable name, which is the abbreviation of the data column name. “Time” is the running time of the centrifuge; “Run” is the running speed of the motor; “Spin” is the rotational speed of the centrifuge; “Pressure” is the pressure of the lubricating oil; “Temperature” is the temperature of the lubricating oil; “Horizontal” is horizontal vibration; “Vertical” is vertical vibration; “Scurrent” is the starting current; “Rcurrent” is the running current. “Front” is the seal pressure of front part dry gas; “Back” is the seal pressure of back part dry gas; “Feed” is the dry gas seal pressure of feed; “Status” is a centrifuge running in normal or fault.

There may be some problems in the collected source data, such as the lack of some data, the appearance of NaN in Python, the type of text that some data types cannot execute, and the data being too large or too small, which will affect subsequent operations, so it is necessary to preprocess these collected data.

Because the centrifuge data collected in this paper do not have missing values, there is no need to deal with them. However, there is a character type data which is “str” type data, it cannot be recognized in machine algorithm. It can only recognize an “int” type or a “float” type. So the data needs to be processed.

The result of data type conversion obtained is shown in Fig. 3.

| | Time | Run | Spin | Pressure | Temperature | Horizontal | Vertical | Scurrent | Rcurrent | Front | Back | Feed | Status |
|---|------|-----|------|----------|-------------|------------|----------|----------|----------|-------|------|------|--------|
| 0 | 1 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | 1 |
| 1 | 2 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | 1 |
| 2 | 3 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | 1 |
| 3 | 4 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | 1 |
| 4 | 5 | 300 | 390 | 0.4 | 25.5 | 1.2 | 1.7 | 420 | 75 | 0.15 | 0.15 | 0 | 1 |

Fig. 3. The result of data type conversion

4.2 Visual Display of Data

Data visualization plays a very important role in data mining, and it can directly express the correlation between various features [10]. Before data mining, the data will be visualized to show the overall situation of the data or the relationship between the variables. The heat map of the dataset is drawn.

The result obtained is shown in Fig. 4.

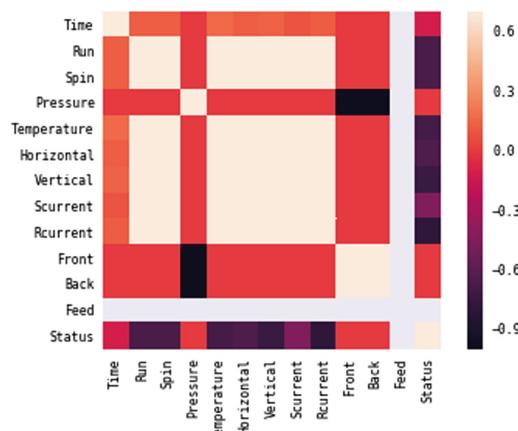


Fig. 4. Correlation between data variables

As can be seen from the heat map, the variable “Status” has a great relationship with the variables “Temperature” and “Recurrent”, showing a negative correlation. With the increase of temperature and current, centrifuges are more likely to fault.

5 Random Forests Applications

This paper mainly discusses the application of random forests in Python. Predict the state of the centrifuge through random forests algorithm and some important sample features such as the normal operating current of the centrifuge.

The random forests code is as follows:

```
from sklearn.ensemble import Random Forests Regressor as RFR
# Introduce the random forests classifier library from the sklearn library and use
# "RFR" to represent
np.random.seed(190)
random_forest = RFR()
random_forest.fit(X_train, y_train)
y_pred = random_forest.predict(X_test)
Visualize the prediction results. The code is as follows:
y_pred = lin_reg.predict(X_test)
plt.rcParams['figure.figsize'] = 10,8
plt.scatter(data[‘Time’], y_pred, c = r, linewidths = 4)
plt.scatter(data[‘Time’], y_test, linewidths = 2)
plt.title(‘Observed and Predicted’)
plt.legend(‘pt’)
```

The result obtained is shown in Fig. 5.

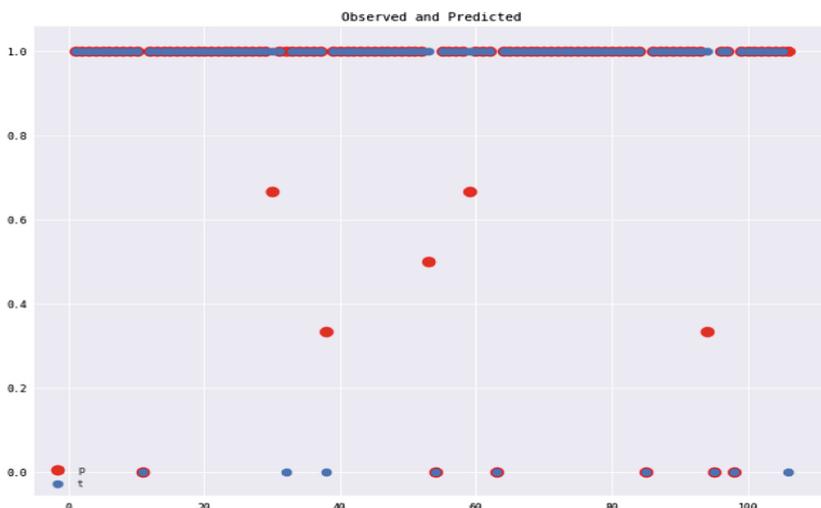


Fig. 5. Random forests prediction result

The output result show that the normal running current is about 0.85 times of the rated current and the fault current is between 0.95 and 1.05 times of the rated current. The blue point in the figure is the actual running state. There are only two cases, “0” and “1”. “0” means the centrifuge is in a fault state, “1” means the centrifuge is in normal operation. The red point is the predicted value. According to the above conditions, the fault state is 0.4 or less, the uncertain state is between 0.4 and 0.6, and the normal operation state is 0.6 or more. From the output results, random forests algorithm can correctly predict the probability of its fault to reach 88.89%, which indicates that the diagnosis result of the model is ideal, faster and more accurate than the traditional diagnosis method. When the predicted result is between 0.4–0.6 and below 0.4, that is, the uncertain state and the fault state, an alarm signal should be issued, and the corresponding pre-planning and measures should be started immediately to avoid more serious safety accidents and economic losses.

It can verify the advantages of random forests in centrifuge fault diagnosis by analogizing random forests with other intelligent algorithms.

6 Linear Regression

Linear regression is the most basic and important algorithm in machine learning. It can analyze one independent variable and the dependent variable, and can also be used to analyze multiple independent variables, which can be applied in a wide range [11]. Linear regression algorithm is simple and easy to understand, but can only fit linear data, otherwise under-fitting or over-fitting problems may occur [12].

The result obtained is shown in Fig. 6.

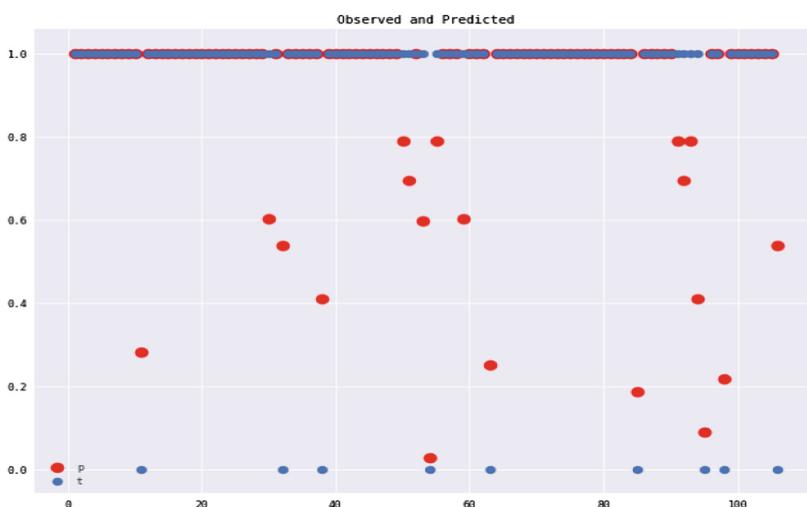


Fig. 6. Linear regression prediction result

As can be seen from the above figure, the probability that linear regression algorithm can correctly predict the fault of the centrifuge is 67%, and the effect is not satisfactory.

7 Logistic Regression

Logistic regression algorithm is a classic two-classification algorithm which is also commonly used in data mining. Logistic regression algorithm is simple, convenient to apply and can be applied to various forms of independent variables, but it has two shortcomings: overfitting and variable selection. Overfitting means that the fit of the logistic regression classifier to the training set can often show higher precision, but the accuracy of the data set outside the training set is poor. The variable selection problem refers to the model parameters obtained after the classifier is trained. Most of these model parameters are not 0 [13]. In machine learning algorithm selection, logistic regression algorithm is the first consideration, and then the complex algorithm is used.

The result obtained is shown in Fig. 7.

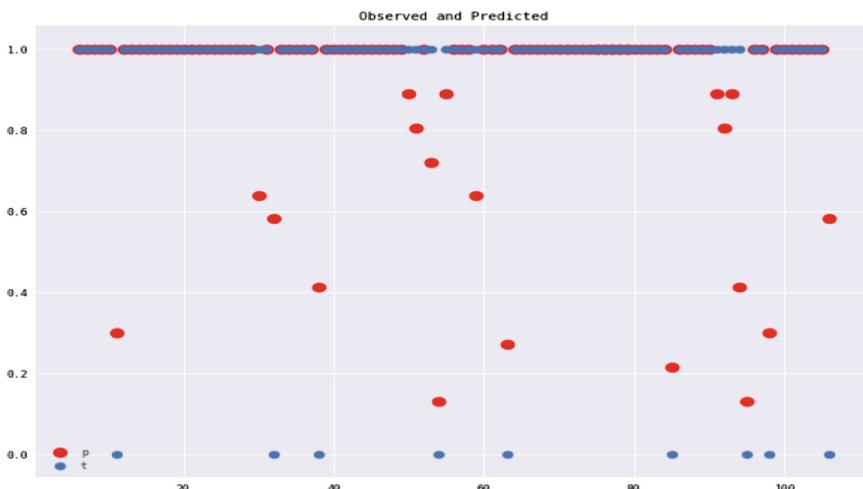


Fig. 7. Logistic regression prediction result

As can be seen from the above figure, the probability that logistic regression algorithm can correctly predict the fault of the centrifuge is also 67%.

For the data collected in this paper, the centrifuge fault diagnosis method based on random forests algorithm, which is better than the learning model constructed by linear regression and logistic regression, has a good diagnostic accuracy rate.

In practice, there is no intelligent algorithm that solves all problems well. The main reason is that the performance of intelligent algorithms depends largely on the amount of data and data structure. Therefore, we can evaluate whether the selected intelligent algorithm is the best algorithm by the accuracy of experimental prediction [14]. If you

use a large random forest, that is, many decision trees get better performance, but it will make it slower and require more memory [15]. However, linear regression needs to get some information about the relationship between feature variables and data structure. Otherwise, it may be difficult to design polynomial regression for non-linear data. At the same time, when data complexity is high, the performance of the model constructed by linear regression is inferior to that of other models [16]. For logistic regression, when the data features are missing or the feature space is large, the performance is not good, and it is easy to underfitting, and the classification accuracy is not high [17]. In short, different intelligent algorithms have their own adaptive occasions, so we need to choose the best intelligent algorithm according to the actual situation.

8 Conclusion

Aiming at the frequent problem of centrifuge fault, the intelligent diagnosis model of centrifuge fault was established. Random forests algorithm is used to train and learn the data of the centrifuge, determine the fault threshold, and perform fault diagnosis on the centrifuge according to the real-time running data, which provides certain technical support for the monitoring of the centrifuge fault. Random forests can be used to diagnose identifiable faults, and the results prove that the algorithm is effective.

Acknowledgments. This work was supported by the Science and Technology Development Project of Jilin Science and Technology Department 20180623009TC.

References

1. Zhou C, Ling Y, Shen W, Liu M (2014) Study on sludge dewatering simulation of horizontal spiral sedimentation centrifuge. Chin J Mech Eng 50(16):206–212
2. Lu H (2013) Study on reliability of ultracentrifuge. Hunan University
3. Huang L, Jiang B, Lu S, Liu Y, Li D (2018) A review of in-depth learning-based recommendation system. J Comput Sci 41(07):1619–1647
4. Wanxiang, Hu N, Han P, Zhang H, Li S (2016) Application of big data mining technology in turbine performance optimization. Chin J Electr Eng 36(02):459–467
5. Yuan Z, Fan L, Lu J (2018) Fault diagnosis technology and development trend of electrical equipment. Precis Manuf Autom (04):51–53 + 64
6. Liu S (2018) Research on vibration body analysis and fault diagnosis method of horizontal vibration centrifuge. Coal Science Research Institute
7. Han Y (2012) Research on Fault Diagnosis Expert System of Ultracentrifuge. Hunan University
8. Zhao T, Wang L, Zhang Y, Tian S (2016) Identification of user electricity related factors and prediction of electricity consumption using mutual information and random forests algorithm. Chin J Electr Eng 36(03):604–614
9. Luo J (2012) Application of data mining in dynamic detection of malicious web pages. Shanghai Jiao Tong University
10. Lu J (2012) Identification and prediction of biological activities and toxicity of some small organic molecules. Shanghai University

11. Lan H, Wang X (2013) Insect image segmentation method based on multiple linear regression. *Comput Appl Softw* 30(07):193–196+208
12. Huang Y, Liu M, Zheng W, Liu X, Chen X (2015) Fuzzy linear regression dynamic load parameter prediction based on confidence interval. *J Electr Technol* 30(24):196–205
13. Mao Y, Chen W, Guo B, Chen Y (2014) Logistic regression model based on density estimation. *J Autom* 40(01):62–72
14. Wang S (2019) Discussion on machine learning algorithm based on big data analysis. *Inf Comput (Theoret Edn)* 04:59–60
15. Chaudhary A, Kolhe S, Kamal R (2016) An improved random forest classifier for multi-class classification. *Inf Process Agric* 3(4):215–222
16. Waser M, Garn H (2013) Removing cardiac interference from the electroencephalogram using a modified Pan-Tompkins algorithm and linear regression
17. Abhilash CB, Rohitaksha K, Biradar S (2014) A comparative analysis of data sets using Machine Learning techniques. In: 2014 IEEE international advance computing conference (IACC)



Medical Image Super-Resolution Based on the Generative Adversarial Network

Xu Huang¹, Qianyao Zhang², Guoli Wang^{2,3(✉)}, Xuemei Guo^{2,3},
and Zhonghua Li¹

¹ School of Electronics and Information Technology, Sun Yat-sen University,
Guangzhou 510006, China

² School of Data and Computer Science, Sun Yat-sen University,
Guangzhou 510006, China
isswg1@mail.sysu.edu.cn

³ Key Laboratory of Machine Intelligence and Advanced Computing,
Ministry of Education, Guangzhou, China

Abstract. In order to assist doctors to read medical pathological images with low resolution, this paper proposes a medical image super-resolution (SR) reconstruction method based on generative adversarial network (GAN). Considering that the pathological image has large non-organized regions, we design a medical pathological image preprocessing system to extract tissue area image patches. And, we improve discriminator with small batch relative discrimination to enhance the quality of reconstructed images by learning more prior information. We use Huber loss instead of the original MSE which can keep the network training stable. We find the feature similarity (FSIM) is suitable as an image quality evaluation way for medical image reconstruction research. And, the experimental results show the advantages of our method in the restoration of color and intercellular texture details.

Keywords: Medical image · Super-resolution · Generative adversarial network · Huber loss · Feature similarity

1 Introduction

Due to the hardware limitations of the optical device and the diffraction effects of the light, we generally acquired medical image with low resolution. However, the low-resolution medical image reduces the accuracy and efficiency of the doctor's diagnosis to a certain extent. In order to solve this problem, it is necessary to improve the quality of medical images. Currently, we can improve the quality of medical images from two aspects. On the one hand, we can try to break the hardware limitations of the device. But, this method is costly and has a very limited ability to improve image quality. On the other hand, we can improve image quality by using image super-resolution (SR) reconstruction from the perspective of software. Relatively, this method is low cost and effective.

We all know that SR aims to recover a high-resolution (HR) image from one low-resolution (LR) image and focus on natural images mainly. However, Medical image reconstruction is not exactly the same as natural image reconstruction. Firstly, medical

images have fewer targets and simple structural information, and the size, shape and position of target are highly variable. Secondly, the difference between the target and non-target information is very small, and it is easy to misjudge the same target. Therefore, in order to assist the doctor in recognizing the image information, the reconstructed image needs to restore sufficiently clear texture details and feature information.

In recent years, more and more SR methods based on deep learning algorithms are proposed. Those methods have continuously improved the performance of SR, especially the Peak Signal-to-Noise Ratio (PSNR) [1–6]. However, as PSNR is defined based on pixel-level image differences, the ability to capture perceptual information, e.g. texture details, feature information, is very limited [7–9]. Later on, some scholars have proposed to optimize the perceptual loss of super-resolution models in feature space rather than pixel space to solve this problem [10, 11]. It has been found that the reconstructed image is more realistic by using the generative adversarial network (GAN) [1, 12]. So, this paper combines the GAN and characteristics of medical image to achieve the reconstruction of medical images.

In this work, we proposed a medical image reconstruction method based on SRGAN model. Firstly, in order to obtain a effective dataset, we design a medical pathological image preprocessing system to extract tissue area image patches. Different from the previous method, the discriminator of this paper uses small batch relative discrimination (BatchGAN). This can not only make the adversarial loss learn the prior information when guiding the network training, but also reduce the difference between the reconstructed pathological image blocks. And, we use Huber loss to calculate the pixel loss part of content loss, which effectively alleviates the problem of training instability of GAN and makes the loss function converge to smaller value. Then, we use feature similarity (FSIM) as an evaluation standard of medical image reconstruction quality. Because medical images will be used for visual observation by doctors, and FSIM is more in line with human subjective visual perception.

Overall, the contributions of this paper are mainly in four aspects:

- (1) We designed a medical pathological image preprocessing system to extract tissue area image patches.
- (2) We proposed discriminator with small batch relative discrimination which can improve image quality by learning more prior information of sample data and help doctors read the reconstructed image by reducing the difference between reconstructed image blocks.
- (3) We use Huber loss instead of the original MSE to keep the network training stable.
- (4) We use FSIM to better evaluate the recovery of feature information after medical image reconstruction.

2 Related Work

2.1 Image Super-Resolution

In this paper, we mainly focus on single image super-resolution (SISR) based on deep neural network. Since Dong et al. [13] proposed the SRCNN model to learn the

mapping from LR image to HR image in an end-to-end manner, deep convolution network have been widely used for image SR. Based on SRCNN, Kim [14] proposed the VDSR model, and extending the network to 20 layers effectively improved the quality of reconstructed images. Then, unlike SRCNN and VDSR, which need to first enlarge the LR image into HR image by interpolation method, and then use HR image to make convolution calculation through network model. Shi et al. [15] proposed ESPCN to do convolution calculation directly on LR image.

Later, in the SR field, there were network structures such as deep networks with residual learning, residual blocks, and residual dense networks. Especially, the EDSR model proposed by Lim et al. [16] removed the Batch Normalization (BN) layer in the residual block, which improved the generalization ability of the network. Zhang et al. [17] proposed the use of effective residual-dense blocks on SR problems, and further explored the effective channels in deep networks to achieve the best PSNR performance. And, Yu et al. [18] proposed the WDSR model based on EDSR. WDSR removed the redundant convolutional layer in EDSR, and modified the residual block. The improved model can obtain a wider feature map. In order to further improve the quality of the generated image, Ledig et al. [1] proposed the SRGAN model, which uses the GAN to reconstruct the HR image. The method proposes to use the perceptual loss to guide the network training so that the reconstructed image satisfies the signal to noise ratio and restores rich high frequency detail information.

Domestic SR research on medical images is relatively mature in terms of optical characteristics, and is mainly published in China's *«Journal of Optics»* and other publications. There are methods for super-resolution reconstruction of structured light illumination microscope images, random optical reconstruction microscopy, protein fluorescence and super-resolution microscopy and so on. Then, the wide application of deep learning in SR problem also provides a new solution for medical image super-resolution research. In 2017, the team of Yan Jinghui of Tianjin University proposed a single image reconstruction method combining wavelet features and clustering dictionary for the characteristics of magnetic resonance medical images. Yair Rivenson et al. [19] also proposed that the method of deep neural network can improve the spatial resolution of microscopic images in large field of view (FOV) and depth of field (DOF), and the trained model has excellent portability and robustness.

2.2 Generative Adversarial Network (GAN)

Firstly, the generative adversarial network based on a differential generator network is a generative model. Inspired by the two-player game in game theory, GAN consists of generative network and discriminative network. When GAN is successfully trained to converge, it will perform very well in many computer vision tasks. But in fact, GAN is difficult to train and the training process is unstable. Because GAN's loss function can guide the training of the network, but it does not reflect the schedule of training of the network, which is easy to cause model collapse [20].

2.3 Image Quality Evaluation Method

Image quality evaluation methods are often used to assess the quality of an image. And these methods can be divided into two categories: subjective and objective [21]. The subjective evaluation is mainly the mean opinion score (MOS), and the objective evaluation includes PSNR [22] and the structural similarity (SSIM) [23, 24].

3 Method

In this section, we will introduce our complete processing system. There are three parts in our system:

- (1) Tissue area image patch extraction. This operation extract patches from tissue area of pathological slice to accurately processing of the sliced tissue area.
- (2) Patch reconstruction. This part implement the patch reconstruction of high-resolution image from low-resolution image.
- (3) Result evaluation. This operation is used to evaluate the quality of the reconstruction.

3.1 Tissue Area Image Patch Extraction

Unlike natural images, medical pathology images have many non-organized areas such as the gastroscopic tissue section shown in Fig. 1 that make it difficult to accurately process tissue information.

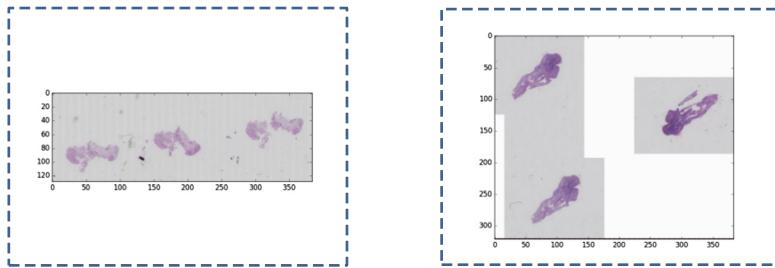


Fig. 1. Gastroscopic tissue sections stained with H&E

The gastroscopic tissue sections used in this paper were scanned by Hamamatsu scanner at 40X magnification with a resolution of 226 nm/pixel, a DPI size of 112389, and a format of NDPI. The average image size is about 300 MB. And image in NDPI format is stored similar to that in TIFF format, storing multi-resolution images in the form of image pyramids in a file. In this paper, the OpenSlide framework is used to load and preprocess the pathological slice image, and the image of the resolution on the fourth layer of the pyramid level is selected as the sample data. Before acquiring image patches, we need to preprocess the image, specifically shown in Fig. 2. Firstly, using median filtering and Gaussian filtering to remove noise and using the stain

normalization algorithm [25] to evenly color the image. Secondly, in order to visually reflect the color information of the image, the image is converted from the RGB to the HSV color space. Thirdly, binary image obtained by edge detection and image segmentation based on Canny operator. Then Morphological processing can be used to extract the fine contours of the foreground area. Finally, extracting image patches with a sliding window of a certain step size in the segmented foreground area.

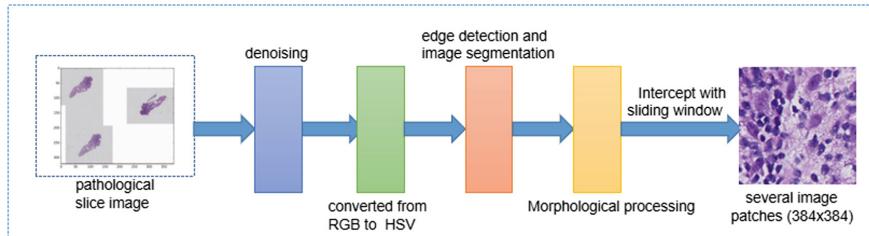


Fig. 2. The medical pathological image preprocessing system

3.2 Patch Reconstruction

Discriminator Network. In this section, we propose discriminator with small batch relative discrimination. Different from discriminator in SRGAN, small batch relativistic discriminator tries to predict the average probability that a set of sample images HRs are relatively more realistic than generated images SRs, as shown in Fig. 3.

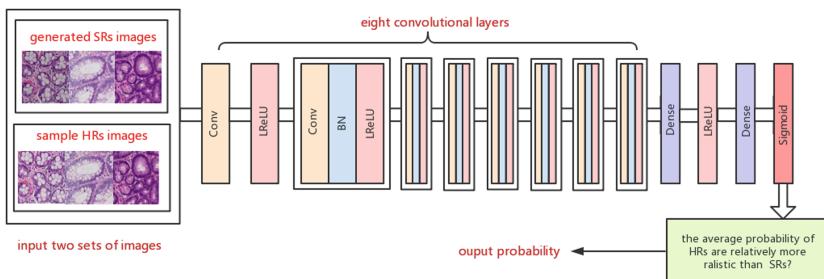


Fig. 3. The small batch relativistic discriminator

In the first part we have introduced process that the complete pathology slice is divided into several image blocks. If the previous discriminator is used for single absolute discrimination, the difference between the reconstructed image blocks is large, which affects the overall visual observation. However, the small batch relativistic discriminator we proposed can not only reduce the difference between reconstructed image blocks, but also learn the prior information of the sample images when guiding the generator training, which is beneficial to improve the quality of generated image. The discriminant formula of the improved discriminator is as follows:

$$\begin{aligned} D(HRs, SRs) &= \frac{1}{N} \sum_{i=1}^N \sigma(C(HR_i) - C(SR_i)) \rightarrow 1 \\ D(SRs, HRs) &= \frac{1}{N} \sum_{i=1}^N \sigma(C(SR_i) - C(HR_i)) \rightarrow 0 \end{aligned} \quad (1)$$

Where N represents the value of batch, σ is the sigmoid function and $C(\cdot)$ is the non-transformed discriminator output. Therefore the loss function for discriminator is:

$$Loss_D = -\log D(HRs, SRs) - \log(1 - D(SRs, HRs)) \quad (2)$$

Perceptual Loss. Inspired by perceptual similarity [26, 27], Ledig et al. [1] propose perceptual loss in SRGAN. Perceptual loss includes the content loss and the adversarial loss, and makes the reconstructed image richer and more realistic. Therefore, the loss function for generator is:

$$L_G = \alpha l_{pixel} + \lambda l_{VGG} + \eta l_{gen} \quad (3)$$

Where the ground-truth α , λ and η are the coefficients to balance different losses. And l_{pixel} , l_{VGG} and l_{gen} represent the pixel loss, feature loss and adversarial loss.

Pixel Loss. The MSE is used to calculate the pixel error between the generated image and the target image in SRGAN. However, MSE has a large loss of outliers, so the robustness is poor when dealing with large different data. And, when MSE deals with medical pathology images of which the shape, size and position of the same target vary widely, and the difference is large, it is difficult to exert the advantages of stability in the training process. In order to keep the stability of training, we propose to use Huber loss instead of MSE. The pixel loss is defined as:

$$l_{pixel} = \begin{cases} \frac{1}{2r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{HR} - I_{x,y}^{SR})^2, |I^{HR} - I^{SR}| \leq \delta \\ \frac{1}{r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} \delta |I_{x,y}^{HR} - I_{x,y}^{SR}| - \frac{1}{2} \delta^2, |I^{HR} - I^{SR}| > \delta \end{cases} \quad (4)$$

Where W, H are the width and height of input LR image, I^{HR} represents target HR image and I^{SR} represents generated HR image, r is the upscale factor, x, y represent the pixel coordinates of the image and δ represents the absolute error threshold between I^{HR} and I^{SR} .

Feature Loss. The feature loss is the loss of the ground truth HR image and the generated image features obtained by the VGG19 network at the fourth layer, so that the generated image detail features are closer to the real image. The pixel loss is defined as:

$$l_{VGG} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\theta_{i,j}(I_{x,y}^{HR}) - \theta_{i,j}(I_{x,y}^{SR}))^2 \quad (5)$$

Where $W_{i,j}, H_{i,j}$ represent the dimension of the feature map in the VGG19 network and $\theta_{i,j}$ represents the feature map obtained by the j-th convolution layer (after activation) before the i-th max pooling layer within the VGG19 network.

Adversarial Loss. The adversarial loss is based on the output probability of the improved discriminator. The adversarial loss is defined as:

$$l_{gen} = -\log(1 - D(HRs, SRs)) - \log D(SRs, HRs) \quad (6)$$

Training. The training method used in this paper is a back-propagation stochastic gradient descent algorithm, which performs gradient descent on the generation network and the discriminator network in turn. The algorithm calculates the gradient calculation formula and weight update formula of the j-th layer in the i-th iteration as following:

$$\Delta_{i+1} = \gamma\Delta_i - \eta \frac{\partial L}{\partial W_i^j} \quad (7)$$

$$W_{i+1}^j = W_i^j + \Delta_{i+1} \quad (8)$$

Where Δ, W represent the gradient and weight, γ, η represent the momentum and learning rate.

3.3 Result Evaluation

We have already introduced three methods for image quality evaluation. Among them, MOS is relatively large in workload and the results will be affected by factors such as the knowledge background and observation environment of the observer. PSNR and SSIM results are inconsistent with human subjective evaluation. The reconstructed medical images are mainly used for visual observation by doctors, and the reconstruction effect of the region of interest of the human eye directly affects subjective judgment. And, the human eye is more sensitive to the feature information of the image than the pixel information and structural information of the image. Therefore, the assessment of feature similarity (FSIM) [28] is more meaningful for medical image reconstruction studies.

4 Experiments

4.1 Experimental Condition

Firstly, let us introduce the experimental environment. The algorithm is run on a lab server with Ubuntu 16.04 and hardware devices configured with four Titan Xp Pascal

graphics cards. Our programming language is Python 2.7, and the network framework is TensorFlow 1.4.

Following SRGAN, we achieve the goal with upscale factor of $\times 4$ between LR image and HR image. And the input and output resolution of generator network are 96×96 and 384×384 , respectively. The value of batch is set to 8 in Eq. (1), the generator is trained using the loss function in Eq. (3) with $\alpha = 1$, $\lambda = 2 \times 10^{-6}$ and $\eta = 10^{-3}$, and the absolute error threshold is set to 2.5 in Eq. (4). We use the learning rate with 10^{-4} and momentum with 0.99.

4.2 Data

The Beijing Hospital which cooperated with the project provided 111 sections of gastroscope tissue. Then we obtain the dataset by the medical pathological image preprocessing system. The dataset is distributed as follows (Table 1):

Table 1. The dataset distribution

| | |
|------------------|------|
| Training dataset | 1254 |
| Testing dataset | 418 |
| Sum | 1672 |

4.3 Results

Firstly, we compare the variation of the loss function between our method and SRGAN. And the loss function visualization change graph is generated by tensorboard.

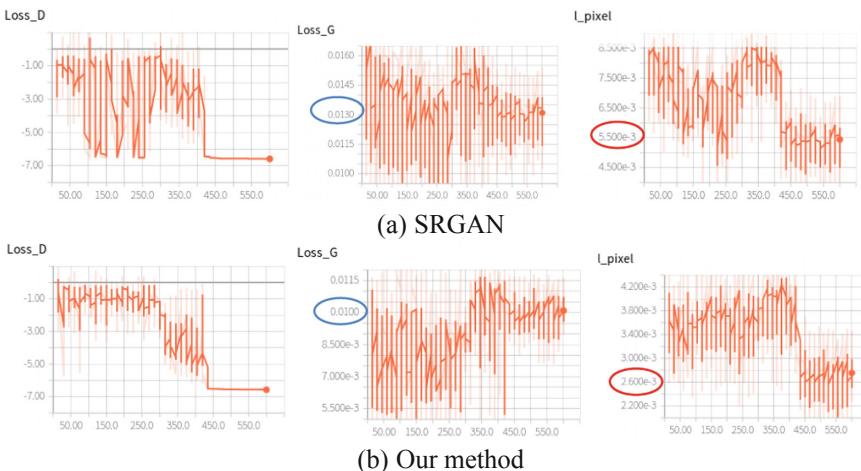


Fig. 4. The loss function including Loss_D, Loss_G and l_{pixel} of SRGAN and our method

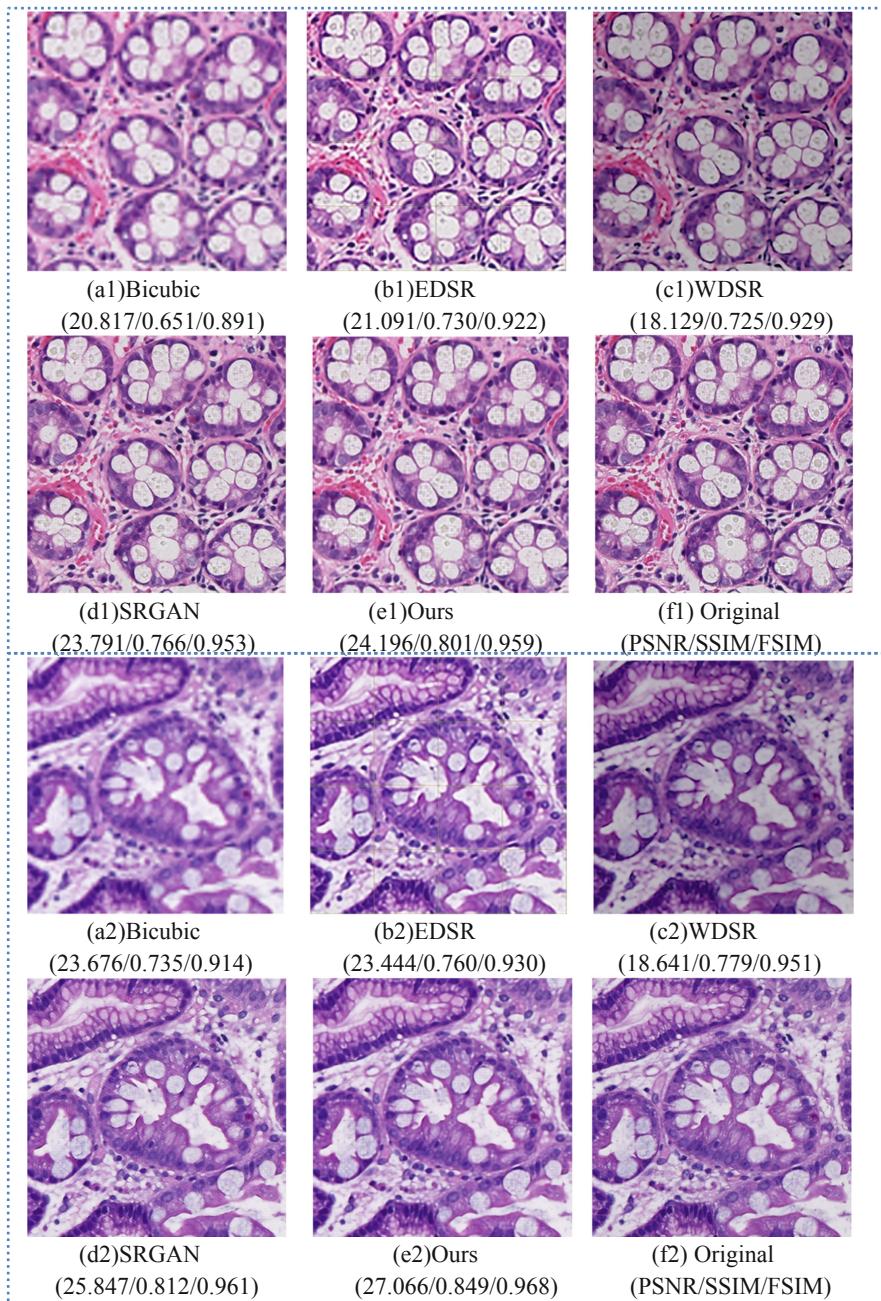


Fig. 5. The results of our method. Our method is better than the other methods in the three objective evaluation ways and the reconstructed image is closer to the original image in terms of color and cell texture restoration.

It can be observed from Fig. 4 that the Loss_D of our method with Huber loss is less volatile and relatively more stable during the training process. And we can find that the convergence value of the Loss_G and l_pixel of our method is smaller than SRGAN. These show that using Huber loss instead of MSE can maintain the training stability and converge the loss function to a smaller value. And we all know that the better the network converges, the better the performance of the network.

Then, we compare our model on our dataset with other SR methods including Bicubic, EDSR [16], WDSR [18], SRGAN [1]. And there we use PSNR, SSIM and FSIM to evaluate the results objectively.

We can observe that our method has improved the quality of reconstructed images to some extent, especially the restoration of color and intercellular texture details. Visually, the Fig. 5(c1) is more similar to the original image than the Fig. 5(b1), but the value of PSNR and SSIM inconsistent with visual perception. Therefore, the FSIM is more suitable for medical image reconstruction studies.

5 Conclusion

We have presented a method based on GAN to achieve better reconstruction results of medical images. Firstly, we designed a medical pathological image preprocessing system to extract tissue area image patches. And we proposed discriminator with small batch relative discrimination to improve the quality of reconstructed images by learning more prior information. Then we used Huber loss instead of the original MSE to maintain network more stable. Finally, we found that FSIM, which is consistent with the subjective perception of human vision, is more suitable for medical image reconstruction studies.

References

1. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z et al (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: CVPR
2. Kim J, Kwon Lee J, Mu Lee K (2016) Accurate image super-resolution using very deep convolutional networks. In: CVPR
3. Lai WS, Huang JB, Ahuja N, Yang MH (2017) Deep Laplacian pyramid networks for fast and accurate super-resolution. In: CVPR
4. Kim J, Kwon Lee J, Mu Lee K (2016) Deeply-recursive convolutional network for image super-resolution. In: CVPR
5. Tai Y, Yang J, Liu X (2017) Image super-resolution via deep recursive residual network. In: CVPR
6. Tai Y, Yang J, Liu X, Xu C (2017) Memnet: a persistent memory network for image restoration. In: ICCV
7. Gupta P, Srivastava P, Bhardwaj S, Bhateja V (2011) A modified PSNR metric based on HVS for quality assessment of color images. In: IEEE international conference on communication and industrial application (ICCIA), pp 1–4

8. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612
9. Wang Z, Simoncelli EP, Bovik AC (2003) Multi-scale structural similarity for image quality assessment. In: IEEE asilomar conference on signals, systems and computers, vol 2, pp 9–13
10. Johnson J, Alahi A, Fei-Fei L (2016) Perceptual losses for real-time style transfer and super-resolution. In: *ECCV*
11. Bruna J, Sprechmann P, LeCun Y (2015) Super-resolution with deep convolutional sufficient statistics. In: *ICLR*
12. Sajjadi MS, Schölkopf B, Hirsch M (2017) EnhanceNet: single image super-resolution through automated texture synthesis. In: *ICCV*
13. Dong C, Loy CC, He K, Tang X (2014) Learning a deep convolutional network for image super-resolution. In: *ECCV*
14. Kim J, Lee JK, Lee KM (2016) Accurate image super-resolution using very deep convolutional network. In: Computer vision and pattern recognition. IEEE, pp 1646–1654
15. Shi W, Caballero J, Huszar F et al (2016) Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Computer vision and pattern recognition. IEEE, pp 1874–1883
16. Lim B, Son S, Kim H et al (2017) Enhanced deep residual networks for single image super-resolution. In: Computer vision and pattern recognition, vol 1, no 2. IEEE, p 3
17. Zhang Y, Tian Y, Kong Y, Zhong B, Fu Y (2018) Residual dense network for image super-resolution. In: Computer vision and pattern recognition
18. Yu JH, Fan YC et al (2018) Wide activation for efficient and accurate image super-resolution. In: Computer vision and pattern recognition. IEEE
19. Rivenson Y, Göröcs Z, Günaydin H, Zhang Y, Wang H, Ozcan A (2017) Deep Learning Microscopy, Optical Society of America, pp 2334–2536
20. Metz L, Poole B, Pfau D et al (2016) Unrolled generative adversarial networks. arXiv preprint [arXiv:1611.02163](https://arxiv.org/abs/1611.02163)
21. Wang Z, Bovik AC (2006) Model image quality assessment. Morgan and Claypool Publishing Company, New York, pp 20–30
22. Avcibas I, Sanbur B, Sayood K (2002) Statistical evaluation of image quality measure. *J Electron Imaging* 11(2):206–213
23. Wang Z, Bovik AC, Sheikh HR et al (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612
24. Wang Z, Bovik AC (2002) A universal image quality index. *IEEE Trans Image Process* 9 (3):91–95
25. Macenko M, Niethammer M, Marron JS et al (2009) A method for normalizing histology slides for quantitative analysis. In: IEEE international conference on symposium on biomedical imaging: from nano to macro. IEEE Press, pp 1107–1110
26. Runa J, Sprechmann P, LeCun Y (2015) Super-resolution with deep convolutional sufficient statistics. In: *ICLR*
27. Gatys L, Ecker AS, Bethge M (2015) Texture synthesis using convolutional neural networks. In: *NIPS*
28. Morrone MC, Owens RA (1978) Feature detection from local energy. *Pattern Recogn* 6 (5):303–313



A Self-tuning Controller Design Method Based on LQG/LTR and Back Propagation Algorithm

Weimin Liu, Zhongzhi Hu, Jiqiang Wang^(✉), and Shengyi Liu

Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, China
jiqiang.wang@nuaa.edu.cn

Abstract. A self-tuning controller design method, based on the Back Propagation (BP) algorithm, was proposed to tune the LQG/LTR gain matrices directly without the need of selecting weighting matrices. The proposed controller design methodology is illustrated through the application to a turbo-shaft engine, and the simulation results demonstrate the improved design efficiency and the better controller performance.

Keywords: Aeroengine control · LQG/LTR · Back propagation algorithm · Self-tuning controller · Optimization design

1 Introduction

Controller design for an aeroengine has always been challenging since it is a complex nonlinear and dynamic system due to its operation in harsh environment with strong coupling between thermal, fluid, and mechanical effects. To meet the ever-demanding requirements on aeroengine performance and operability, multivariable control methods were used to control its fuel flow and geometrical variables [1]. With simple structure and considerable robustness [2], Linear Quadratic Gaussian/Loop Transfer Recovery (LQG/LTR) method was used in the controller design for different types of aeroengines, including turboshaft engines GE T700 and Allison T406, and turbofan engine PW F100 [3–5]. However, there are two potential issues when using LQG/LTR method: one is that during the controller design process, selecting its weighting matrices, which have a great influence on the controller performance, is experience dependent and therefore time consuming; and the other is that the controller obtained at the design point for a specified operating condition may lose performance and stability robustness if there are large modeling errors [6].

A lot of research work has been carried out by modifying the LQG/LTR method to address those two issues as mentioned above [7–15]. However, those methods are either computationally intensive or partially experience dependent when selecting the gain matrices.

In this paper, a self-tuning aeroengine controller design method, based on the BP algorithm, is proposed to tune the LQG/LTR gain matrices automatically without the need of selecting weighting matrices. In the following sections, the self-tuning process for the LQG/LTR gain matrices, based on the BP algorithm, is described in detail; The

controller design process is illustrated through the application to a turbo-shaft engine, GE T700, and the simulation results demonstrate the improved design efficiency and the better controller performance.

2 Self-tuning LQG/LTR Through BP Algorithm

In LQG/LTR controller, the key points of calculating K_f and K_c is to select the weight matrix Q and R . However, the selection of Q and R depends on the designer's experience, so it is time-consuming and labor-intensive to design LQG/LTR controller. Also, if there is a large modeling error, the performance of LQG/LTR controller will decline sharply. To solve this problem, this paper regards K_f and K_c as mapping matrices, and introduces BP to optimize K_f and K_c matrices, automatically.

The principle of optimizing gain matrix based on BP algorithm is shown in Fig. 1:

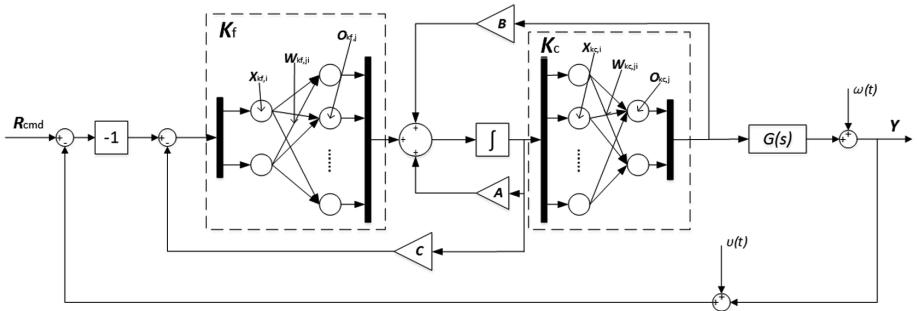


Fig. 1. Principle of optimizing gain matrix based on BP algorithm

In this paper, K_f and K_c are regarded as mapping matrices, so that they can be optimized by BP algorithm in LQG/LTR controller [16, 17].

Steps to tune gain matrices using BP algorithm are as follows:

Step 1: Obtain arbitrary K_f and K_c : Obtain arbitrary K_f and K_c based on LQG/LTR controller design method.

Step 2: Calculate forward propagation: for each moment, calculate forward propagation of information and get system actual output

Step 3: Calculate the error: calculate error according to the error definition. This article uses mean square error definition, which is commonly used while solving practical problems.

$$E(k) = \sum_{i=1}^m [\hat{y}_i(k) - y_i(k)]^2 \quad (1)$$

In Eq. (1), m represents output vector dimension, $\hat{y}_i(k)$ represents the i -th expected output value at current time, $y_i(k)$ is the i -th actual output at current time. In this paper, the expected output is represented by the command signal.

Step 4: Calculate back propagation: calculate back propagation of error and calculate adjusting value of each element in mapping matrix. The calculation of error back propagation to \mathbf{K}_f is taken as an example.

$$\Delta w_{kf,ji} = -\eta_{kf} \frac{\partial E}{\partial w_{kf,ji}} \quad (2)$$

In Eq. (2), η_{kf} represents optimization step size. $\frac{\partial E}{\partial w_{kf,ji}}$ can be calculated from Eq. (1):

$$\frac{\partial E}{\partial w_{kf,ji}} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial O_{kf,j}} \frac{\partial O_{kf,j}}{\partial w_{kf,ji}} \quad (3)$$

In Eq. (3), the first term is the partial derivative of the mean square error to the output:

$$\frac{\partial E}{\partial y_j} = \frac{\partial}{\partial y_j} \left\{ \sum_{i=1}^m [\hat{y}_i(k) - y_i(k)]^2 \right\} = -2 \left[\sum_{i=1}^m \hat{y}_i(k) - y_i(k) \right] \quad (4)$$

The second term is the partial derivative of the plant output to \mathbf{K}_f output:

$$\frac{\partial y_j}{\partial O_{kf,j}} \approx \frac{\Delta y_j}{\Delta O_{kf,j}} = \frac{y_j(k) - y_j(k-1)}{O_{kf,j}(k) - O_{kf,j}(k-1)} \quad (5)$$

When $O_{kf,j}(k)$ and $O_{kf,j}(k-1)$ differs little or equal, $\frac{\partial y_j}{\partial O_{kf,j}}$ tends to infinity, which is unreasonable during operation. Considering that the convergence speed of mapping matrix elements can be adjusted by the optimization step size η_{kf} , so this term only has to decide the convergence direction [18]. Therefore, in this paper, a sign function is used instead of $\frac{\partial y_j}{\partial O_{kf,j}}$:

$$\frac{\partial y_j}{\partial O_{kf,j}} = \text{sgn} \frac{y_j(k) - y_j(k-1)}{O_{kf,j}(k) - O_{kf,j}(k-1)} \quad (6)$$

The third term is the output value of the mapping matrix to elements in mapping matrix:

$$\frac{\partial O_{kf,j}}{\partial w_{kf,ji}} = \frac{\partial}{\partial w_{kf,ji}} \left[\sum_{k=1}^n w_{kf,jk} \cdot x_{kf,k}(k) \right] = x_{kf,i}(k) \quad (7)$$

Step 5: Parameter update: adjust each element of each mapping matrix according to calculation results in Step 4.

$$w_{kf,ji}(k+1) = w_{kf,ji}(k) + \Delta w_{kf,ji} \quad (8)$$

$$w_{kc,ji}(k+1) = w_{kc,ji}(k) + \Delta w_{kc,ji} \quad (9)$$

Step 6: Optimize judgment: judge whether the optimization end condition is satisfied. If it is not satisfied, jump to Step 2 to perform a new round of optimization operation, and if it is satisfied, stop optimization.

3 Aeroengine Controller Design

3.1 Design Point Simulation

The linear state and output equations for an engine are expressed as follows [19]:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + \Gamma w(t) \\ y(t) = Cx(t) + Du(t) + v(t) \end{cases} \quad (10)$$

In Eq. (10), x represents state vector, y represents output vector, u represents control vector, w and v represent independent system noises.

In Matlab/Simulink(2017b), for GE T700 turbo-shaft engine's linear model [3], optimization simulation of LQG/LTR controller gain matrices is completed.

First, an arbitrary LQG/LTR controller is designed based on 90% N_g design model of GE T700 turboshaft engine. Then optimize the LQG/LTR controller based on the optimization method proposed in this paper and set the optimization step size as 0.002. Comparison of system response before and after optimization are shown in Fig. 2:

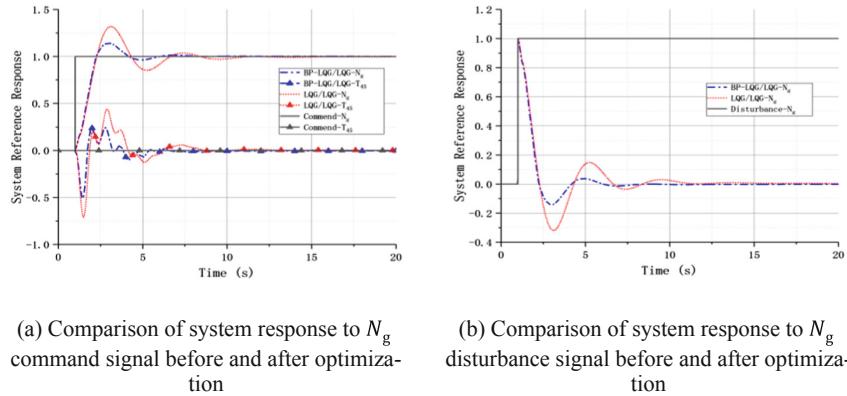


Fig. 2. Comparison of system response before and after optimization

The simulation results in Fig. 2 show that the optimized LQG/LTR controller has obvious advantages on overshoot and settling time when facing a step rotational speed command signal. The turbine inlet gas temperature can quickly converge at the equilibrium position without generating large fluctuations; Also, when facing a disturbance signal, the optimized LQG/LTR controller can control the plant back to equilibrium state. So, the optimized LQG/LTR controller has better performance.

3.2 Off-Design Point Simulation

It is inevitable that modeling errors will occur during engine modeling and linearization. When the modeling error is large, the performance of LQG/LTR controller will deteriorate. The linear state and output equations for an engine with modeling error can be expressed as:

$$\begin{cases} \dot{x}(t) = (\mathbf{A} + \Delta\mathbf{A})x(t) + (\mathbf{B} + \Delta\mathbf{B})u(t) \\ y(t) = \mathbf{C}x(t) \end{cases} \quad (11)$$

$\Delta\mathbf{A}$ and $\Delta\mathbf{B}$ can be obtained in reference [3]. In this section, high and low rotational speed modeling error, marked as state 1 and state 2, are obtained to verify proposed method.

Utilize the optimization algorithm proposed in this paper and set the optimization step size as 0.002. Comparison of system response before and after optimization with modeling error are shown in Fig. 3.

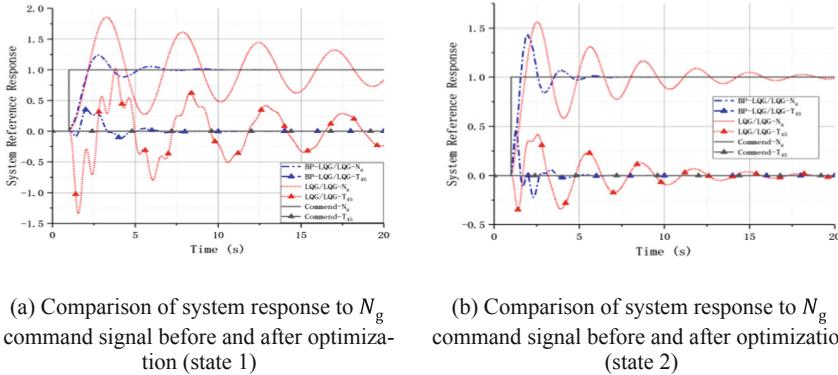


Fig. 3. Comparison of system response before and after optimization with modeling error

As can be seen in Fig. 3, when modeling error is large, the arbitrary LQG/LTR controller will have poor performance, even in some cases, it will directly diverge. However, the optimization method proposed in this paper can effectively optimize the LQG/LTR controller and improve the control performance, significantly.

4 Conclusion

In this paper, a self-tuning aeroengine controller design method, based on the BP algorithm, is proposed to tune the LQG/LTR gain matrices automatically without the need of selecting weighting matrices. In the following sections, the self-tuning process for the LQG/LTR gain matrices, based on the BP algorithm, is described in detail; The controller design process is illustrated through the application to a turbo-shaft engine,

GE T700, and the simulation results demonstrate the improved design efficiency and the better controller performance.

References

1. Yao H (2014) Full authority digital electronic control system for aero-engine. Aviation Industry Press, Beijing
2. Fan S (2008) Aero-Engine control system. Aviation Industry Press, Beijing
3. Pfeil WH, Athans M, Spang HA (1986) Multi-variable control of the GE T700 engine using the LQG/LTR design methodology. NASA Technical Report, NASA-CR-177080, NAS 1.26:177080, LIDS-P-1547
4. Moellenhoff DE, Rao SV, Skarvan CA (1991) Design of robust controllers for gas turbine engines. *J Eng Gas Turbines Power* 113(2):283–289
5. Athans M, Kapasouris P, Kappos E et al (1986) Linear-quadratic Gaussian with loop-transfer recovery methodology for the F-100 engine. *J Guidance Control Dyn* 9(1):45–52
6. Doyle J, Stein G (1979) Robustness with observers. *IEEE Trans Autom Control* 24(4):607–611
7. Zhang M, Sun P, Cao R et al (2010) LQG/LTR flight controller optimal design based on differential evolution algorithm. In: International conference on intelligent computation technology and automation, pp 613–616
8. Rongfu T, Weiming P, Junjian Z et al (2016) The LQG/LTR control optimization based on genetic algorithm for aeroengine. *Comput Technol Autom* 35(4):33–38
9. Das S, Pan I, Halder K et al (2013) LQR based improved discrete PID controller design via optimum selection of weighting matrices using fractional order integral performance index. *Appl Math Model* 37(6):4253–4268
10. Pan I, Das S (2013) Design of hybrid regrouping PSO–GA based sub-optimal networked control system with random packet losses. *Memetic Comput* 5(2):141–153
11. Haessig D (1995) Selection of LQG/LTR weighting matrices through constrained optimization. In: Proceedings of the American control conference, no 1. IEEE, pp 458–460
12. Ma R, Wu HT, Ding L (2016) Optimal LQG/LTR controller for small-scale unmanned helicopter based on artificial bee colony algorithm. *Control Decis* 31(12):2248–2254
13. Das S, Halder K (2014) Missile attitude control via a hybrid LQG-LTR-LQI control scheme with optimum weight selection. In: International conference on automation, control, energy and systems. IEEE, pp 115–120
14. Jianqiao Y, Guanchen L, Yuesong MEI (2011) Surface-to-air missile autopilot design using LQG/LTR gain scheduling method. *Chin J Aeronaut* 24(3):279–286
15. Yingqing G, Dan W, Huz Z (2002) Design of LQG/LTR multivariable feedback controller. *Aeroengine* 4:44–47
16. Kubat M (1999) Neural networks: a comprehensive foundation, by Simon Haykin, Macmillan. ISBN 0-02-352781-7. Cambridge University Press
17. Rumelhart DE, Hinton, GE (1986) Parallel distributed processing. The MIT Press, Cambridge
18. Cong S, Liang Y (2009) PID-like neural network nonlinear adaptive control for uncertain multivariable motion control systems. *IEEE Trans Ind Electron* 56(10):3872–3879
19. Jaw LC, Mattingly JD (2009) Aircraft engine controls: design, system analysis, and health monitoring. American Institute of Aeronautics and Astronautics



Research on Tracking Guidance Method for On-orbit Service Autonomous Homing Based on Proportional Guidance and Trajectory Planning

Qi Song, Xinlong Chen^(✉), Qiang Chen, Yajing Ren,
and Mengping Zhu

China Academy of Space Technology, Beijing 100094, China
chenxinlong@qxslab.cn

Abstract. It is a trend that the high orbit and high performance satellites are applied to on-orbit service. There are some advantages for tracking spacecraft to use proportional navigation timely in autonomous homing phase of rendezvous docking. In this paper, proportional navigation is derived from inertial coordinate system using optimal control method. The approximate treatment of C-W equation shows that proportional navigation is suitable for high orbit. The simulation results of proportional navigation and trajectory planning show that the fuel optimal effect of proportional navigation is good. There is no accumulated error in proportional navigation, which results in a good accuracy of reaching the target point. Because the terminal velocity cannot be effectively controlled, it is suitable for long-distance autonomous homing phase. Timely use of proportional navigation has certain advantages.

Keywords: Proportional navigation · Trajectory planning · Autonomous homing phase of rendezvous docking

1 Introduction

High-value satellites need on-orbit services. Tracking spacecraft is a main spacecraft in on-orbit service applications. The main operational processes of tracking spacecraft include the stages of long-range guidance, short-range guidance, fly around and final approximation [1–3]. Among which, the short range guidance phase in some of the references known as autonomous homing [4]. Autonomous homing begins from the tracking spacecraft to establish relative navigation with target spacecraft. Considering such factors as orbital altitude, tracking spacecraft performance and system scheme, the range of autonomous homing is also different. A large amount of literature [5–9] at home and abroad have studied the guidance and control of autonomous homing stage of space rendezvous, and most of the studies are based on trajectory planning method.

If the target spacecraft is a non-cooperative target, trajectory planning algorithm needs to be updated in real time, and the guidance and control algorithm is complicated, which requires real-time performance and high navigation accuracy, therefore, some researchers use proportional guidance method with simple engineering implementation,

follow-up flight and error correction ability to study autonomous homing phase [10–14]. In this paper, the optimization mechanism of proportional navigation is studied, and the feasibility of the application of proportional navigation in high orbit is revealed by formula deduction and simulation, and including the following ability and error tolerance of proportional navigation. The comparison between proportional navigation and trajectory planning shows the comprehensive advantages and constraints of proportional navigation in the application process.

2 Proportional Guidance Optimization Model

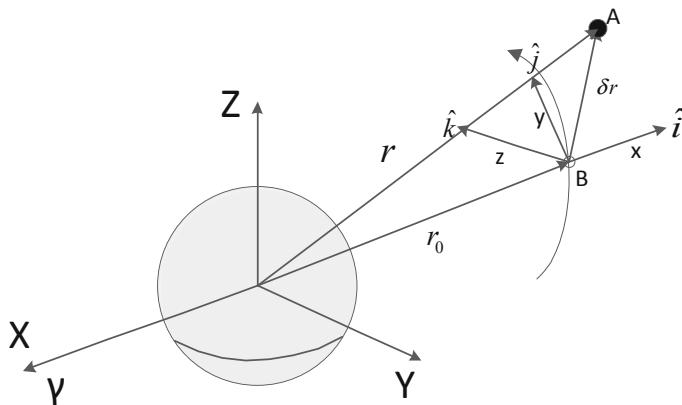


Fig. 1. Observation of A from a motion coordinate system fixed to B (Orbital coordinate system)

The mass center of the target spacecraft lies at the origin of the orbital coordinate system (Fig. 1). The X-axis points from the center of the earth to the target spacecraft. The Y-axis is perpendicular to the x-axis in the orbit plane towards the velocity direction. Assuming proportional navigation applies in orbit plane and C-W equation is as follows:

$$\delta\ddot{x} - 3n^2\delta x - 2n\delta\dot{y} = a_{cx} \quad (1a)$$

$$\delta\ddot{y} + 2n\delta\dot{x} = a_{cy} \quad (1b)$$

The orbital angular velocity is n , and the high-orbit angular velocity is very small. Assuming $n \approx 0$. Thrust acceleration on the x-axis is a_{cx} . Thrust acceleration on the y-axis is a_{cy} . C-W equation is transformed as follows:

$$\delta\ddot{x} = a_{cx} \quad (2a)$$

$$\delta\ddot{y} = a_{cy} \quad (2b)$$

The above formula shows that the orbital coordinate system approximates the inertial coordinate system (Fig. 2).

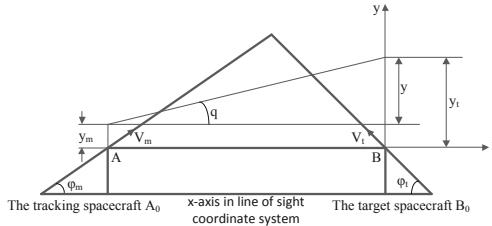


Fig. 2. The relative relationship between the tracking spacecraft and the target spacecraft (Line of sight coordinate system)

As shown above, in line-of-sight coordinate system, assuming V_m and V_t are respectively the speed of the tracking spacecraft and the target spacecraft, and also they are constants. At time t , the tracking spacecraft reaches point A, the target spacecraft reaches point B, the origin of coordinate system is point B, and the x-axis is the line of sight. According to the hypothesis, the xBy coordinate system is inertial.

According to small perturbation hypothesis on Y axis for tracking spacecraft and target spacecraft, the perturbation amounts are y_m and y_t . The reason for perturbation on the y-axis is that the y-axis is perpendicular to the line of sight and is the direction of force applied by proportional navigation. The position error of the two perturbations is $y = y_t - y_m$. The relative velocity is $V = \dot{y}$. The relative acceleration is $a_t - a_m$. Assuming that the target has no acceleration, it satisfies $a_t = 0$. The dynamic model is as follows:

$$\begin{cases} \dot{y} = V \\ \dot{V} = -a_m \end{cases} \quad (3)$$

The objective function of the optimal guidance rate is:

$$\min J = \min \left[S \frac{y(t_f)^2}{2} + \frac{1}{2} \int_0^{t_f} a_m^2(t) dt \right] \quad (4)$$

$S \frac{y(t_f)^2}{2}$ is the success rate of tracking the target. If S tends to be infinite, the success rate of reaching the target is 100%. According to the minimum principle in the optimal control, the optimal control is obtained as follows.

$$a_m = \frac{3}{(t_f - t)^2} y(t) + \frac{3}{(t_f - t)} V(t) \quad (5)$$

According to the hypothesis of small perturbation, the line-of-sight angle is $q = \frac{y(t)}{\dot{\rho} \cdot (t_f - t)}$. $\dot{\rho}$ is the relative velocity between tracking spacecraft and target spacecraft in line-of-sight coordinate system, also known as line-of-sight velocity rate. The derivation of time by q is

$$\dot{q} = \frac{y(t)}{\dot{\rho} \cdot (t_f - t)^2} + \frac{V(t)}{\dot{\rho} \cdot (t_f - t)} \quad (6)$$

Formula (6) is introduced into Formula (5) to obtain:

$$a_m = 3\dot{\rho}\dot{q} \quad (7)$$

a_m is the acceleration generated by thrust in the vertical line of sight.

The benefits of proportional navigation are as follows: (1) It is easy to obtain \dot{q} from detection technology in engineering. (2) It is not necessary to consider the initial time state of tracking spacecraft and target spacecraft.

3 Trajectory Planning

The optimal performance index function of energy consumption is defined as follows:

$$J = \frac{1}{2} \int_{t_0}^{t_f} u^T R u dt \quad (8)$$

R is a weighted diagonal matrix. $R = \text{diag}\{R_x, R_y, R_z\}$

The optimal problem can be described as follows:

$$\min_u J = \int_{t_0}^{t_f} L(x, u, \tau) d\tau = \frac{1}{2} \int_{t_0}^{t_f} u^T R u dt \quad (9)$$

The constraint conditions are: $\dot{x} = Ax + Bu$

The above formulas are written in the form of matrix produced by C-W equation, the expansion forms are as follows:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 3n^2 & 0 & 0 & 0 & 2n & 0 \\ 0 & 0 & 0 & -2n & 0 & 0 \\ 0 & 0 & -n^2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \\ \delta \dot{x} \\ \delta \dot{y} \\ \delta \dot{z} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{cx} \\ a_{cy} \\ a_{cz} \end{bmatrix} \quad (10)$$

The initial position is $x(0) = x_0$. The terminal position is $x(t_f) = x_{t_f}$.

$$\text{In the above formula, } A = \begin{bmatrix} 0_{3 \times 3} & I_{3 \times 3} \\ A_1 & A_2 \end{bmatrix}, B = \begin{bmatrix} 0_{3 \times 3} \\ I_{3 \times 3} \end{bmatrix}$$

The approximate analytical solution of the above optimization problem can be obtained. According to the minimum principle of optimal control, the following conclusions can be obtained:

$$u = -R^{-1}\lambda_v \quad (11)$$

$$\lambda_v = [\Phi_{11}, \Phi_{12}]A_0 \quad (12)$$

$$A_0 = \left\{ \int_{t_0}^{t_f} \begin{bmatrix} \Phi_{12}(t_f - s) \\ \Phi_{22}(t_f - s) \end{bmatrix} [\Phi_{11}(s) \quad \Phi_{12}(s)] ds \right\}^{-1} [\Phi x_0 - x(t_f)] \quad (13)$$

In the above formulas, the following sub-formulas can be obtained.

$$\Phi_{11}(t) = \begin{bmatrix} 4 - 3 \cos nt & 0 & 0 \\ 6(\sin nt - nt) & 1 & 0 \\ 0 & 0 & \cos nt \end{bmatrix} \quad (14)$$

$$\Phi_{12}(t) = \begin{bmatrix} \frac{\sin nt}{n} & \frac{2}{n}(1 - \cos nt) & 0 \\ \frac{2}{n}(\cos nt - 1) & \frac{4}{n}\sin nt - 3t & 0 \\ 0 & 0 & \frac{1}{n}\sin nt \end{bmatrix} \quad (15)$$

$$\Phi_{21}(t) = \begin{bmatrix} 3n \sin nt & 0 & 0 \\ 6n(\cos nt - 1) & 0 & 0 \\ 0 & 0 & -n \sin nt \end{bmatrix} \quad (16)$$

$$\Phi_{22}(t) = \begin{bmatrix} \cos nt & 2 \sin nt & 0 \\ -2 \sin nt & 4 \cos nt - 3 & 0 \\ 0 & 0 & \cos nt \end{bmatrix} \quad (17)$$

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \quad (18)$$

4 Simulation Analysis

Using proportional navigation, in the orbital coordinate system, without considering the influence of gravity term, the coordinate system is approximately an inertial system, assuming that the initial position of the tracking spacecraft is $[-7500; -7500]$, the initial velocity is $[40; 1]$, and the termination position is set to $[0; -100]$. It takes

453.46 s to reach the target spacecraft. For tracking spacecraft, the minimum distance to the target spacecraft is 0.0637 m, the termination position is [0.0143; -99.938], the termination velocity is [4.9114; 21.2662], the total fuel consumption ΔV is 41.0701 m/s, and the maximum thrust acceleration that is needed is 0.2275 m/s². The above data shows that proportional navigation is a very effective guidance method in inertial coordinate system. Because high orbit is mainly considered in this paper, the simulation diagram is not put here to illustrate.

Using proportional navigation, in the orbital coordinate system, the gravity term is considered and the C-W equation is used to simulate the tracking target spacecraft. The initial position of the tracking spacecraft is assumed to be [-7500; -7500], the initial velocity is [40; 1], and the termination position is set to [0; -100]. It takes 468.39 s to reach the target spacecraft. The minimum distance to the target spacecraft is 0.0371 m, the termination position is [0.0066; -99.9634], the termination velocity is [3.6848; 20.4803], and the total fuel consumption ΔV is 43.34 m/s. The maximum thrust acceleration is 0.2275 m/s². The simulation results above show that the proportional navigation can guide the autonomous homing phase of the high orbit very well (Figs. 3 and 4).

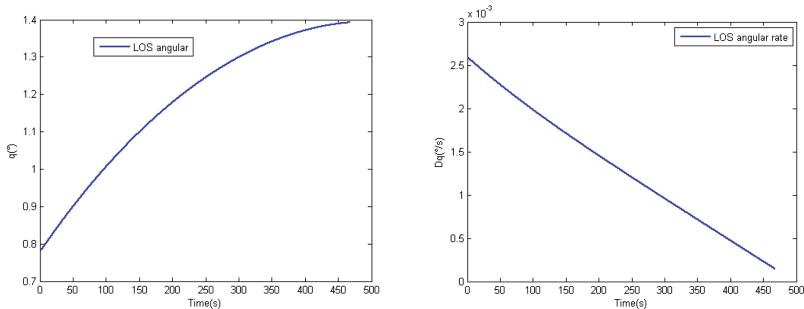


Fig. 3. Relationship between line of sight angle and line of sight angular acceleration with time

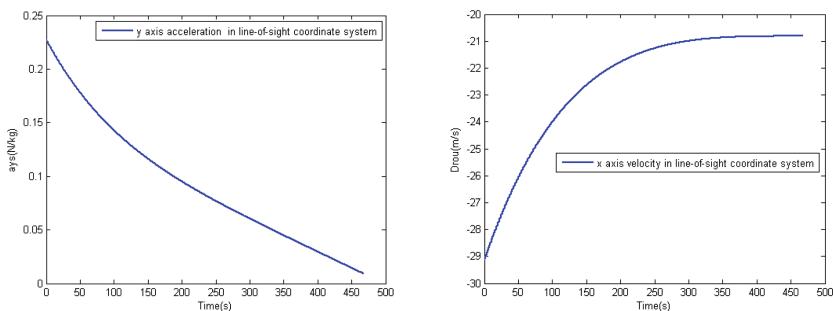


Fig. 4. Relationship between vertical line-of-sight acceleration and line-of-sight X-axis velocity of proportional navigation with time

The simulation of proportional navigation for the orbit at a height of 780 km shows that the proportional navigation is invalid due to the gravity term, which indicates that

proportional navigation can't be used in low-earth orbit. The simulation details are not described here. The disadvantages of proportional navigation are that the velocity of the termination position and all autonomous homing phase are uncontrollable.

Using the trajectory planning, considering the gravity term in the orbital coordinate system, the C-W equation is used to simulate, assuming that the initial position of the tracking spacecraft is $[-7500; -7500]$, the initial velocity is $[40; 1]$, the termination position is set to $[0; -100]$, and the termination velocity is set to $[3.6848; 20.4803]$. It takes 468.39 s to reach the target spacecraft. The minimum distance to the target spacecraft is 0.2332 m, the termination position is $[-0.1697; -99.8401]$, the termination velocity is $[3.6841; 20.481]$, and the total fuel consumption ΔV is 44.0565 m/s . The maximum thrust acceleration is 0.1894 m/s^2 . Thrust acceleration of trajectory planning is relatively small, which reduces the performance requirements of the engine, this is one of the advantages of trajectory planning. But on the whole, the simulation results show that there are accumulative errors in trajectory planning, so the simulation results of some parameters are not as good as that of proportional navigation.

The simulation results show that the fuel consumption of proportional navigation is less than that of trajectory optimization, which also proves that proportional navigation is the result of the comprehensive optimization of the minimum fuel and the reach to the target point, and the trajectory planning method is not as effective as proportional navigation method (Figs. 5 and 6).

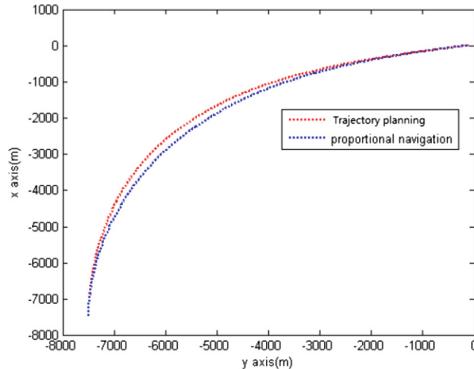


Fig. 5. Trajectory planning and proportional guidance trajectory algorithm

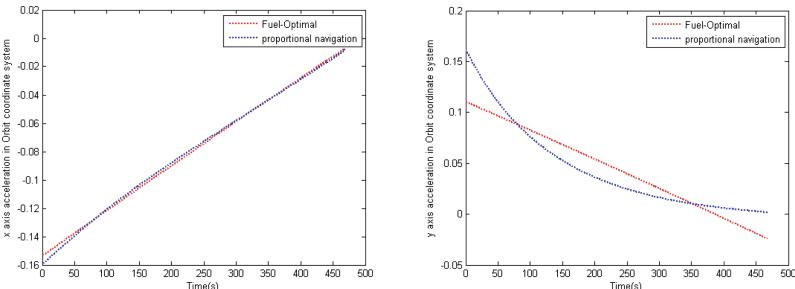


Fig. 6. Time-varying charts of thrust acceleration X and Y directions for trajectory planning and proportional navigation

The trajectories caused by the models of the trajectory planning and proportional navigation are different, and the difference of thrust acceleration of Y axis is the biggest.

Proportional guidance cannot effectively control terminal speed, but trajectory planning can. If the tracking spacecraft needs zero speed at the terminal of the autonomous homing phase to adjust the next operation, then the impulse thrust applied by proportional navigation easily exceeds the capability of the engine, and trajectory planning control can make the terminal reach zero very well. In addition to setting the termination speed as [0; 0], keeping the other parameters unchanged, the trajectory planning simulation results are as follows: it takes 468.39 s to reach the target spacecraft. The minimum distance to the target spacecraft is 0.4808 m, at the position of [-0.1145; -99.533], the termination speed is [0; 0.002], the total fuel consumption ΔV is 65.73 m/s, and the maximum thrust acceleration is 0.2435 m/s² through the fuel optimal trajectory planning.

Proportional guidance has ability of maneuvering followed flight. Assuming that the initial position of the tracking spacecraft is [-7500; -7500], the initial velocity is [40; 1], and the initial position of the rendezvous target is set to [0; -100], the rendezvous target moves forward to the Y axis at the speed of 4 m/s. The tracking spacecraft follows and eventually rendezvous with the target. It takes 624.87 s to get close to the target in a minimum distance of 0.0219 m, the termination position is [0; 2399.5], the termination speed is [-0.0031; 18.7884], and the total fuel consumption ΔV is 46.8956 m/s, in which the maximum thrust acceleration is 0.2275 m/s². The simulation results show that the proportional navigation has the ability of long-range follow-up and the relative speed at the rendezvous is large. Proportional guidance can be used in missile interception, so it has strong follow-up ability. The simulation only shows that proportional guidance has follow-up ability in high orbit space (Fig. 7).

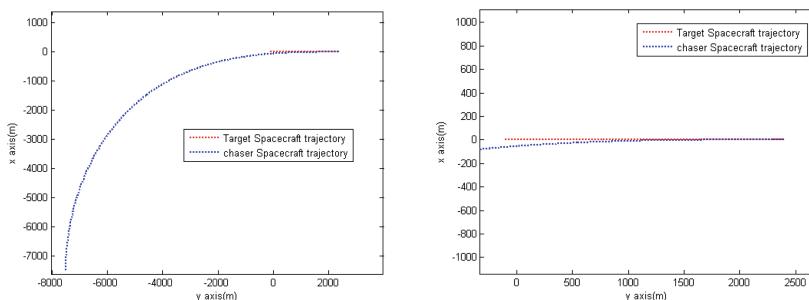


Fig. 7. Tracking spacecraft follow-up trajectory ability

5 Conclusion

From the perspective of optimization mechanism, this paper deduces and analyses the advantages of proportional navigation in high orbit, and compares it with the trajectory planning method based on C-W equation. The conclusions are as follows:

- (1) Proportional guidance conforms to the optimal control theory and takes into consideration both fuel optimum and follow-up capability. Under the condition of high-orbit, the fuel optimal trajectory planning is excellent in terms of fuel saving, but the trajectory planning has no follow-up capability, and the trajectory planning has accumulated errors. If the accuracy of trajectory planning measurement and calculation cannot be improved, the final effect is not as good as proportional navigation. Proportional guidance has the advantage of not having to consider the initial state parameters in detail, but it cannot effectively control the speed of the rendezvous point and the rendezvous time. Trajectory planning needs to know the location, speed and time of initial and rendezvous points, so that it can effectively control the position, speed and time of rendezvous points. The speed of proportional navigation rendezvous cannot be effectively controlled. For on-orbit service mission, it is only suitable for autonomous homing phase guidance with relatively long distance, and there is a margin of final approaching distance to prevent collision between tracking spacecraft and target spacecraft.
- (2) The traditional proportional navigation acceleration is applied in the vertical line of sight direction, so the speed close to the target point is not effectively controlled. If the acceleration in line of sight direction is controlled, the time and rendezvous speed can be effectively controlled. However, the optimal control of fuel and follow-up flight for traditional proportional navigation cannot be guaranteed. The traditional proportional navigation has only one chance to hit the target. If the target is repeatedly followed after rendezvous, the force in the line of sight must be adjusted. It is too expensive and risky to test directly in GEO orbit. Therefore, it's a common idea to verify technology in low orbit and apply it in high orbit. However, proportional navigation algorithm cannot be tested in low-orbit under the influence of gravity term.
- (3) The measurement of line-of-sight motion parameters and GNC algorithm are simple, so micro-nano spacecraft have certain advantages in using proportional navigation in autonomous homing phase of orbit, and have a wide application prospect. There are still a lot of research and development space based on traditional proportional navigation.

References

1. Zhu R (2007) Rendezvous and docking techniques of spacecraft. National Defense Industry Press
2. Zhu R, Tang Y (2005) Design methods for the closing and fly-around of space rendezvous. Chinese Space Sci Technol 25(1)
3. Cheng Y (2014) Study on relative navigation and flying-around of non-cooperative space crafts. Harbin Institute of Technology
4. Tang G, Luo Y, Zhang J (2007) Rendezvous and docking the mission planning for space flight. Science Press, Beijing
5. Zhou L (2011) Design and optimization of orbit transfer and flying trajectory of spacecraft. Harbin Institute of Technology

6. Wu S (2012) Research on control of chaser spacecraft with proximity to space target. Harbin Institute of Technology
7. Chen T, Xu S (2008) Research on guidance law of autonomous approach for spacecrafts flying in elliptical orbits. *J Astronaut* 29(6)
8. Wang C, Qu Y (2013) Trajectory optimization method for spacecraft orbit transfer with finite thrust. *J Southwest Jiaotong Univ* 48(2)
9. Lei T (2009) Study on optimal orbit design for spacecraft with finite thrust. Harbin Institute of Technology
10. Qi Z (2018) Design of missile guidance and control system. China Astronautic Publishing House
11. Li J, Li H (2009) Homing mission design of space rendezvous based on proportional guidance. *Flight Dyn* 27(4)
12. Li X (2011) Guidance method of fast fly-around relative to target spacecraft. Harbin Institute of Technology
13. Tian Y (2010) Study on guidance methods of fly-by approaching observation of space target. Graduate School of National University of Defense Technology
14. Tang Y (2006) Research on method of spacecraft's rendezvous with maneuvering target based on sight coordinate system. Harbin Institute of Technology



SOC Estimation for Lithium Battery Based on Segmented Model UKF Filter

Xue Li^(✉), Sijia Cao, Zhishu Xu, Zongshuai Yue, and Qiaochu Guo

The 18th Institute of China Academy of Launch Vehicle Technology,
Beijing, China
13488889707@163.com

Abstract. The rapidness and accuracy of the state of charge (SOC) estimation is the key technology in the battery management system. To solve the problems in SOC estimation caused by the inherent dynamic and nonlinear property of a lithium battery, first, an improved accurate measurement model is proposed by using segmented model method, in order to mitigate the negative influence of Polarization Effect. And then, the detail procedures and algorithms for lithium battery SOC estimation based on segmented model Unscented Kalman Filter (UKF) are given. Finally, the accuracy and the convergence rate of the segmented model UKF filter are analyzed. Experiments show that the segmented model UKF filter can be used to estimate the SOC quickly and accurately with an estimation error of 3%, which saves half time than the global model. Meanwhile, the adjustment of the initial value of UKF filter does not affect the estimation accuracy.

Keywords: Agricultural robot · UKF filter · Lithium battery · State of Charge

1 Introduction

Lithium batteries are widely used in communication, defense, robotic industries as the main power source for their advantages such as long cycling life, good temperature characteristic, high stability and relatively low cost. The working status of lithium battery in the above applications are an important factor related to system life and reliability. The modelling of lithium battery and estimation of State of Charge (SOC) are methods to ensure the safety and reliability of lithium battery, and hence make this technology one of the most key technologies in lithium battery management system [1, 2]. The voltage, output current, internal resistance of battery and the environmental temperature and etc. are all related to SOC, and they have non-linear characteristic.

There are 4 commonly used SOC estimation methods:

- (1) Estimation based on the measurement of identification parameters of battery, e.g. open circuit voltage method and internal resistance method.
- (2) Estimation based on ampere-hour metering, the estimation result will be affected by accumulated time-elapse error.

- (3) Kalman filter based SOC estimation. This method exhibits excellent estimation accuracy and can fix the initial error, but the estimation accuracy highly relies on the correctness of battery model [3].
- (4) Neural-network based SOC estimation [4–6]. Neural-network is an algorithm with non-linear behavior, parallel structure and learning ability. However, this method highly relies on the training method and training data.

In recent years, as the improvement of Kalman filter algorithm, various Kalman filter algorithms used for non-linear systems are being introduced. Among them, the most popular algorithms are: Unscented Kalman Filter (UKF) and Cubature Kalman Filter (CKF). Compared to Extended Kalman Filter (EKF), UKF and CKF need similar amount of calculations but with the result accuracy of 3-order Taylor series (while EKF only has 1-order accuracy) [11]. In lower dimensions ($n < 3$) where UKF gets a higher accuracy, while for higher dimensions ($n \geq 3$) CKF has a better result or requires less calculations [12].

In this article, a lithium battery that is installed in autonomous agricultural rover which can be used in crop harvesting, soil sampling and analyzing is to be considered. The structure of lithium battery is 3S2P and the voltage of which is 11.1 V. To achieve high accuracy of SOC estimation, firstly a segmented fitting method is proposed to enhance the accuracy of battery model, which lowered the error to 107 mV from original value of 224 mV; then non-linear Kalman filter algorithm is adopted; in order to implement SOC estimation on embedded equipment, a low-order UKF algorithm is used; and at last the estimation accuracy and robustness are verified by experiments.

2 SoC Model of Battery

The main parameters that being used for estimation of the SOC are open circuit voltage, output current and ambient temperature. In the agriculture rover, the lithium battery works in the environment with stable ambient temperature of 20–30 °C, so as to achieve the maximum output power. Therefore only the relationship of open circuit voltage and output current with the SOC is to be discussed.

2.1 Process Model

The SOC of battery with the elapsing of time t can be represented as [13]:

$$x(t) = x(0) - \int_0^t \frac{\eta i(\tau)}{Q_n} d\tau \quad (1)$$

In Eq. (1):

η —Discharge ratio(not considering ambient temperature)

Q_n —Nominal capacity of battery

$i(\tau)$ —Output current under time τ

After discretize Eq. (1), we will get:

$$x_{k+1} = x_k - \left(\frac{\eta \Delta t}{Q_n} \right) i_k \quad (2)$$

In Eq. (2):

Δt —Sample time

In this article, thanks to the working environment with constant ambient temperature, only the output current is taken into account. The steady state output current of agriculture rover is about 0.1C, while the output current with maximum power output of agriculture rover is 0.35C. The discharge ratio is acquired by testing different discharge current and then using Eq. (3)

$$\eta = 1/(ai^2 + bi + c) \quad (3)$$

2.2 Estimation Model

The relationship between estimated open circuit voltage and SOC status satisfies the empirical equations. There are several empirical equations [13]:

Shepherd Model:

$$z_k = E_0 - Ri_k - \frac{K_i}{x_k} \quad (4)$$

Unnewehr Model:

$$z_k = E_0 - Ri_k - K_i x_k \quad (5)$$

Nernst Model:

$$z_k = E_0 - Ri_k - K_2 \ln x_k + K_3 \ln(1 - x_k) \quad (6)$$

Combined Estimation Model:

$$z_k = K_0 - Ri_k - \frac{K_1}{x_k} - K_2 x_k + K_3 \ln x_k + K_4 \ln(1 - x_k) \quad (7)$$

In order to get better estimation accuracy, the Combined Estimation Model is adopted. In Eq. (7):

x_k —SOC

R —Internal resistance of battery

K_i —Polarization resistance

The above parameters can be calculated by least square fitting.

In Fig. 1, the global fitting estimation resulted in large error at two ends of X-axis which represents SOC, the maximum error was as large as 224 mV. While by using

segmented fitting method, the error in whole SOC range was well controlled, the maximum error was reduced to 107 mV.

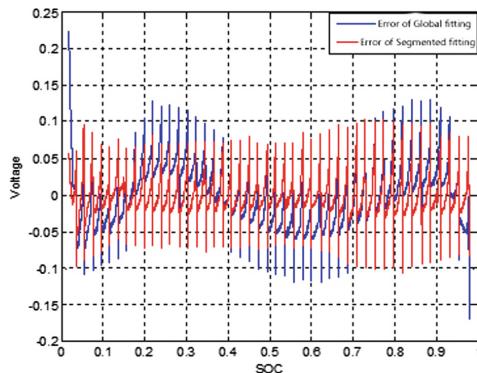


Fig. 1. Estimation error comparison between global fitting and segmented fitting results

Thanks to the large non-linearity of SOC with open circuit voltage, especially when SOC approaches 100% or 0%, so there would be a large estimation error if global fitting was adopted. In order to get better model accuracy, a segmented fitting method is proposed. The SOC is segmented into 5 segments of [90,100], [70,90], [30,70), [15,30), [0,15), and the least square fitting is applied to each of the 5 segments respectively.

3 SoC Estimation Based on UKF

UKF is a kind of non-linear Kalman filter, which is based on the concept of “It is much easier of probability distribution approximation compared to non-linear function approximation”. The core of UKF is UT transformation, i.e. choose $2n + 1$ weighted Sigma points to approximate the average value of state variables. The mean square value and variance of Sigma points being non-linear function propagated can result in 3-order accuracy of the true value of non-linear function, which makes the accuracy higher than EKF and mitigated the drawback of divergence of EKF. The easy divergence of EKF makes it only fitted for less non-linearity functions.

3.1 Theory of UKF

Assume the state equation and estimation equation of non-linear system is:

$$\begin{cases} x_{k+1} = f(x_k, u_k) + w_k \\ z_k = h(x_k) + v_k \end{cases} \quad (8)$$

In Eq. (8),

x_k —State vector

u_k —Input vector

w_k —System noise

v_k —Measurement noise

w_k and v_k are zero-average-value white noise and do not relate to each other, and their mean square root values are Q_k and R_k respectively.

(1) Initialization

$$\hat{x}_0 = Ex_0 \quad (9)$$

$$P_0 = E(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T \quad (10)$$

(2) Calculate sampling point (Sigma point)

$$\chi_k = \left[\hat{x}_k \hat{x}_k + \sqrt{(n + \lambda)(P_k)} \hat{x}_k - \sqrt{(n + \lambda)(P_k)} \right] \quad (11)$$

(3) Time update

$$\chi_{k+1|k} = f(\chi_k, u_k) \quad (12)$$

$$\hat{x}_{k+1|k} = \sum_{i=0}^{2n} W_i^m \chi_{i,k+1|k} \quad (13)$$

$$P_{k+1|k} = \sum_{i=0}^{2n} W_i^c \left[\chi_{i,k+1|k} - \hat{x}_{k+1|k} \right] \left[\chi_{i,k+1|k} - \hat{x}_{k+1|k} \right]^T + Q_k \quad (14)$$

$$\mathcal{Z}_{k+1|k} = h[\chi_{k+1|k}] \quad (15)$$

$$\hat{z}_{k+1|k} = \sum_{i=0}^{2n} W_i^m \mathcal{Z}_{i,k+1|k} \quad (16)$$

(4) Measurement update

$$P_{yy} = \sum_{i=0}^{2n} W_i^c \left[\mathcal{Z}_{i,k+1|k} - \hat{z}_{k+1|k} \right] \left[\mathcal{Z}_{i,k+1|k} - \hat{z}_{k+1|k} \right]^T + R_k \quad (17)$$

$$P_{xy} = \sum_{i=0}^{2n} W_i^c \left[\chi_{i,k+1|k} - \hat{x}_{k+1|k} \right] \left[\mathcal{Z}_{i,k+1|k} - \hat{z}_{k+1|k} \right]^T \quad (18)$$

$$\mathcal{K}_{k+1} = P_{xy} P_{yy}^{-1} \quad (19)$$

$$\hat{x}_{k+1} = \hat{x}_{k+1|k} + \mathcal{K}_{k+1} (z_{k+1} - \hat{z}_{k+1|k}) \quad (20)$$

$$P_{k+1} = P_{k+1|k} - \mathcal{K}_{k+1} P_{yy} \mathcal{K}_{k+1}^T \quad (21)$$

In the above equations, the variables are calculated as follows:

$$\lambda = \alpha^2(n + \kappa) - n \quad (22)$$

$$W_0^m = \lambda / (n + \lambda) \quad (23)$$

$$W_0^c = \lambda / (n + \lambda) + (1 - \alpha^2 + \beta) \quad (24)$$

$$W_i^m = W_i^c = 1 / \{2(n + \lambda)\} \quad i = 1, 2, \dots, 2n \quad (25)$$

3.2 UKF Based Segmented Fitting Estimation of SOC

According to the process model and estimation model of lithium battery, assume the SOC to be the state variable x_k , and output current i_k to be the input variable u_k , open circuit voltage to be estimated variable z_k , set sample time to be 0.1 s, and the equation for lithium battery will be derived as follow:

$$\begin{cases} x_{k+1} = x_k - \left(\frac{0.1u_k}{au_k^2 + bu_k + c} \right) \\ z_k = K_0 - Ru_k - \frac{K_1}{x_k} - K_2 x_k + K_3 \ln x_k + K_4 \ln(1 - x_k) \end{cases} \quad (26)$$

Perform estimation according to Eqs. (11)–(21), in which the system order n is set to 1, α is set to 1e-3, κ is set to 0, β is set to 2 by Gauss distribution.

In the process of time update, when calculating $z_{k+1}|k$ using Eq. (15), different segmented fitting estimation model should be chosen according to the variation of current estimated state variable x_k .

4 Analysis of Experimental Results

4.1 Acquisition of Experimental Data

In order to verify the effectiveness of the proposed algorithm, the lithium battery was tested at first. The lithium battery module used in testing features a nominal voltage of 11.1 V and nominal capacity of 20 Ah, the standard charging current was 0.2C and maximum discharging current was 0.5C.

The SOC estimation algorithm was tested under various discharging current. Under room temperature of 25 °C, tests were performed under different discharging ratios (0.1C, 0.2C and 0.3C) respectively. According to HPPC test procedure, the battery under test was charged to cut-off voltage using constant current at first, and then charge to cut-off current in constant voltage charging mode. The battery was intermittent discharged, i.e. discharging for 5 min and disconnected for 10 min. During the test, high accuracy voltage and current measuring devices were utilized to record the working status of lithium battery with 10 Hz sampling frequency. In Fig. 2, the data of real-time open circuit voltage, current and SOC under 0.1C are shown.

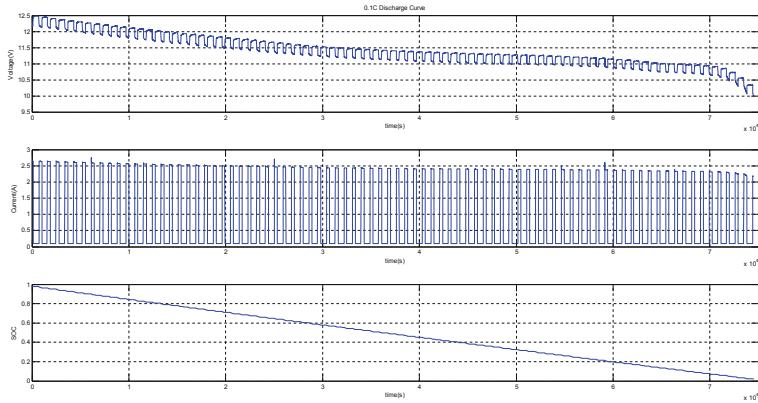


Fig. 2. Curve under 0.1C of discharging

4.2 Experimental Results

In order to verify the effectiveness of segmented model UKF, by using the measured battery discharging current and open circuit voltage under 0.1C and 0.3C, the segmented model UKF was adopted to estimate the SOC. The estimated results and errors are shown in Fig. 3.

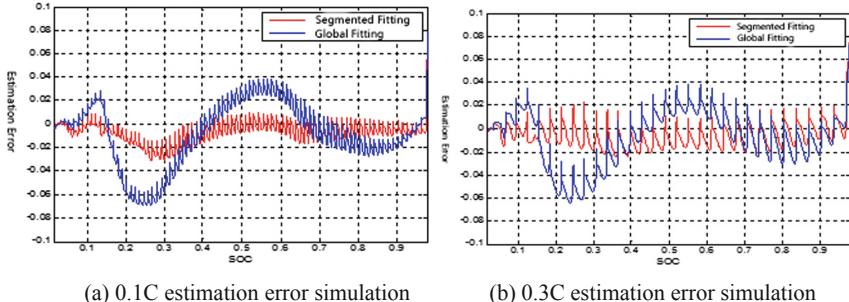


Fig. 3. UKF estimation error simulation under 0.1C and 0.3C discharging current

In Table 1, we can see that the estimation accuracy of segmented model is better than global model. Take 0.2C discharging current as an example, the maximum estimation error of segmented model is 2.5% meanwhile which of global model is 6.7%, hence the segmented model has 2.7 times more accurate compared to global model.

In Table 2, it shows that the convergence speed becomes slower along with the increasing of P_0 , and the initial value of SOC estimation has little effect on the convergence speed. Meanwhile there is little effect for different initial values have on the convergence speed, finally they will be converged to the same curve.

Table 1. Comparison of estimation errors under different discharging current

| | 0.1C | 0.2C | 0.3C |
|-------------------|------|------|------|
| Segmented fitting | 3.1% | 2.5% | 2.3% |
| Global fitting | 6.9% | 6.7% | 6.4% |

Table 2. The effect on convergence accuracy and speed of different initial values

| | Convergence speed in seconds | Accuracy (%) |
|-------------------------------|------------------------------|--------------|
| $\hat{x}_0 = 90\%, P_0 = 1$ | 2 | 2.5% |
| $\hat{x}_0 = 90\%, P_0 = 0.1$ | 8 | 2.5% |
| $\hat{x}_0 = 90\%, P_0 = 5$ | 10 | 2.5% |
| $\hat{x}_0 = 50\%, P_0 = 1$ | 11 | 2.5% |
| $\hat{x}_0 = 10\%, P_0 = 1$ | 60 | 2.5% |

5 Conclusions

A method which firstly using segmented fitting method to enhance the accuracy of lithium battery model and then utilizing non-linear UKF algorithm to estimate the SOC is proposed. Thanks to the polarization of lithium battery when reaches full capacity or empty, the relationship between open circuit voltage and SOC changes dramatically, this would lead a large model error at the two ends if global fitting was used. By utilizing segmented fitting, the model error is reduced to 107 mV from 224 mV, which provides a more accuracy model for UKF algorithm. Non-linear UKF algorithm comes with the concept of using possibility distribution approximation instead of the method of linearization of non-linear function which used in EKF, which makes it have better estimation accuracy. The experiment results show: SOC estimation by Segmented model based UKF algorithm will exhibits faster time and higher accuracy, the maximum error is around 3%, which saves half time compared to global model estimation; The initial estimated value will not affect the accuracy of SOC estimation, and will slightly affect the convergence speed.

References

- Chen X, Sun D, Chen X (2015) Modeling and state of charge robust estimation for lithium-ion batteries. *Trans China Electrotech Soc* 30(15):141–147
- He Z, Chen D, Pan C (2016) State of charge estimation of power Li-ion batteries using a hybrid estimation algorithm based on UKF. *Electrochim Acta* 211:101–109
- Zhou X, Zhao Y (2014) Study on the SOC estimation of power battery for electric vehicle. *Mech Sci Technol Aerospace Eng* 33(2):263–266
- Kuang L, Deng Q, Li Z (2013) New method for estimating SOC of lithium battery. *Computer Engineering and Applications* 49(6):249–252
- Liu z, Yang J, Zhang Q, Zhang L, Zhang A (2013) Estimation for SOC of lithium battery based on QPSO-BP neural network. *J Electron Measur Instrum* 27(3):224–228

6. Huang Y, Huang H, Zhang H, Weng G (2013) Prediction of lithium-ion battery SOC in EV based on genetic neural network. *J Mech Electr Eng* 30(10):1255–1258
7. Plett G (2004) Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs. Part1: background. *J Power Sources* 134(2):252–261
8. Plett G (2004) Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs. Part2: modeling and identification. *J Power Sources* 134(2):262–276
9. Plett G (2004) Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs. Part3: state and parameter estimation. *J Power Sources* 134(2):277–292
10. Fan B, Tian X, Ma J (2010) EKF-based estimation of lithium-ion traction-battery SOC. *Chin J Power Sources* 34(8):797–799
11. Song Q, Han J (2005) UKF-based active modeling and model-reference adaptive control for mobile robot. *ROBOT* 27(3):226–230
12. Sun F, Tang L (2013) Comparison of cubature Kalman filter and unscented Kalman filter on estimation accuracy. *Decis Control* 28(2):303–308
13. Gao M, He Z, Xu J (2011) Sigma point Kalman filter based SOC estimation for power supply battery. *Trans China Electrotech Soc* 26(11):161–168
14. Zheng J, Ning F (2015) Application of adaptive UKF algorithm on aircraft attitude estimation system. In: International conference on intelligent systems research and engineering, pp 976–979
15. Julier SJ, Uhlmann JK (2004) Unscented filtering and nonlinear estimation. In: Proceedings of the IEEE aerospace and electronic systems 92(3):401–422
16. Chen M, Rincon-Mora GA (2006) An accurate electrical battery model capable of predicting runtime and I-V performance. *IEEE Trans Energy Convers* 21(2):504–511



Simplification of Gauss Spherical Point Cloud Based on K-Mean

Tao Ma, Xiaobin Li^(✉), and Tianyang Yu

School of Shanghai Institute of Technology, Shanghai 201418, China
lixiaobinauto@163.com

Abstract. To simplify the point cloud data of three-dimensional reconstruction of object surface, a point cloud simplification method based on K-means and Gauss sphere is proposed. Without changing the geometric features of the model and retaining the details of the object, the redundant data of the point cloud is removed and the point cloud data is greatly simplified. Firstly, the classical point cloud simplification algorithm is introduced, and its advantages and disadvantages are analyzed. Then, according to the characteristics of various algorithms, PCA method is used to estimate the normal vector features, K-means and Gauss sphere clustering are used to simplify the point cloud. Finally, the validity of the algorithm is validated by the point cloud data of Stanford Bunny model and Dragon model.

Keywords: Point cloud simplification · Detail features · Redundant data

1 Introduction

Three-dimensional reconstruction of object surface has always been an important research topic in SLAM field. The original point cloud data collected by the equipment is huge and redundant, which seriously affects the processing speed of point cloud stitching and registration, and three-dimensional reconstruction of object surface. Therefore, it is very important to simplify the clustering of point clouds. The classical point cloud simplification methods are triangular patch simplification method, direct method and combination of the two methods. However, the construction of triangular patches takes a long time and has a large complexity. The direct point cloud simplification method has no process of triangulation point cloud data processing, so it is more efficient and more widely used. Martin [1] and others used median filtering to propose a uniform mesh method. Select a median point in each grid to represent all points in the grid. This method does not take into account the feature points, and the simplified point cloud is reconstructed to find that the details of the surface are not well preserved. Lee et al. [2] proposed a non-uniform grid method based on the uniform grid method, and then subdivided the uniform grid again until the error ratio of point cloud is smaller than the threshold value and the grid division is stopped. Shi [3] proposed an adaptive simplification algorithm for point clouds. This method is based on K-means clustering principle, which can preserve the boundary of the model well, but is vulnerable to noise interference. Zhou et al. [4] established an octree-based K-Neighborhood model to simplify point clouds based on mean curvature. The

algorithm can protect the boundary data points, and its reliability and accuracy are verified by experiments. Shi [3] and others subdivide the point clouds whose deviation is greater than the threshold value into two categories again, and then iterate again until the normal vector deviation is smaller than the threshold value. Finally, for flat and non-flat regions, the points with maximum deviation of normal vector and clustering center are retained respectively. Document [5] presents a vector method to simplify point clouds.

2 Method

2.1 K-Mean Method for Spatial Clustering

The purpose of K-means spatial clustering is to divide point clouds into several subsets, so that the points of some subsets are similar in space. The most important part of K-means spatial clustering is how to initialize the clustering center. It is the simplest method to choose K value arbitrarily as the initial clustering center, but the stability is not high. JAINAK proposed k-means++ method [6] to improve stability. The most important difference between K-means and k-means++ is the selection of initial clustering centers. K - means++ chooses the point with the greatest distance, which is satisfactory for the case of less point cloud data. But when the point cloud data is large, this method takes a long time. In addition, in order to converge quickly, it is better to let the initial clustering distribute uniformly throughout the model. Therefore, the points in the local neighborhood can be selected as the initial clustering center. K-d tree is a fast neighborhood search method. This method is used to find the initial clustering center.

2.2 Selection of Feature Points

In order to realize K-means clustering of point clouds quickly, the feature points set should be determined first. In the case of extracting feature points from point cloud data, if the change of point normal vector is relatively flat in the local area, the area will be flat. If the change of the normal vector in the point region is large, the region is uneven. According to this discovery, the points with large change of normal vector direction are usually regarded as feature points.

2.3 PCA Estimation Normal Vector and Characteristic

PCA method [7] assumes that the sampling surface of point cloud is smooth, and the normal vector corresponding to the plane corresponds to the normal vector of the point. The plane is fitted by the neighborhood of the point cloud. This method is accurate in estimating the direction of normal vectors in flat regions. However, at high curvature, i.e. K-neighborhood on multiple planes, the method is not smoothed. Figure 1 is a sketch of normal vector estimation of point clouds. Red dots represent feature points and blue dots represent neighborhood points. In Fig. 1(a), K neighborhood is located in two planes. As shown by dotted lines, the normal vectors of fitting plane feature points are smoothed, and the normal vectors of feature points deviate greatly from the normal

vectors of flat areas. In order to estimate the normal vectors of point clouds more accurately, the paper [8] proposes an opposite neighborhood. As shown in Fig. 1(b), only one surface has a neighborhood with characteristic points. Moreover, the normal vectors of the feature points are almost perpendicular on the surface and slightly equal to the normal vectors of the adjacent points.



(a) PCA method for estimating normal vectors (b) Accurate normal vectors of point clouds

Fig. 1. A Normal vector diagram.

PCA method estimates normal vectors: for point sets $P = \{p_1, p_2, p_3, \dots, p_n\}$. n is the total amount of point clouds, the nearest K_1 neighborhood of point p_i is represented by $N_b(p_i)$, the center of gravity of $N_b(p_i)$ is represented by u , and the covariance matrix of point p_i is defined as C .

$$u = \frac{1}{k_1} \sum_{p_i \in N_b} p_i \quad (1)$$

$$C = \frac{1}{k_1} \sum_{i=1}^{k_1} (p_i - u)(p_i - u)^T = \frac{1}{k_1} \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}^T \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}, p_i \in P \quad (2)$$

2.4 Gauss Sphere Clustering

The non-flat clustering of point clouds is subdivided into several parts by Gauss mapping, which ensures that the flat points and feature points belong to different clustering. Suppose that a front side of a surface is represented by $T = T(U, V) \subset R^3$, and the normal vector direction of the surface is the same. The beginning of the unit normal vector on the point $Q < T$ surface moves to the center of the unit sphere $T^2 = \{(x, y, z) | R^3 | x^2 + y^2 + z^2 = 1\}$, and the point on the surface corresponds to the point on the sphere. This corresponds to the Gaussian mapping $G : T \rightarrow T^2$ of surface T . However, there are always errors in the direction of the normal vector estimated by this method. Gauss map is a clustering of point clouds densely distributed on the Gauss sphere. After K-means clustering, the cluster centers of flat regions can represent the whole cluster, but if the cluster centers contain feature points, then replacing the whole

cluster with the cluster centers will produce surface volume shrinkage. Therefore, clustering needs to be divided into several flat regions and feature points. According to PCA method, clustering centers on Gauss spheres are mapped to multiple clusters in flat regions, and feature points are mapped to one or more clusters independently. Since the number of sub-clusters can not be specified in advance, it is necessary to use parameter clustering algorithm to automatically segment the data on the Gauss graph. In this paper, adaptive mean shift algorithm is used to cluster Gauss graphs.

Let the point clouds of a K-means clustering method which are out of direction consistent be expressed as $Q = (q_i, 1, 2, 3, \dots, n).G(Q) = [x(q), q \in Q] = \{x_i, 1, 2, 3, \dots, m\}$ denotes the Gaussian graph of point set Q, the normal vector of point q is $x(q)$, the normal vector of point i in point set Q is x_i and the number of point clouds in point set Q is m . Then the mean drift vector of a point x can be expressed as:

$$m_{g,h}(x) = \left[\sum_{i=1}^m \frac{1}{h^{d+2}} x_i g\left(\left\|\frac{x - x_i^2}{h}\right\|\right) / \sum_{i=1}^m \frac{1}{h^{d+2}} g\left(\left\|\frac{x - x_i^2}{h}\right\|\right) \right] - x \quad (3)$$

Where: the kernel function is $g(x)$, the bandwidth is h , and the spatial dimension is d . d is determined to be 3. In document [9], the bandwidth is represented by $h_i = |x_i - x_{ij}|$, x_{ij} are the neighborhood points of x_i . In this paper, the maximum value is bandwidth, the k_2 neighborhood of point x_i is $N(x_i)$, and the Epanechnikov kernel [9] is used.

3 Algorithmic Analysis

3.1 Simplified Results Analysis

The point cloud data in this paper can be downloaded from the website provided in document [5]. Among them, Bunny model contains more connected surfaces with high curvature, Dragon contains more sharp surfaces, and the sharp feature points are Dragon model dorsal fin, tail, forelimb, etc. As can be seen from Table 1 the method in this paper has a high simplification rate. From the simplified process diagram, it can be seen that the simplified model has clear features, the feature points are well preserved, and the point cloud data are retained at high curvature and sharp surface (Figs. 2 and 3 and Table 1).

3.2 Comparative Analysis of Effectiveness

In order to analyze the effectiveness of the proposed method, the comparison between the proposed method and the literature [5] method is made. Two groups of completely different models are selected for analysis: one is Bunny model with high curvature connected surface, which has uniform point cloud distribution and smooth transition edge of surface intersection; the other is Dragon model with sharp feature points, which has sharp transition edge. The effectiveness of the two methods is compared by observing whether the feature points of the simplified model are still clearly visible and calculating the simplification rates of different models. The simplified results of the two methods are shown in Figs. 4 and 5 respectively. It can be seen that both the methods

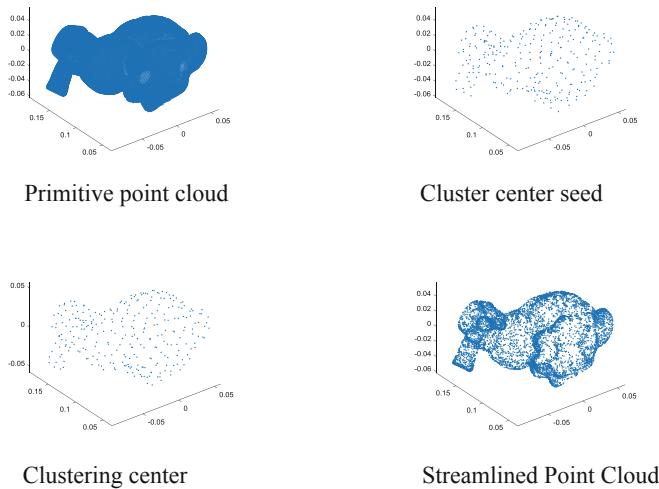


Fig. 2. Simplification of Bunny point clouds.

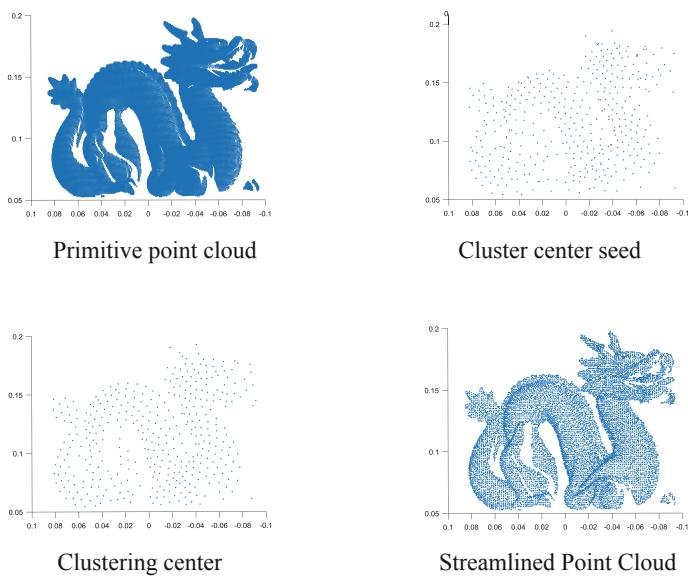


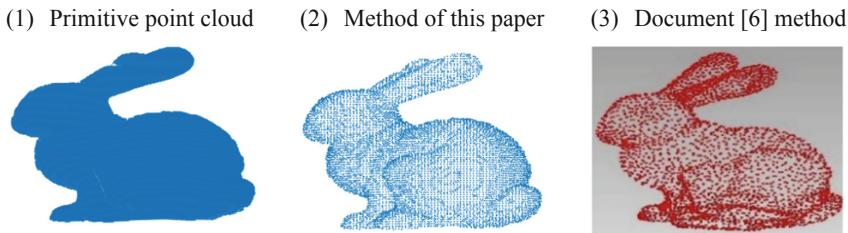
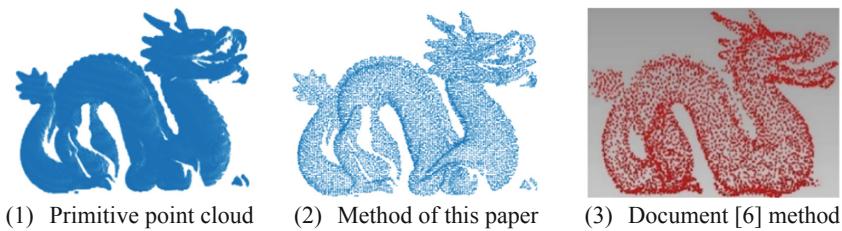
Fig. 3. A Dragon point cloud simplification process.

presented in this paper and in reference [5] can retain more points on high curvature surfaces and flat surfaces. However, the feature points on sharp surfaces are poorly preserved by the method in reference [5], and the simplified point cloud loses the detailed features of the surfaces. The method of this paper can preserve the feature

Table 1. The simplification rate of model point cloud of this paper's method

| Model | Number of original point clouds | Number of simplified point clouds | Simplification rate |
|--------|---------------------------------|-----------------------------------|---------------------|
| Bunny | 34835 | 5976 | 82.8441% |
| Dragon | 49990 | 8670 | 82.6558% |

points of objects well on high curvature surface and sharp surface (dragon horn, dragon tail, dragon forelimb, dragon dorsal fin). It can remove the redundant data of point cloud and effectively retain the detail features of point cloud. The method of this paper is more effective.

**Fig. 4.** Simplified comparison of Bunny model.**Fig. 5.** Simplified comparison of Dragon model.

4 Conclusion

In this paper, the method of preserving edge point cloud simplification based on normal vector in point cloud simplification algorithm is improved, and a comprehensive algorithm is studied. Through actual verification and comparison, it is concluded that the proposed method can preserve the feature points of the object, remove the redundant data of the point cloud and retain the detailed features of the object with a high simplification rate, no matter in flat area, high curvature surface and sharp surface, without changing the features of the object, so it has more advantages and universality.

Acknowledgments. This paper has been supported and guided by my tutor Professor Li Xiaobin. Here, thanks to my tutor. The rigorous scientific research spirit, profound knowledge and modest attitude of my tutor are the realms of life that I pursue. In addition, I would like to express my heartfelt thanks to all those who have helped me. Thank you!

References

1. Martin R, Stroud I, Marshall A (1996) Data reductions for reversing engineering. RECCAD. Deliverable document 1 COPERUNICUS project. No. 1068. Computer and Automation Institute of Hungarian Academy of Science
2. Lee K, Woo H, Suk T (2001) Point data reductions using 3D grids. *Int J Adv Manuf Technol* 18(3):201–210
3. Shi BQ, Liang J, Liu Q (2011) Adaptive simplification of point cloud use-means clustering. *Comput Aided Des* 43(8):910–922
4. Zhou Y, Zhang W, Du F et al (2010) Curvature simplification algorithm for scattered point cloud data. *J Beijing Univ Technol* 30(7):785–789
5. Han H, Han X, Sun F, Huang C (2015) Point cloud simplification with preserved edge based on normal vector. *Optik - Int J Light Elect* (19):2157–2162
6. Jainak (2010) Data clustering: 50 years beyond K-means. *Pattern Recogn Lett* 31(8):651–666
7. Hoppeh H, Deroiset T, Duchampt T et al (1992) Surface reconstruction from unorganized points. ACM
8. Nurunnabia A, Westg G, Beltond D (2015) Outlier detection and robust normal-curvature estimation in mobile laser scanning 3D point Cloud data. *Pattern Recogn* 48(4):1404–1419
9. Anand S, Mittal S, Tuzelo O et al (2013) Semi-supervised kernel mean shift clustering IEEE Trans Pattern Anal Mach Intell 36(6):1201–1215



Trajectory Planning for Seven-DOF Robotic Arm Based on Seventh Degree Polynomial

Shuang Fang, Xianghua Ma^(✉), Jiarui Qu, Sijing Zhang, Na Lu,
and Xin Zhao

School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai 201418, China
xhuam@sit.edu.cn

Abstract. Trajectory planning can ensure the robot arm move smoothly and quickly to the target position. The difficulty is to use inverse kinematics to convert the path point into the joint angle and fit a smooth function to each joint. The paper deals with research on trajectory planning for seven-DOF robotic arm. In the joint space, based on the kinematics analysis, the joint space trajectory planning is realized by the cubic polynomial and the seventh degree polynomial, and the simulation is realized on the MATLAB platform. The simulation results show that the robot link parameters are designed reasonably, the kinematics model is correctly established and the seventh degree polynomial interpolation method effectively solves the problem of acceleration discontinuity, and obtains a continuous smooth trajectory curve of each joint, which visually verifies the effect of trajectory planning and provides an efficient and feasible trajectory planning method.

Keywords: Trajectory planning · Seven-DOF robotic arm · Polynomial

1 Introduction

Manipulator trajectory planning means calculating the trajectory according to the requirements of the expected operation of the task and studying the generate method of displacement, angular velocity and angular acceleration in movement [1]. Reasonable trajectory planning can make the movement more stable, reduce wear of the mechanism, and improve the stability and reliability of the robot arm effectively. So for grasping motion of the robot, to make it fast, smooth and accurate movement to the target position, trajectory planning is very important [2].

The trajectory planning is generally carried out in the operating space and in the joint space. As Gasparetto pointed out in Ref. [3], the trajectories were planned in the joint space in most cases. Moreover, trajectory planning in the joint space would allow to avoid the problems arising with kinematic singularities and manipulator redundancy [4]. A manipulator with more degrees of freedom than necessary is called a kinematic redundant manipulator [5]. Redundant manipulators have added flexibility, which can be useful in avoiding obstacles or reaching inaccessible locations [6]. The use of kinematic redundant manipulators is expected to increase dramatically in the future because of their ability to avoid the internal singularity configurations and obstacles,

high manipulability, and to optimize dynamic performance [7, 8]. So it is important to develop sophisticated motion planning algorithms to be able to utilize the advantages of redundant manipulators.

Many researchers focus on the study of trajectory planning. Ho et al. [9] used cubic polynomial and four-order polynomials for trajectory planning, but the cubic polynomial did not obtain a smooth acceleration curve. Gasparetto [4] used five power B-spline function for planning. In order to acquire a smooth trajectory, the second derivative of the speed is introduced into the objective function, and he added the execution time in the objective function to constraint angular velocity, angular acceleration, and other information. In many robotics motion planning problems, the robot need to pass an intermediate point(s) between the starting pose and the ending pose, such as avoiding obstacles. Angeles [10] derives a 4-5-6-7 seventh-order polynomial to fit two via points and ensure finite joint jerk at the start and end of motion. In order to deal with via points that the robot does not need to stop at the path point, Craig [11] suggests matching two third-order polynomials. Although there is a wide range of different trajectory designing algorithms, currently there exists no general method, which outperforms all others for all problem instances. In fact, each method has various advantages and drawbacks that make it best suited for specific types of problems [12].

In order to meet the smooth, continuous trajectory requirements of the seven-DOF manipulator during operation, the article adopted the method of seventh degree polynomial. The specific method is when the initial position and posture of the robot arm are known, the seven joint angles of the inverse kinematics of a set of manipulators are used to represent the ideal position of the robot arm at the target. The end trajectory of the robot arm can be planned by matching the proper conditions with the end conditions and boundary conditions of the robot arm [13].

This paper is organized as follows. In Sect. 2, the robot arm kinematics is analyzed in a coordinate system established by the D-H method, and the forward kinematics equations are deduced. In Sect. 3, two methods of polynomial trajectory planning are introduced. In Sect. 4, a geometric model of seven degrees of freedom robot is established first, in the joint space, cubic polynomial and the seventh degree polynomial functions are used to carry out arm trajectory planning, the simulation and experimental results are presented with relevant discussion. Finally, Sect. 5 outlines the main conclusions.

2 Kinematics Analysis

2.1 Positive Kinematics Analysis

Robot manipulator has seven degrees of freedom. For the solution of kinematics equation [14, 15], this paper uses D-H [16] method to establish a manipulator link coordinate system. Between the two links, the position and orientation are described by a 4×4 uniform transformation matrix in the D-H method. As shown in formula (1).

$$A_i = \text{Rot}(z, \theta_i) \text{Trans}(0, 0, d_i) \text{Trans}(a_{i-1}, 0, 0) \text{Rot}(x, \alpha_{i-1})$$

$$= \begin{bmatrix} c\theta_i & -s\theta_i c\alpha_{i-1} & s\theta_i s\alpha_{i-1} & a_{i-1} c\theta_i \\ s\theta_i & c\theta_i c\alpha_{i-1} & -c\theta_i s\alpha_{i-1} & a_{i-1} s\theta_i \\ 0 & s\alpha_{i-1} & c\alpha_{i-1} & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

The transformation matrix of the end effector coordinate system relative to the base coordinate system can be obtained by multiply the coordinate transformation matrix of each adjacent links:

$$T = A_1 A_2 A_3 A_4 A_5 A_6 A_7 \quad (2)$$

Now, putting the D-H parameters and joint variables of manipulator substitution of the Table 1 in formula (2). Therefore, the initial pose of the end effector is:

$$T = \begin{bmatrix} 1 & 0 & 0 & -0.5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0.14 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

The configuration of the robot is shown in Fig. 1.

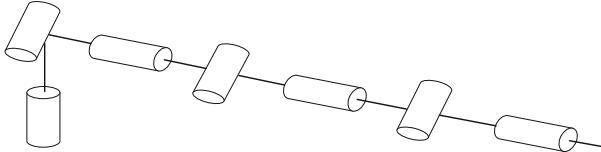


Fig. 1. The configuration of manipulator.

D-H parameters and associated values for manipulator have given in Table 1.

Table 1. The link parameters of robot.

| Link i | a_i/m | $\alpha_i (\circ)$ | d_i/m | $\theta_i (\circ)$ |
|----------|----------------|--------------------|----------------|--------------------|
| 1 | 0 | 0 | 0.14 | 0 |
| 2 | 0 | 90 | 0 | 0 |
| 3 | 0 | 90 | 0 | 0 |
| 4 | -0.25 | 90 | 0 | 0 |
| 5 | -0.25 | 90 | 0 | 0 |
| 6 | 0 | 90 | 0 | 0 |
| 7 | 0 | 90 | 0 | 0 |

2.2 Inverse Kinematics Analysis

The inverse kinematics problem is more difficult than the direct kinematics problem because no single explicit systematic procedure analogous to the D-H algorithm is available. As a result, each robot or a generically similar class of robots has to be treated [17]. In the manipulator control, it is often known to reach the target position and orientation, and the joint variables are used to drive the motor of each joint. Therefore, it is the basis of the manipulator motion planning and trajectory control. The concrete method can be found in [15].

3 Kinematics Analysis

3.1 Cubic Polynomial Interpolation

Unstable movement will not only aggravate joint spot attrition, moreover cause vibration and impact of various arm [18]. Polynomials are natural choices for providing smooth, continuous motion, with some level of continuous derivatives [19]. On the process of joint space interpolation, the corresponding posture of the arm joint angles to the starting point and the termination point can be obtained by inverse kinematics. So the trajectory description can be represented for smoothing interpolation function $\theta(t)$ to the termination and starting joint angle. There are four unknowns in the cubic polynomial equation. We define the initial time of the robotic arm joint movement is t_0 , the start point of the joint angle is θ_0 , the termination time is t_f , the joint angle is θ_f . The description of the trajectory can be represented by a smooth interpolation function of the start angle and end angle. According to these four known information, the equation can be solved. The functional equation for joint angle and velocity is as follows:

$$\begin{cases} \theta(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \\ \dot{\theta}(t) = a_1 + 2a_2 t + 3a_3 t^2 \end{cases} \quad (4)$$

Then, in order to achieve a smooth motion of a single joint, the trajectory function $\theta(t)$ satisfies at least four constraints:

$$\begin{cases} \theta(t_0) = \theta_0 \\ \theta(t_f) = \theta_f \\ \dot{\theta}(t_0) = 0 \\ \dot{\theta}(t_f) = 0 \end{cases} \quad (5)$$

Substituting the known initial and end conditions into Eq. (4) gives the coefficients:

$$\begin{cases} a_0 = \theta_0 \\ a_1 = 0 \\ a_2 = \frac{3(\theta_f - \theta_0)}{t_f^2} \\ a_3 = \frac{-2(\theta_f - \theta_0)}{t_f^3} \end{cases} \quad (6)$$

3.2 Seventh Degree Polynomial Interpolation

In the case where the trajectory is stricter and the constraints are increased, the cubic polynomial interpolation cannot satisfy the requirement, and the high-order polynomial is used for interpolation. There are eight unknowns in the seventh degree polynomial. Therefore, in addition to the known angle and velocity of a joint at the initial motion time t_0 and the end time t_f , the acceleration values of the starting point and the end point are also set. The functional equation of the joint angle, velocity and acceleration is as follows:

$$\begin{cases} \theta(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5 + a_6 t^6 + a_7 t^7 \\ \dot{\theta}(t) = a_1 + 2a_2 t + 3a_3 t^2 + 4a_4 t^3 + 5a_5 t^4 + 6a_6 t^5 + 7a_7 t^6 \\ \ddot{\theta}(t) = 2a_2 + 6a_3 t + 12a_4 t^2 + 20a_5 t^3 + 30a_6 t^4 + 42a_7 t^5 \\ \dddot{\theta}(t) = 6a_3 + 24a_4 t + 60a_5 t^2 + 120a_6 t^3 + 210a_7 t^4 \end{cases} \quad (7)$$

Six known conditions are:

$$\begin{cases} \theta(t_0) = \dot{\theta}(t_0) = \ddot{\theta}(t_f) = \dddot{\theta}(t_0) = \ddot{\theta}(t_f) = \dddot{\theta}(t_f) = 0 \\ \theta(t_f) = \theta_f \end{cases} \quad (8)$$

Substituting known conditions into Eq. (7), the coefficients are as follows:

$$\begin{cases} a_0 = \theta_0 \\ a_1 = v_0 \\ a_2 = \frac{a_0}{2} \\ a_3 = \frac{j_0}{6} \\ a_4 = \frac{210\theta_f - t_f [(30a_0 - 15a_1)t_f + (4j_0 + j_1)t_f^2 + 120v_0 + 90v_1]}{t_f^4} \\ a_5 = \frac{-1680\theta_f + t_f [(20a_0 - 14a_1)t_f + (2j_0 + j_1)t_f^2 + 90v_0 + 78v_1]}{2t_f^5} \\ a_6 = \frac{4200\theta_f - t_f [(20a_0 - 39a_1)t_f + (4j_0 + 3j_1)t_f^2 + 216v_0 + 204v_1]}{6t_f^6} \\ a_7 = \frac{-1200\theta_f + t_f [30(12a_0 - 12a_1)t_f + (j_0 + j_1)t_f^2 + 60v_0 + 60v_1]}{6t_f^7} \end{cases} \quad (9)$$

4 Prototype and Test Results

4.1 Kinematic Modeling

Set starting point $p_0 = [-0.5, 0, 0.14]$, the target point $p_1 = [0.25, -0.25, 0.14]$, in the working space of the known robotic arm. The joint angle values of the joints at two points are obtained by the ikine function in MATLAB Robotics Toolbox as shown in

Table 2. Using the link function to construct the 3D model, its position and orientation at different points are shown in Fig. 2.

Table 2. The joint angle values of the joints at two points.

| Joint angle | Start point | Target point |
|-------------|-------------|--------------|
| θ_1 | 0 | 1.5708 |
| θ_2 | 0 | 1.5708 |
| θ_3 | 0 | 1.5708 |
| θ_4 | 0 | 1.5708 |
| θ_5 | 0 | 1.5708 |
| θ_6 | 0 | 1.5708 |
| θ_7 | 0 | -3.1416 |

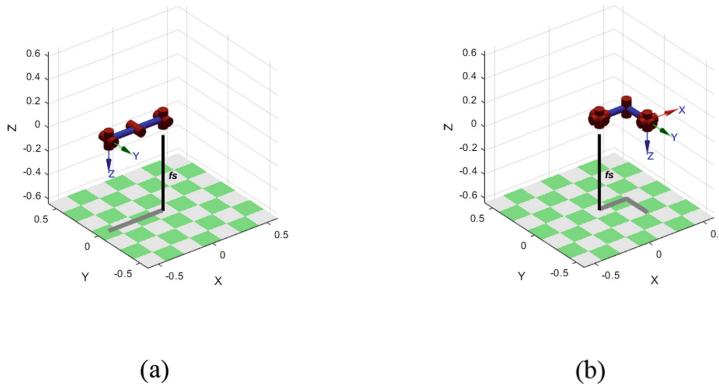


Fig. 2. Start point (a) and target point (b).

4.2 Simulation Results

The cubic polynomial interpolation method and the fifth-order polynomial interpolation method are used to realize the motion planning of the robot from the starting point to the target point. At the same time, the angle, velocity and acceleration curves are obtained. The results are shown in Figs. 3 and 4.

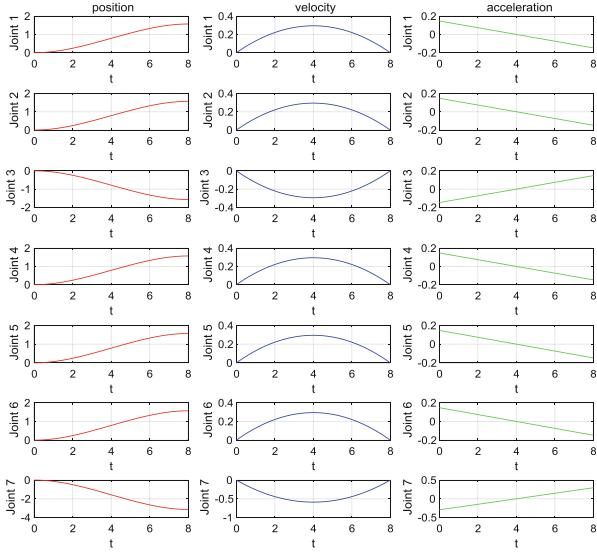


Fig. 3. Cubic polynomial 1–7 joint track curve.

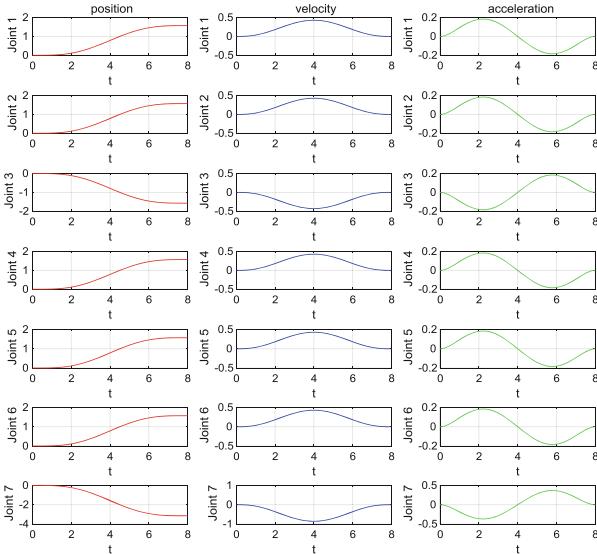


Fig. 4. Cubic polynomial 1–7 joint track curve.

As shown in Fig. 3, the angular displacement function, angular velocity function and angular acceleration function of each joint are continuous, and the motion of the robot in this motion is smooth. Figure 4 shows that the joint curve obtained by the seventh degree polynomial is smoother than the joint curve obtained by the cubic polynomial, and the joint acceleration curve is not straight but curved, and has a better effect.

5 Analyses of Experimental Results

In this paper, the kinematics analysis of the seven-DOF robotic arm is carried out. Based on the kinematics analysis, the joint space trajectory planning is realized by the cubic polynomial and the seventh degree polynomial, and the simulation is realized on the Matlab platform. The simulation results show that the robot link parameters are designed reasonably and the kinematics model is correctly established. An important contribution of this paper is to find out that the seventh degree polynomial method has an important influence on the trajectory planning of the seven-DOF robotic arm through comparative analysis. We obtain a smooth continuous trajectory planning curve through the seventh degree polynomial, which shows that the method is feasible in the similar seven-DOF robotic arm trajectory planning. The method can make the robotic arm run smoothly, reduce the occurrence of jitter and vibration, and improve the stability of the system. It provides a reference for the future trajectory planning of the humanoid robotic arm and hand gripping.

Acknowledgments. This work was supported by the affiliate program LM201768.

References

1. Lu JY, Xiong XS, Zhu D (2013) Five degrees of freedom picking manipulator joint space trajectory planning. *Light Ind Mach* 31:11–14
2. Ma SH, Yang L, Liu ZG, Wang JK (2014) Trajectory planning of 6-DOF manipulator based on combination function method. In: The 26th Chinese Control and Decision Conference (2014 CCDC)
3. Gasparesco A, Boscaroli P, Lanzutti A, Vidoni R (2015) Path planning and trajectory planning algorithms: a general overview. *Mech Mach Sci* 29:3–27
4. Gasparesco A, Zanotto V (2007) A new method for smooth trajectory planning of robot manipulators. *Mech Mach Theory* 42(4):455–471
5. Agarwal V (2011) Trajectory planning of redundant manipulator using fuzzy clustering method. *Int J Adv Manuf Technol* 61:5–8
6. Wunderlich J, Boncelet C, Wunderlich J, Boncelet C (1996) Local optimization of redundant manipulator kinematics within constrained workspaces. In: Proceedings of the IEEE international conference on robotics and automation, Minneapolis, MN, vol 1, pp 127–132
7. Dubey R, Luh JYS (1988) Redundant robot control using task based performance measures. *J Robot Syst* 5(5):409–432
8. Ding H, Chan SP (1996) A real-time planning algorithm for obstacle avoidance of redundant robots. *J Intell Robot Syst* 16(3):229–243
9. Ho CY, Cook CC (1982) The application of spline functions to trajectory generation for computer-controlled manipulators. *Digit Syst Ind Autom* 1(4):325–333
10. Angeles J (2007) Fundamentals of robotic mechanical systems, 3rd edn. Springer, New York
11. Craig JJ (2005) Introduction to robotics: mechanics and control, 3rd edn. Pearson Prentice Hall, Upper Saddle River
12. Morales M, Tapia L, Pearce R, Rodriguez S, Amato NM (2005) A machine learning approach for feature-sensitive motion planning. In: Algorithmic foundations robotics VI, pp 361–376

13. Zhu HJ (2017) Motion trajectory planning and simulation of 6 DOF manipulator arm robot. *Acad J Manuf Eng* 15(3)
14. Jiang XS (1994) Introduction to robotics. Liaoning Science and Technology Press, Shenyang
15. Cai ZX (2000) The robotics. Tsinghua University Press, Beijing
16. Craig JJ (1986) Introduction to robotics: mechanics and control. Addison-Wesley Publishing Company, Boston
17. Aljawi AAN, Balamesh AS, Almatrafi TD, Akyurt M (2002) Symbolic modeling of robot manipulators. In: The 6th Saudi engineering conference, vol 4, pp 597–611
18. Zhang ZY, He DJ, Tang JL, Meng LS (2014) Picking robot arm trajectory planning method. *Sensors Transducers* 162(1):11–20
19. Yan XS, Wu QH, Liu HM (2012) An improved robot path planning algorithm based on genetic algorithm. *Telkomnika* 10(8):1948–1955



A Hybrid Approach for Solving the Flexible Job-Shop Scheduling Problem with Considering Transporting Time and Deteriorating Effect

Zhuang Huang and Jianjun Yang^(✉)

School of Mechanical Engineering and Automation, Beihang University,
Beijing 100191, China
jjyang@buaa.edu.cn

Abstract. To study the flexible job-shop problem(FJSP) with considering the transporting time and deteriorating effect, a novel model with the objective of minimizing the makespan is developed by disjunction graph and a linear deteriorating function whereas a hybrid algorithm (IGAVNS) is proposed combining immune genetic algorithm (IGA) with variable neighborhood search strategy (VNS). First, the operation-based encoding method is used to encode chromosome individual, solving the operation sequencing and machine selection. Then, elite individual is searched by the proposed algorithm with five neighborhood structures. Simulations were carried out by different algorithms on the cases of different sizes. The results were compared and analyzed, verifying the effectiveness of the proposed algorithm.

Keywords: Transporting time · Deteriorating effect · Flexible job-shop · Immune genetic algorithm · Variable neighborhood search

1 Introduction

In actual production, an operation can be usually processed on multiple machines. It is necessary to select suitable machines for the processing according to the production situation on the spot. Thus, a flexible job shop scheduling problem (FJSP) is developed on the basis of the traditional job-shop scheduling problem (JSP).

The classical FJSP is divided into two sub-problems of operation sequencing and machine selection [1]. The makespan is generally used as the objective of optimization.

FJSP has been proved to be a NP-Hard problem. The main solutions are swarm intelligence optimization algorithm, such as genetic algorithm (GA) [2], ant colony algorithm (ACO) [3], particle swarm optimization (PSO) [4]. However, the transporting time exists objectively and cannot be ignored in actual production. Unfortunately, it was neglected by most of the studies on FJSP. As we all know, the theoretical research will be lack of guidance if too many objective factors are ignored in assumptions. Moreover, the long transporting time will directly affect the quality of products in some specific production areas.

At present, transporting time was considered in just few studies on FJSP. Zhang et al. [5] proposed an improved genetic algorithm to solve the mathematical model of FJSP problem with transporting time. Ji et al. [6] analyzed the problem using a batching and scheduling model involving both batch supply and batch delivery to minimize the sum of the total weighted inventory cost and transport cost. Hurink et al. [7] assumed that transporting time exists between material handling systems, took minimizing completion time as the optimization objective, and used tabu search algorithm to find local optimal solution.

The effects of job characteristics, equipment wear and operator fatigue are considered by Gupta et al. [8], who assumed that the processing time of the job is a linear increasing function of its starting time. They proposed scheduling model with deteriorating job characteristics. This deteriorating scheduling model makes up for the lack of practicability of the traditional scheduling model, and has a wide application background. However, the current research on deteriorating scheduling problem is mostly based on a single machine [9–11] and assembly line [12, 13], less on job shop [14]. Mosheiov [15] stated that the JSP with minimizing the latest completion time is NP-complete problem when the actual processing time is linear with its start time. The FJSP with exponential deterioration of job processing time was considered in [16]. Existing research on JSP with deteriorating characteristics assumed that the deteriorating factors of jobs on all devices are the same. However, in the actual production environment, the deteriorating factor of the operation is often affected by equipment wear, operator fatigue and the complexity of the process. Therefore, different deteriorating factors is required to use for different operations of each job.

Besides some simple problems can be solved by operational research method, genetic algorithm (GA) is an effective method to solve FJSP considering transporting time. In practical, it is more practical to obtain quickly the near optimal solution. Nevertheless, the standard genetic algorithm (SGA) has the characteristics of premature convergence. Therefore, we focus on solving the FJSP with considering the transporting time and deteriorating effect in this paper. To solve FJSPs, a novel model with the objective of minimizing the makespan is developed by disjunction graph whereas a hybrid algorithm (IGAVNS) is proposed combining immune genetic algorithm (IGA) with variable neighborhood search strategy (VNS).

2 The Model of FJSP Considering Transporting Time and Deteriorating Effect

There are n independent jobs $J = \{J_1, J_2, \dots, J_n\}$ and m machines $M = M_1, M_2, \dots, M_m\}$, each job J_i ($i = 1, 2, \dots, n$) contains n_i operations $\{O_{i1}, O_{i2}, \dots, O_{in_i}\}$. The operation O_{ij} of the job J_i is processed on the machine k of the optional machine set M_{ik} ($k \in \{1, 2, \dots, m\}$, $M_{ik} \subset M$).

In order to simplify the scheduling problem, besides the common assumptions, there are also the following assumptions for scheduling considering transporting time:

1. The sequence of all operations is not allowed to change. After each operation has been completed, it can be transported to the next equipment and start the processing of the next operation.
2. The first operation does not need to be transported before processing, and the last operation does not need to be transported after processing.
3. Starting from the first operation, once the job has been processed, it will be transported immediately to the next operation processing equipment, regardless of the batch transportation.
4. Processing time varies with the choice of equipment while transporting time varies with the choice of equipment before and after the operation.

The disjunctive graph model [17] is convenient to express the sequence constraints and equipment resource constraints. Therefore, in order to express the transporting time constraints, a time set T is added on the basis of the classical disjunctive graph, which is expressed with TDG.

$$\text{TDG} = (N, A, E, T) \quad (1)$$

where, N represents the set of nodes in disjunction graph, including virtual start and end, A is the set of connecting arcs, representing the set of operations of each job. E is the set of disjunction arcs, representing the set of operations processed on the same machine, T represents the set of time in the process of processing, and here is the transporting time of operations. The transporting time of optional processing equipment in two adjacent operations of the same job is expressed by time matrix.

A flexible production scheduling model with 3 jobs and 3 operations is taken as an example, whose the disjunctive graph model is given in Fig. 1. There are two optional machines in the process of O₂₁. According to the hypothesis, neither of them needs transporting time before processing. Therefore, two optional operations in the process of O₂₁ are represented by (0, 0). There are three optional machines in process O₂₂, and together with the two machines in process O₂₁, six possible machine selection schemes are formed. The corresponding transporting time of the six schemes is expressed by

2×3 matrix $\begin{bmatrix} 4 & 2 & 1 \\ 5 & 3 & 4 \end{bmatrix}$. Similarly, there are two optional machines in the process of O₂₃, which form a 3×2 matrix $\begin{bmatrix} 3 & 2 \\ 1 & 5 \\ 4 & 3 \end{bmatrix}$ with O₂₂. By analogy, the transporting

time constraints between all two adjacent nodes can be expressed by the transporting time matrix combined with the disjunctive graph model.

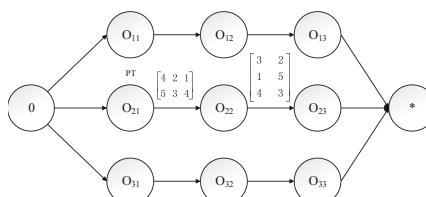


Fig. 1. Disjunctive graph scheduling model with considering transporting time.

In [16], the actual processing time is defined as an exponential function of the total processing time on the machine, but there is still a problem. When the time exponent increases with the accumulation of time, the actual processing time will be a very large. Therefore, a time formula is proposed in the paper to calculate actual working hours considering deteriorating effect, shown as (2).

$$P'_{ijk} = P_{ijk} + \beta \sum P_{ijk} \quad (2)$$

where P'_{ijk} represents the actual processing time of the j th operation of the i th job on machine k , P_{ijk} denotes the standard processing time, and β is the non-negative deteriorating factor.

This formula indicates that the actual working hours of the current operation on this machine are linear functions of deteriorating factor β with the increase of the processing time.

3 Algorithm Design

3.1 Global Flow Chart of the Proposed Algorithm

The global flow chart of IGA VNS is shown as the Fig. 2.

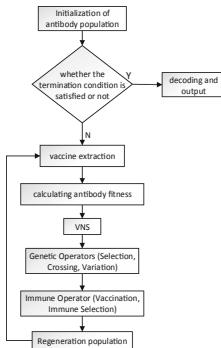


Fig. 2. Flow chart of IGA based on variable neighborhood search

Algorithm consists of six steps as follows:

Step 1: Antibody coding and initialization of antibody population. Integer is used for individual coding. Each chromosome represents the processing sequence of all jobs. The first half of the chromosome represents the processing sequence of all the jobs on the machine, and the second half represents the processing machine serial number of each working procedure of the job.

Step 2: Vaccine extraction. Vaccine is a kind of basic characteristic information extracted from the experience knowledge of people about the problem to be solved. In practical calculation, it is usually a few genes of the best individual. In this paper, through the analyses about the application of FJSP in practical production, the principle of the shortest processing time (SPT) [18], which is often used in the priority rule of JSP, is used as the vaccine of our algorithm.

Step 3: Antibody fitness calculation.

Step 4: Variable neighborhood search

Individuals with the smallest antibody fitness are selected as the initial solution of variable neighborhood search. The process is detailed in Sect. 3.2.

Step 5: The design of genetic operators

- (1) selection. Roulette gambling is used to select chromosomes with better fitness.
- (2) crossover. The method of single point crossover (random selection of a point, with this point as the boundary, and exchange variables with each other) is adopted here.
- (3) mutation. The mutation operator first randomly selects the mutation individuals from the population, then chooses randomly two mutation positions. Finally, the processing operations of the two positions in the individuals and the corresponding processing machine serial numbers are exchanged.

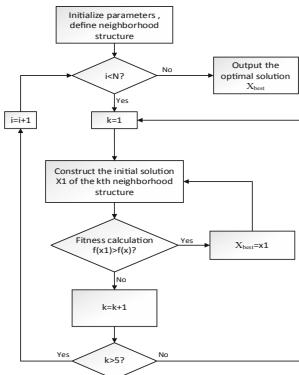
Step 6: The design of immune operator. It is divided into two parts: vaccination and immune selection.

- (1) The process of vaccination is to randomly select individuals in the population with a certain probability and compare them with the best individuals, so that the new obtained individuals have higher adaptability with a larger probability. For each row in the processing sequence of jobs on each machine, the sequence is sorted according to the principle of SPT, and the chromosome is restored according to the processing constraints. This adjustment process is the injection process of immune vaccine.
- (2) The process of immune selection is to compare the fitness values of individuals before and after vaccination. If the fitness values become smaller, the individuals before inoculation will remain unchanged. If the fitness values become larger, the individuals before inoculation will be replaced by those after inoculation.

3.2 Variable Neighborhood Search

The design of neighborhood structure is the key of VNS algorithm. Five neighborhood structures are adopted in this paper shown as Fig. 3: (1) N_1 insertion. (2) N_2 reverse. (3) N_3 disruption. (4) N_4 exchange. (5) N_5 Minimum processing time.

Step 1: determine the type of neighborhood structure N_k ($k = 1, 2, 3, 4, 5$), initial the number of loops N and initial solution x , where x is the result of immune genetic algorithm (IGA), $X_{best} = x$.

**Fig. 3.** Flow chart of VNS

Step 2: if $i > N$, then the best individual (X_{best}) is output, otherwise step 3 is executed.

Step 3: construct the new solution x_1 of the k th neighborhood structure, compare the fitness values $f(x_1)$ and $f(X_{best})$. If $f(x_1) > f(X_{best})$, the new solution will replace the initial solution ($X_{best} = x_1$) and the search will continue in the neighborhood, otherwise, $k = k + 1$.

Step 4: if $k > 5$, $i = i+1$, and the next loop starts, otherwise, the search enters the next neighborhood structure (back to step 3).

4 Simulation Results and Analysis

To verify the performance of the proposed algorithm, 10 jobs 7 machines are taken as an example. The value of transporting time matrix is generated randomly between the maximum and minimum of processing time. The algorithm parameters of IGAVNS are generated after running some pilot tests and showed in Table 1.

Table 1. Algorithm basic parameters of IGAVNS

| Algorithm basic parameters of IGAVNS | |
|--------------------------------------|------|
| Individual number | 1000 |
| Gens of GA | 200 |
| Crossover rate | 0.9 |
| Aberration rate | 0.8 |
| Counts of VNS | 20 |
| Number of N_k | 5 |
| Immune probability | 0.95 |

The experimental results of different algorithms under different factors are shown in Table 2.

Table 2. Results (makespan) of different algorithms under different factors

| Factor | Method | | |
|--------------|--------|-----|--------|
| | GA | IGA | IGAVNS |
| PT | 712 | 606 | 603 |
| PT & TT | 957 | 691 | 681 |
| PT & TT & DT | 1072 | 842 | 839 |

PT: only consider processing time, PT & TT: consider processing time and transporting time, PT & TT & DT: consider processing time, transporting time and deteriorating effect.

It can be seen from Table 2, IGAVNS results are superior to IGA and GA under the same factor. The makespan is obviously larger for PT & TT & DT than for PT & TT. Figures 4, 5 and 6 present the Gantt charts of the optimal solution by IGAVNS algorithm under different factors. The transporting time in Fig. 5 and Fig. 6 is represented by the corresponding bar above the operation.

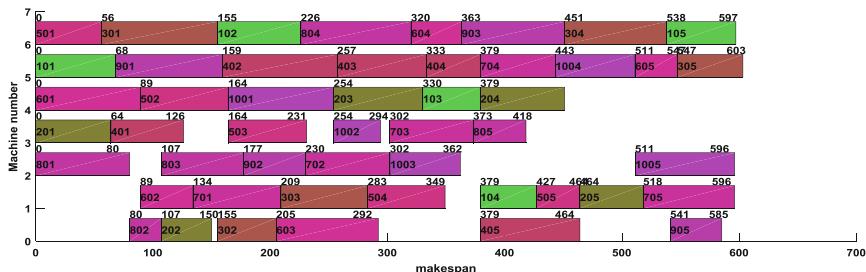


Fig. 4. The Gantt chart of IGAVNS algorithm under only PT

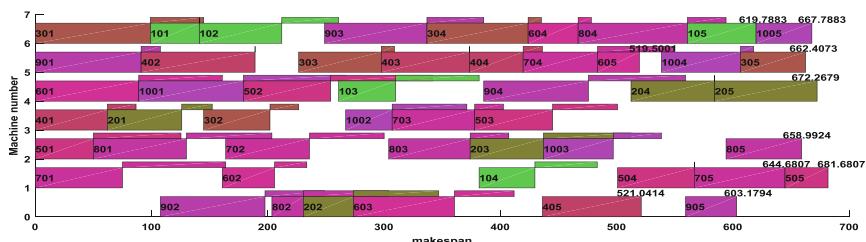


Fig. 5. The Gantt chart of IGAVNS algorithm under PT & TT

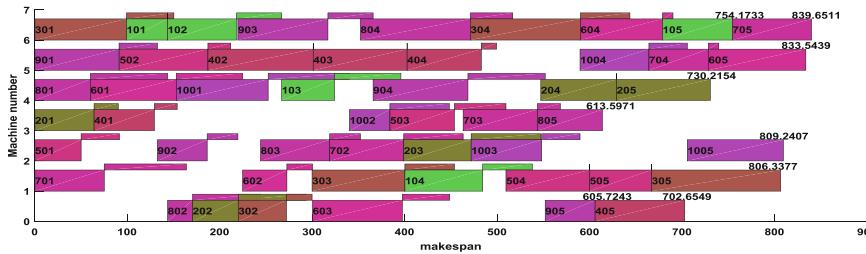


Fig. 6. The Gantt chart of IGAVNS algorithm under PT & TT & DT

The problem under different size are analyzed to verify the comprehensiveness of the proposed algorithm. 30 test cases are given for the combination of $n \times m = \{5,10,15,20,25\} \times \{5,6,7,8,9,10\}$. n is job amount while m is machine amount. Each case is marked in the form of “n_m”, such as “10_7”, which means the size of the problem is 10 jobs and 7 machines. The amount of operations for each job subjects to the discrete uniform distribution, DU (2 m). The transporting time is selected between the maximum and the minimum value of the processing time.

As is known to all, different parameter settings have a different impact on algorithm to solve the problem. Therefore, in order to test different parameters configuration, a lot of comparative experiments were carried out in advance. Since population size is sensitive to the scale of the problem, three different sizes (100, 200 and 300 respectively) are considered here, that is, small ($n \times m = \{5\} \times \{5,6,7,8,9,10\}$), medium ($n \times m = \{10, 15\} \times \{5,6,7,8,9,10\}$), large ($n \times m = \{20, 25\} \times \{5,6,7,8,9,10\}$). The result of each algorithm is presented as Table 3. Each result is the average of results that each algorithm is performed on each problem for 30 times independently.

Table 3. Results under different factors and different algorithms.

| Size | GA | | | IA | | |
|------|------------|------------|--------------|-----|---------|--------------|
| | PT | PT & TT | PT & TT & DT | PT | PT & TT | PT & TT & DT |
| 5_5 | 242 | 323 | 330 | 242 | 323 | 330 |
| 5_6 | 166 | 692 | 753 | 166 | 682 | 753 |
| 5_7 | 187 | 605 | 375 | 187 | 610 | 375 |
| 5_8 | 364 | 515 | 624 | 364 | 513 | 624 |
| 5_9 | 265 | 390 | 444 | 265 | 390 | 444 |
| 5_10 | 232 | 406 | 852 | 232 | 395 | 852 |
| 10_5 | 527 | 780 | 674 | 540 | 654 | 656 |
| 10_6 | 525 | 760 | 1188 | 516 | 725 | 1054 |
| 10_7 | 960 | 311 | 975 | 852 | 315 | 869 |
| 10_8 | 488 | 1210 | 2012 | 452 | 1045 | 1872 |

(continued)

Table 3. (*continued*)

| Size | GA | | | IA | | |
|-------|-------------|------------|--------------|------|---------|--------------|
| | PT | PT & TT | PT & TT & DT | PT | PT & TT | PT & TT & DT |
| 10_9 | 1239 | 1389 | 843 | 1201 | 1200 | 823 |
| 10_10 | 1044 | 960 | 1091 | 950 | 865 | 1045 |
| 15_5 | 935 | 900 | 2013 | 932 | 881 | 1954 |
| 15_6 | 532 | 793 | 526 | 501 | 756 | 510 |
| 15_7 | 1077 | 1649 | 3558 | 985 | 1550 | 3249 |
| 15_8 | 320 | 834 | 453 | 320 | 753 | 435 |
| 15_9 | 1438 | 610 | 1560 | 1420 | 515 | 1401 |
| 15_10 | 264 | 1141 | 392 | 267 | 1022 | 405 |
| 20_5 | 1366 | 1556 | 2036 | 1321 | 1437 | 1872 |
| 20_6 | 712 | 811 | 3981 | 705 | 825 | 3011 |
| 20_7 | 909 | 1672 | 2755 | 852 | 1587 | 2247 |
| 20_8 | 670 | 1499 | 2963 | 667 | 1369 | 2453 |
| 20_9 | 1002 | 1782 | 2267 | 1010 | 1645 | 2106 |
| 20_10 | 753 | 2116 | 2061 | 749 | 1801 | 1897 |
| 25_5 | 1652 | 685 | 1505 | 1546 | 710 | 1455 |
| 25_6 | 909 | 1394 | 3351 | 895 | 1254 | 2753 |
| 25_7 | 1353 | 1731 | 675 | 1321 | 1546 | 697 |
| 25_8 | 851 | 2566 | 3337 | 823 | 2178 | 2890 |
| 25_9 | 912 | 541 | 1050 | 925 | 545 | 913 |
| 25_10 | 1697 | 1149 | 2109 | 1569 | 1021 | 1783 |

From the experimental results, under the same factors, the proposed algorithms IGAVNS is superior to GA, Immune Algorithm (IA) and IGA on the whole (except for the Bold part). It means that local search can be better performed and better solutions can be found by the combination of genetic operator, immune operator and variable neighborhood search in IGAVNS. Comparing results of the same algorithms under different factors, the makespan is the smallest when only the processing time is considered. The makespan required is the longest when the processing time, the transporting time between machines and the deteriorating effect are considered. The makespan is medium when the processing time and the machine transporting time are considered. Furthermore, the scheduling result obtained by IGAVNS is more realistic when the transporting time and the deteriorating effect are considered (Table 4).

Table 4. Results under different factors and different algorithms

| Size | IGA | | | IGAVNS | | |
|-------|------------|------------|--------------|--------|---------|--------------|
| | PT | PT & TT | PT & TT & DT | PT | PT & TT | PT & TT & DT |
| 5_5 | 242 | 323 | 330 | 242 | 323 | 330 |
| 5_6 | 166 | 679 | 753 | 166 | 625 | 733 |
| 5_7 | 187 | 605 | 375 | 187 | 605 | 375 |
| 5_8 | 364 | 514 | 620 | 366 | 515 | 620 |
| 5_9 | 265 | 390 | 444 | 265 | 390 | 444 |
| 5_10 | 232 | 388 | 850 | 232 | 388 | 842 |
| 10_5 | 537 | 589 | 562 | 524 | 596 | 550 |
| 10_6 | 501 | 635 | 922 | 510 | 630 | 968 |
| 10_7 | 822 | 311 | 779 | 737 | 311 | 804 |
| 10_8 | 422 | 960 | 1515 | 433 | 916 | 1452 |
| 10_9 | 1041 | 1059 | 726 | 1035 | 1055 | 708 |
| 10_10 | 933 | 765 | 910 | 906 | 806 | 855 |
| 15_5 | 930 | 870 | 1657 | 866 | 828 | 1578 |
| 15_6 | 479 | 690 | 458 | 462 | 671 | 458 |
| 15_7 | 972 | 1387 | 2259 | 930 | 1261 | 2393 |
| 15_8 | 318 | 696 | 418 | 318 | 715 | 418 |
| 15_9 | 1310 | 466 | 1246 | 1254 | 465 | 1371 |
| 15_10 | 257 | 943 | 391 | 257 | 904 | 391 |
| 20_5 | 1289 | 1307 | 1561 | 1289 | 1287 | 1509 |
| 20_6 | 692 | 730 | 2840 | 636 | 726 | 2328 |
| 20_7 | 822 | 1305 | 2168 | 784 | 1302 | 2138 |
| 20_8 | 662 | 1276 | 1993 | 662 | 1166 | 1686 |
| 20_9 | 955 | 1398 | 1481 | 954 | 1287 | 1377 |
| 20_10 | 633 | 1608 | 1316 | 602 | 1529 | 1199 |
| 25_5 | 1533 | 692 | 1390 | 1455 | 643 | 1356 |
| 25_6 | 880 | 1136 | 2139 | 871 | 1123 | 2521 |
| 25_7 | 1263 | 1384 | 655 | 1240 | 1347 | 637 |
| 25_8 | 760 | 2066 | 2510 | 688 | 1895 | 2174 |
| 25_9 | 806 | 541 | 873 | 804 | 541 | 828 |
| 25_10 | 1456 | 883 | 1401 | 1279 | 817 | 1280 |

5 Conclusion

This article focus on time factors in the FJSP. A disjunctive graph model and a linear deteriorating function was established, considering the processing time, the transporting time and the deteriorating effect. Then, the IGAVNS under the three factors was proposed to solve FJSP. Simulations were carried out by different algorithms on the cases of different sizes. The results were compared and analyzed, conforming that the

proposed algorithms IGAVNS is superior to GA, IA and IGA on the whole. Moreover, the results obtained by IGAVNS are more realistic when the transporting time and the deteriorating effect are considered.

References

1. Kacem I, Hammadi S, Borne P (2002) Approach by localization and multiobjective evolutionary optimization for flexible job-shop scheduling problems. *IEEE Trans Syst Man Cybern Part C (Appl Rev)* 32(1):1–13
2. Pezzella F, Morganti G, Ciaschetti G (2008) A genetic algorithm for the flexible job-shop scheduling problem. *Comput Oper Res* 35(10):3202–3212
3. Xing LN, Chen YW, Wang P et al (2010) A knowledge-based ant colony optimization for flexible job shop scheduling problems. *Appl Soft Comput* 10(3):888–896
4. Jia Z, Chen H, Tang J (2007) A new multi-objective fully-informed particle swarm algorithm for flexible job-shop scheduling problems. In: International conference on computational intelligence & security workshops. IEEE
5. Zhang GH et al (2019) Solving flexible job shop scheduling problems with transportation time based on improved genetic algorithm. *Math Biosci Eng: MBE* 16:1334–1347
6. Ji M, Fang J, Zhang W et al (2018) Logistics scheduling to minimize the sum of total weighted inventory cost and transport cost. *Comput Ind Eng* 120:206–215
7. Hurink J, Knust S (2005) Tabu search algorithms for job-shop problems with a single transport robot. *Eur J Oper Res* 162(1):99–111
8. Gupta JND, Gupta SK (1988) Single facility scheduling with nonlinear processing times. *Comput Ind Eng* 14(4):387–393
9. Wu HP, Huang M (2014) Improved estimation of distribution algorithm for the problem of single-machine scheduling with deteriorating jobs and different due dates. *Comput Appl Math* 33(3):557–573
10. Mosheiov G (1994) Scheduling jobs under simple linear deterioration. Elsevier Science Ltd.
11. Wu CC, Wu WH, Wu WH et al (2014) A single-machine scheduling with a truncated linear deterioration and ready times. *Inf Sci* 256:109–125
12. Cheng M, Tadikamalla PR, Shang J et al (2014) Bicriteria hierarchical optimization of two-machine flow shop scheduling problem with time-dependent deteriorating jobs. *Eur J Oper Res* 234(3):650–657
13. Wang JB, Wang MZ (2013) Solution algorithms for the total weighted completion time minimization flow shop scheduling with decreasing linear deterioration. *Int J Adv Manuf Technol* 67(1–4):243–253
14. Liu CH, Chen LS, Lin PS (2013) Lot streaming multiple jobs with values exponentially deteriorating over time in a job-shop environment. *Int J Prod Res* 51(1):202–214
15. Mosheiov G (2002) Complexity analysis of job-shop scheduling with deteriorating jobs. *Discrete Appl Math* 117(1–3):195–209
16. Tayebi Araghi ME, Jolai F, Rabiee M (2014) Incorporating learning effect and deterioration for solving a SDST flexible job-shop scheduling problem with a hybrid meta-heuristic approach. *Int J Comput Integ Manuf* 27(8):733–746
17. Wang L (2003) Shop scheduling with genetic algorithms. Tsinghua University Press, Beijing (in Chinese)
18. Xu X, Wang W, Guan Q (2005) Adaptive immune algorithm for solving job-shop scheduling problem. In: International conference on natural computation. Springer, Heidelberg



Deep Learning and Binocular Stereovision to Achieve Fast Detection and Location of Target

Qingbin Wang¹, Yueqian Liang¹, Ziteng Wang², Wenyan Li¹,
Zhiguo Jiang³, and Yanjie Zhao¹⁽⁾

¹ China Academy of Electronics and Information Technology,
Beijing 100041, China
zhaoyj_dyk@163.com

² China Railway Information Technology Co. Ltd., Beijing 100038, China
³ Beihang University, Beijing 100191, China

Abstract. Air targets are often fast and varied. For air targets monitoring tasks, traditional methods tend to be slow and resource-intensive. Therefore, this paper proposes a method of target detection and location using binocular synchronous camera as acquisition device and combining SSD, ORB and binocular stereo vision. Firstly, the left and right images collected synchronously are detected by SSD, and the ROI region of the target is taken as a new image. Then, the new sub-images are detected and matched by ORB algorithm, and the matched feature points are corrected. Then, the three-dimensional coordinates of the target are obtained by binocular stereo vision. This method maximizes the speed of target detection and location. For this method, we validate it through simulation experiments. The experimental results show that the speed of this method in single target detection and location can reach 12 frames per second when the left and right image resolutions are 1280×720 respectively. The experimental results show that this method is effective and high-speed.

Keywords: SSD · ORB · Binocular stereovision · Target detection · Target location

1 Introduction

In the field of aerospace or low altitude monitoring, it is often necessary to detect specific target categories. After detecting the target categories, if the distance of the target can be measured, it can help operators make better decisions. However, due to the high speed of aerospace or UAV equipment, in order to help operators make decisions, the detection and measurement algorithms need to be fast enough.

For a long time, the monitoring of aerial targets mainly depends on large size, high power consumption, expensive radar and other products. These products are very bulky, consume a lot of resources, and have a great impact on the human body and the surrounding environment. At the same time, these methods are generally difficult to install on small equipment such as UAV. Moreover, the measurement range of these devices is often relatively small, which further limits the use of UAV.

Camera-based target detection has made great progress. For example, in the application of aerial refueling, there are many cone detection methods based on monocular camera [1]. At the same time, there are many methods based on aerial video to detect and track human body, face and other targets [2]. Because of the advantages of low cost, light weight, convenient installation and low power consumption, camera has been widely used in many civil UAV fields. Some camera-based target detection methods are based on traditional machine learning methods [3–5], and some are based on the recently popular in-depth learning methods [6, 7]. Obviously, the method based on deep learning has better universality. The deep learning methods used in target detection mainly include FAST-RCNN [8], Faster-RCNN [9], YOLO [10], SSD [11] and MASK-RCNN [12]. The speed and accuracy of these methods are gradually improved. It lays a very good foundation for the application of deep learning in the field of target detection.

Monocular camera can detect targets through deep learning, but if the location information of targets can be obtained, it will greatly help operators to make decisions. Obviously, binocular cameras can accomplish such functions. Using binocular stereo vision, the three-dimensional coordinates of the target can be obtained. The process is as follows: firstly, the target is detected in the images of left and right, then some key feature points are extracted by using SIFT [13], SURF [14], ORB [15] and other feature points detection methods, and then matched according to these feature points using Hamming distance and other methods. Finally, the camera's internal and external parameters are corrected, and then the binocular and external parameters are combined and the triangulation method is used to get the object. The three-dimensional coordinates of the target are in place. But if the whole process implements such a function, it often requires a very large time overhead. Therefore, it is necessary to design a fast target detection and location method.

2 Detection and Location Methods

To meet the above requirements, this paper proposes a method of target detection using SSD and target location using binocular stereo vision. The flow chart is shown in Fig. 1. Firstly, for the left image and the right image, the target is extracted by SSD. After the left and right images are extracted, the ROI region containing the target is extracted as a new image, and the ORB feature points are detected. At the same time, the feature points are matched. For the matched points, the calibrated camera parameters and distortion parameters are used to correct the feature points. After obtaining the corrected feature points, the three-dimensional coordinates of the feature points can be obtained by using the calibrated binocular external parameters and the left-right camera internal parameters. For the three-dimensional coordinates of multiple points, the z-coordinates of these points are taken to generate a new array. The points corresponding to the z-coordinates of the median of the array are taken as the central points of the target, and the three-dimensional coordinates of the target are taken as the three-dimensional coordinates of the target. In this way, the processing speed of the algorithm can be greatly reduced, and the universality of the algorithm can be improved. This method has a great use value for long-distance tracking of special targets.

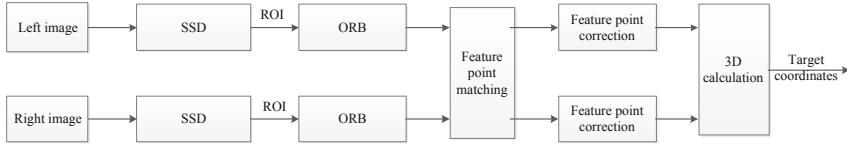


Fig. 1. Flow chart of detection and location methods

2.1 Target Detection

SSD algorithm is a target detection method based on convolution neural network, and its speed is very fast. SSD is a multi-target detection algorithm that directly predicts target categories and bounding box. For the task of detecting targets with different sizes, the general method is to use image pyramid principle to detect targets at different sizes and levels. Finally, the overlapped targets are synthetically optimized by non-maximum suppression (NMS). This method often requires a lot of time and training speed is slow. For this reason, SSD algorithm synthesizes feature maps from different convolution layers, and finally obtains the target location. The SSD algorithm makes full use of the traditional VGG16 [16] network structure, and modifies the network structure of VGG16.

In SSD, the last two full connection layers of VGG16 are transformed into convolution layers, and four convolution layers are added to construct a new network structure. The new network structure is shown in Fig. 2. Five feature maps output from different convolution layers are convoluted with two different 3×3 convolution cores, one for output classification, 21 for each default box, and one for output regression. Each default box generates four coordinate values (x, y, w, h). In addition, these five feature maps generate prior boxes through the priorbox layer. The prebox refers to the default box actually selected. The number of default boxes for each of the five feature maps is given (8732). Finally, the first three calculation results are merged separately and passed to loss layer.

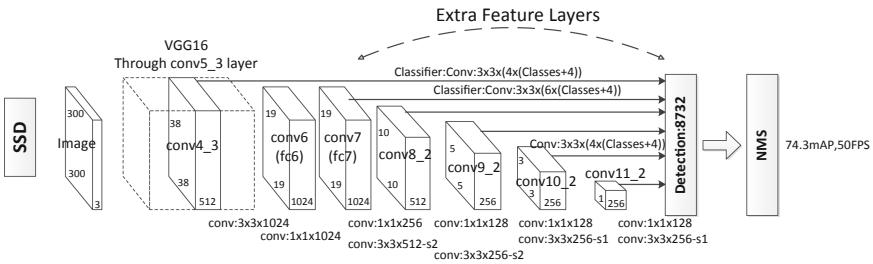


Fig. 2. Network architecture used by SSD

2.2 Feature Points Detection and Matching

ORB is a fast feature extraction and description algorithm. ORB algorithm is divided into two parts: feature extraction and feature description. Feature extraction is

developed by FAST (Features from Accelerated Segment Test) algorithm, and feature point description is improved according to BRIEF (Binary Robust Independent Elementary Features). ORB feature combines FAST feature point detection method with BRIEF feature descriptor, and optimizes them on their original basis. ORB algorithm is divided into four steps: the first step is to select a point P from the image, draw a circle with a radius of 3 pixels clockwise with P as the center of the circle, and label the pixels passing through the circle from 1 to 16, respectively. If the gray value of continuous N pixels on a circle is larger or smaller than that of P , then P is considered as a feature. In order to speed up the extraction of feature points and expel non-feature points quickly, the gray values at positions 1, 9, 5 and 13 are detected first. If P is a feature point, three or more pixel values at these four locations are larger or smaller than the gray values of P . If it is not satisfied, this point is excluded directly. This will extract a large number of feature points, but there are many invalid points. In the second step, ID3 algorithm is used to train a decision tree, and 16 pixels on the circle of feature points are input into the decision tree, which is used to select the optimal FAST feature points. In the third step, non-maximum suppression (NMS) algorithm is used to remove multiple feature points from adjacent positions. In the fourth step, image pyramids are built to extract feature points at each level to achieve multi-scale invariance of feature points. The fifth step is to calculate the centroid of the feature point in the radius by moment, and form a vector from the coordinates of the feature point to the centroid as the direction of the feature point. In order to ensure the invariance of feature rotation, the 256-dimensional descriptor is constructed by rotating the Patch of each feature point to the main direction for the points obtained by the original FAST, using the optimal 256 pairs of feature points selected in step 3 above as τ test. Through the above operations, we can get the position and direction of the features in different scales.

After obtaining the feature points and descriptors of the left and right images respectively, matching point pairs can be obtained by exhaustive method. The most important characteristic of ORB algorithm is its fast computing speed. This is first of all due to the use of FAST to detect feature points. FAST detection speed is as famous as its name is fast. Thirdly, BRIEF algorithm is used to calculate descriptors. The unique representation of binary strings of the descriptors not only saves storage space, but also greatly reduces the matching time. When matching feature points, Hamming distance is used to calculate the similarity of each feature point, and then feature points are matched.

2.3 Target Location

By adding the initial position coordinates of ROI sub graph on the basis of the feature points obtained above, the coordinate positions of the feature points on the original image can be obtained. But these feature points are affected by image distortion. Before using binocular vision to get the three-dimensional coordinates of matching points, the feature points need to be corrected. According to Brown distortion model [17], feature points can be corrected by using camera distortion parameters $(k_1, k_2, k_3, p_1, p_2)$.

After obtaining the corrected left and right binocular matching points, the coordinates of the points in the real world corresponding to each pair of feature points can be obtained by using the binocular internal and external parameters. Let the center point of

the right camera be $P_0^r(x_0^r, y_0^r, z_0^r)$, the coordinates of the feature point on the right camera image are $P_m^r(u_m^r, v_m^r, 1)$, the center point of the left camera is $P_0^l(x_0^l, y_0^l, z_0^l)$, and the coordinates of the feature point on the left camera image are $P_m^l(u_m^l, v_m^l, 1)$. The inner parameters of the left camera are M_{in}^l . The inner parameters of the right camera are M_{in}^r , and the rotation matrix of the right camera relative to the left camera is R and the translation vector is T .

We use the right camera coordinate system as the world coordinate system. The coordinates of the points on the left camera image on the normalized imaging plane are:

$$\mathbf{P}_{1c}^l = (M_{in}^l)^{-1} \mathbf{P}_m^l \quad (1)$$

The coordinates of the points on the right camera image on the normalized imaging plane are:

$$\mathbf{P}_{1c}^r = (M_{in}^r)^{-1} \mathbf{P}_m^r \quad (2)$$

The origin of the left camera coordinate system is ${}^rP_0^l({}^rx_0^l, {}^ry_0^l, {}^rz_0^l)$ in the right camera coordinate system.

$${}^r\mathbf{P}_0^l = R\mathbf{P}_0^l + T \quad (3)$$

The coordinates of the left camera image feature points ${}^rP_{1c}^l({}^rx_{1c}^l, {}^ry_{1c}^l, {}^rz_{1c}^l)$ in the right camera coordinate system are:

$${}^r\mathbf{P}_{1c}^l = R\mathbf{P}_{1c}^l + T \quad (4)$$

So we can get two rays ${}^rP_0^l {}^rP_{1c}^l$ and $P_0^l P_{1c}^l$.

$$\frac{x - {}^rx_{1c}^l}{{}^rx_0^l - {}^rx_{1c}^l} = \frac{y - {}^ry_{1c}^l}{{}^ry_0^l - {}^ry_{1c}^l} = \frac{z - {}^rz_{1c}^l}{{}^rz_0^l - {}^rz_{1c}^l} = t \quad (5)$$

$$\frac{x - {}^rx_{1c}^l}{{}^rx_0^l - {}^rx_{1c}^l} = \frac{y - {}^ry_{1c}^l}{{}^ry_0^l - {}^ry_{1c}^l} = \frac{z - {}^rz_{1c}^l}{{}^rz_0^l - {}^rz_{1c}^l} = s \quad (6)$$

Let

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & -({}^rx_0^l - {}^rx_{1c}^l) & 0 \\ 0 & 1 & 0 & -({}^ry_0^l - {}^ry_{1c}^l) & 0 \\ 0 & 0 & 1 & -({}^rz_0^l - {}^rz_{1c}^l) & 0 \\ 1 & 0 & 0 & 0 & -({}^rx_0^l - {}^rx_{1c}^l) \\ 0 & 1 & 0 & 0 & -({}^ry_0^l - {}^ry_{1c}^l) \\ 0 & 0 & 1 & 0 & -({}^rz_0^l - {}^rz_{1c}^l) \end{pmatrix} \quad (7)$$

$$\mathbf{B} = \left(-{}^r x_{1c}^l - {}^r y_{1c}^l - {}^r z_{1c}^l - x_{1c}^r - y_{1c}^r - z_{1c}^r \right)^T \quad (8)$$

$$\mathbf{X} = (x \ y \ z \ s \ t)^T \quad (9)$$

By solving the above equation with least squares, the solution (x, y, z, s, t) that minimizes $|AX - B|^2$ can be obtained. For the three-dimensional coordinates of multiple feature points, the three-dimensional coordinates corresponding to the feature points in the middle of the z-coordinate row are taken as the three-dimensional coordinates of the target.

3 Experiments

For the above mentioned methods, this paper uses the binocular platform to experiment. The platform realizes the acquisition of binocular video based on FPGA, and achieves synchronous acquisition on hardware. The focal length of the camera is 3.1 mm, and the resolution of each camera is 1280×720 . The frame rate is 30 fps. The goal of this article is to display the aircraft icon on the display screen. SSD uses 17,000 training models for aircraft detection.

In the specific experiment, the aircraft icon is projected on a display screen, and the binocular camera is placed about one meter in front of the display screen. The binocular camera collects real-time images through USB3.0. The experimental results are shown in Figs. 3 and 4. Figure 3 is the result image after SSD target detection. After obtaining the target position, the ROI sub image is obtained by enlarging the target position by 30 pixels. Figure 4 is an image after ORB feature point detection and matching of ROI sub image. The three-dimensional coordinates of the target are $(-86.1658, -22.5565, 1207.31)$, which coincide with the position of the target and the binocular camera. Table 1 lists the time of each step. Among them, the time of detecting left and right images by SSD is 81 ms, and the total time of these steps is 82.42 ms.

Therefore, the proposed method not only meets the real-time requirements in speed, but also performs well in accuracy and stability.

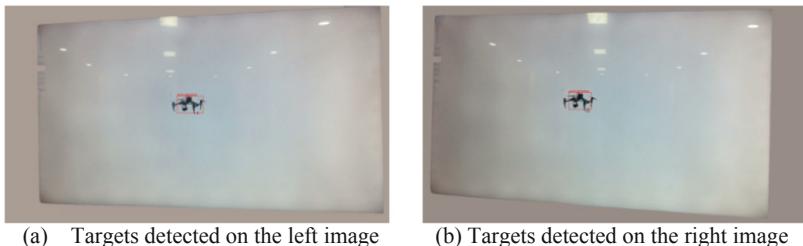


Fig. 3. Left and right result images detected by SSD

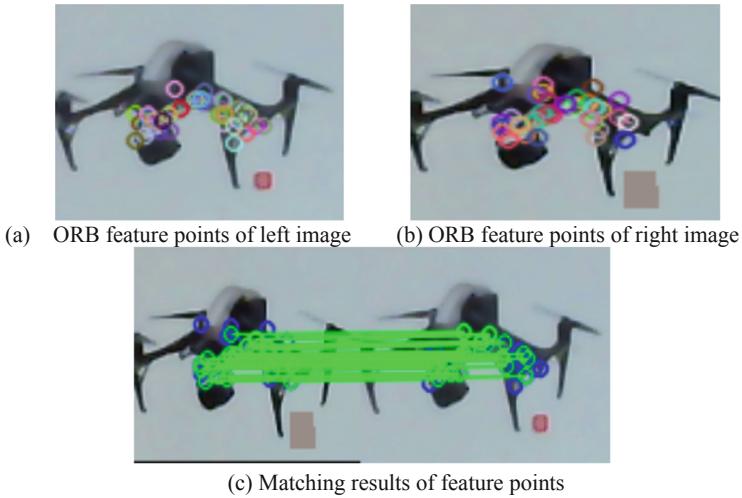


Fig. 4. Feature points of left and right targets detected by ORB and feature points on matching

Table 1. Time spent in each step.

| Item | Time (ms) |
|--------------------------|-----------|
| SSD | 81.0 |
| ORB | 1.22 |
| Matching | 0.06 |
| Feature point correction | 0.02 |
| 3D calculation | 0.12 |
| Sum | 82.42 |

4 Conclusion

The method of combining SSD and binocular stereo vision to detect and locate targets proposed in this paper is not only fast, but also accurate. The SSD algorithm based on VGG16 network has the characteristics of high speed and accuracy, and it is also very suitable for engineering applications. The SSD algorithm based on QT language realizes the engineering application of target detection very well. After detecting the target, we extract ROI region as a new image for ORB algorithm, which further improves the speed of the algorithm. At the same time, ORB algorithm is a very fast feature point detection method. It also has rotation invariance and scale invariance. The feature point matching method based on Hamming distance is also fast. After the matched feature points are obtained, the feature points are corrected separately. This method avoids the correction of the whole image and greatly improves the speed of distortion correction. By using binocular synchronous acquisition of stereo vision method, the three-dimensional coordinates of the target with very high accuracy are obtained. The exact position of the target can be obtained by extracting the median of the three-dimensional coordinates. Target detection and location methods obtained by synthesizing various fast methods

not only improve the speed, but also ensure the accuracy. However, this method has not considered the situation of multi-target and complex background. In the later research, we will use the binocular position relationship to detect and locate multi-target in complex background.

References

1. Wilson DB, Göktoğan AH, Sukkarieh S (2015) Experimental validation of a drogue estimation algorithm for autonomous aerial refueling. In: Proceedings of the IEEE international conference on robotics and automation (ICRA), pp 5318–5323
2. Portmann J, Lynen S, Chli M et al (2014) People detection and tracking from aerial thermal views. In: Proceedings of the IEEE international conference on robotics and automation (ICRA), pp 1794–1800
3. Gossow D, Weikersdorfer D, Beetz M (2012) Distinctive texture features from perspective-invariant keypoints. In: Proceedings of the 21st international conference on pattern recognition (ICPR 2012), pp 2764–2767
4. Langer M, Kuhnert KD (2008) A new hierarchical approach in robust real-time image feature detection and matching. In: Proceedings of the 2008 19th international conference on pattern recognition, pp 1–4
5. Kaur J, Bathla AK (2016) Video stabilization for an aerial surveillance system using SIFT and SURF. In: Proceedings of the 2016 2nd international conference on next generation computing technologies (NGCT), pp 742–747
6. Sangineto E, Nabi M, Culibrk D et al (2019) Self paced deep learning for weakly supervised object detection. *IEEE Trans Pattern Anal Mach Intell* 41(3):712–725
7. Zhou X, Wei G, Fu WL et al (2017) Application of deep learning in object detection. In: Proceedings of the 2017 IEEE/ACIS 16th international conference on computer and information science (ICIS), pp 631–634
8. Girshick R (2015) Fast R-CNN. In: Proceedings of the 2015 IEEE international conference on computer vision (ICCV), pp. 1440–1448
9. Ren S, He K, Girshick R et al (2017) Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell* 39(6):1137–1149
10. Redmon J, Divvala S, Girshick R et al (2016) You only look once: unified, real-time object detection. In: Proceedings of the 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 779–788
11. Liu W, Anguelov D, Erhan D et al (2016) SSD: single shot multibox detector. In: Proceedings of the European conference on computer vision, pp 21–37
12. He K, Gkioxari G, Dollar P et al (2018) Mask R-CNN. *IEEE Trans Pattern Anal Mach Intell*. <https://doi.org/10.1109/tpami.2018.2844175>
13. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110
14. Bay H, Tuytelaars T, Gool LV (2006) SURF: speeded up robust features. In: Proceedings of the European conference on computer vision, pp 404–417
15. Rublee E, Rabaud V, Konolige K et al (2011) ORB: an efficient alternative to SIFT or SURF. In: Proceedings of the international conference on computer vision, pp 2564–2571
16. Chatfield K, Simonyan K, Vedaldi A et al (2014) Return of the devil in the details: delving deep into convolutional nets. In: Proceedings of the British machine vision conference. [arXiv:1405.3531](https://arxiv.org/abs/1405.3531)
17. Brown DC (1971) Close-range camera calibration. *Photogramm Eng* 37:855–866



Position Tracking and Ellipse Fitting of UAV Flying by Elliptical Trajectory

Qingbin Wang¹, Zhiqiang Jia¹, Yueqian Liang¹, Xiangfeng Wang¹,
Zhiguo Jiang², and Yanjie Zhao^{1(✉)}

¹ China Academy of Electronics and Information Technology,
Beijing 100041, China
zhaoyj_dyk@163.com

² Beihang University, Beijing 100191, China

Abstract. The abuse of civil UAVs poses a great threat to many important facilities. Position tracking and intention discrimination of aerial targets such as civil UAVs can help operators stop the damage caused by UAVs to important facilities in time. Aiming at UAV running on elliptical trajectory, this paper presents a method to track its position and fit its elliptical trajectory according to its historical position. Firstly, the binocular camera is used to obtain the three-dimensional coordinates of the target, then the state of the target is estimated by the UKF, and finally the ellipse is fitted according to the historical position of the target. To a certain extent, this method realizes the automatic identification of target intention. The simulation results show that this method can realize UAV position tracking. At the same time, after a certain number of historical positions have been reached, the ellipse can be fitted to its trajectory.

Keywords: UAV · Position tracking · Kalman filter · Ellipse fitting

1 Introduction

Many important facilities, such as civil airports, are often disturbed by UAVs. These unmanned aerial vehicles are small and difficult to monitor, which has caused tremendous trouble to the maintenance personnel of the facilities. Radar is a kind of equipment with huge size, large resource consumption and high cost. When these devices are used to detect the small-scale equipment of UAV, there are often large misjudgments. Because these UAV images on radar often only have contour features, but do not have texture features such as color. But for the camera, not only the contour information of the target can be obtained, but also the color and texture features of the target can be obtained. Researchers have carried out a lot of research work on the detection, tracking and positioning of specific targets. At the same time, in the military field, many camera-based positioning and tracking tasks have also achieved great results, such as in the field of aerial refueling [1] and aerial target recognition, positioning and other fields [2].

For UAVs with special intentions, they tend to have specific laws of motion. For example, when a target moves in a circular or elliptical trajectory, it is often reconnaissance or photography. If we can predict the target's later movement according to its

early movement law, it will provide the operator with very critical hints to prevent UAV injury. 3D coordinates of targets can be obtained by binocular method, which has been applied in many tasks [3, 4], such as unmanned vehicle and so on. Binocular acquisition of UAV's 3D information often results in relatively large noise. In order to reduce noise, a filter can be used. There are many kinds of filters that can be used to reduce noise, such as mean filter, high-pass filter, low-pass filter, KF (Kalman Filter) [5] and PF (Particle Filter) [6]. KF is the most widely used method for state estimation of aerial targets. However, KF is usually only applicable to the state estimation of linear moving targets. The UKF (Unscented Kalman Filter) [7] is a method suitable for state estimation of non-linear targets. This method is faster than PF and better than KF. Therefore, this paper uses the unscented KF to estimate the state of the target and remove the noise from the 3D coordinate acquisition. Aiming at the target with elliptic trajectory, this paper uses ellipse fitting method to obtain the trajectory information of the target.

The fitting of ellipse in 3D space is very different from that of 2D (two-dimensional) ellipse in plane. Traditional ellipse fitting in 2D plane mostly uses Hough ellipse fitting [8] or least-squares ellipse fitting method [9]. The fitting of ellipse in 3D space is to fit an ellipsoid [10], and then a plane [11]. The intersection curve of the ellipsoid and the plane is the ellipse in 3D space.

In summary, this paper uses binocular to obtain the 3D information of the target, and uses UKF to remove the noise in the 3D coordinates. Then the least-square method is used to fit the ellipsoid. At the same time, combining the RANSAC method to fit the plane, the ellipse in 3D space is finally obtained by the intersection of the ellipsoid and the plane.

2 Position Tracking and Ellipse Fitting

The position tracking and ellipse fitting method firstly uses the left and right images to obtain the 3D coordinates of the target according to the precalibrated camera internal and external parameters. After obtaining the 3D coordinates, the state of the target is estimated by using the UKF, and the 3D coordinates with large errors are removed. When a certain number of 3D coordinates are obtained, N of them are randomly selected for ellipsoid fitting. On the basis of successful ellipsoid fitting, the RANSAC method is used to fit the plane, and finally the parameterized ellipse equation is obtained. The flow chart of position tracking and ellipse fitting method is shown in Fig. 1.

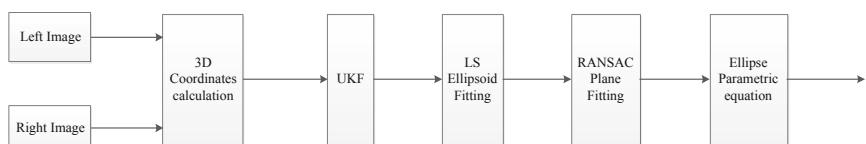


Fig. 1. Position tracking and ellipse fitting process

2.1 3D Coordinates Extraction

This paper uses SSD to detect targets and binocular stereo vision to locate targets. The flow chart is shown in Fig. 2. Firstly, the left and right images are acquired by binocular camera. Then, the target is extracted by SSD. When both left and right images are extracted to the target, the ROI region containing the target is extracted as a new image. ORB feature points are detected and matched for the new image. For the matched points, the calibrated camera parameters and distortion parameters are used to correct the feature points. After obtaining the corrected feature points, the 3D coordinates of the feature points are obtained by using the calibrated binocular external parameters and the left and right camera internal parameters. For the obtained 3D coordinates, the z-coordinates of these points are taken to generate a new array. The point corresponding to the z-coordinate of the array median is the center point of the target, and its 3D coordinate is the 3D coordinates of the target. In this way, the processing speed of the algorithm can be greatly reduced and the universality of the algorithm can be improved. But this method only filters in z-direction.

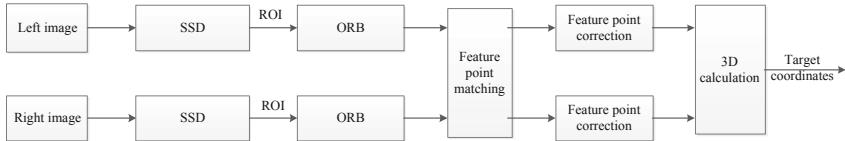


Fig. 2. Binocular stereovision flow chart

2.2 Unscented Kalman Filter

UKF is a non-linear Gauss model [7], which approximates the statistical properties of random variables by using finite parameters. UKF calculates the integral terms in recursive Bayesian by statistical method, and then obtains the mean and variance of posterior probability. For non-linear models, it is difficult to solve Bayesian recursive formulas directly by analytic method, but UT transformation can solve this problem better. UT transform can obtain the mean and variance approximately by sampling and weighting regularly. Moreover, the approximation accuracy of insensitive transformation to statistical moments is higher.

For UAV target, the 3D coordinates of the target have been obtained by binocular stereo vision. The target state obtained for the first time is taken as the initial state. Suppose the target motion system model is:

$$\mathbf{X}(k) = \mathbf{A}\mathbf{X}(k-1) \quad (1)$$

$$\mathbf{Z}(k) = \mathbf{H}\mathbf{X}(k) \quad (2)$$

where:

$$\mathbf{X}(k) = (x(k) \ y(k) \ z(k) \ x_v(k) \ y_v(k) \ z_v(k) \ x_a(k) \ y_a(k) \ z_a(k))$$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & T & 0 & 0 & 0.5T^2 & 0 & 0 \\ 0 & 1 & 0 & 0 & T & 0 & 0 & 0.5T^2 & 0 \\ 0 & 0 & 1 & 0 & 0 & T & 0 & 0 & 0.5T^2 \\ 0 & 0 & 0 & 1 & 0 & 0 & T & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & T & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & T \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Let \mathbf{Q} be the covariance matrix of \mathbf{X} . \mathbf{R} is the covariance matrix of \mathbf{Z} . The mean of \mathbf{X} is $\bar{\mathbf{X}}$. The variance of \mathbf{X} is \mathbf{P} . The detailed process of UKF is divided into five steps.

In the first step, $2n+1$ Sigma point and corresponding weight ω are obtained by UT transformation. The calculation method of Sigma point is as follows:

$$\begin{cases} \chi_i(k-1) = \bar{\mathbf{X}}, i = 0 \\ \chi_i(k-1) = \bar{\mathbf{X}} + (\sqrt{(n+\lambda)\mathbf{P}})_i, i = 1 \sim n \\ \chi_i(k-1) = \bar{\mathbf{X}} - (\sqrt{(n+\lambda)\mathbf{P}})_i, i = n+1 \sim 2n \end{cases} \quad (3)$$

where $(\sqrt{\mathbf{P}})^T(\sqrt{\mathbf{P}}) = \mathbf{P}$, $(\sqrt{\mathbf{P}})_i$ is the first column of the square root of a matrix. The parameter λ is:

$$\lambda = \alpha^2(n + \kappa) - n \quad (4)$$

where, α and κ are constant.

After the Sigma point is obtained, the weights are then calculated. The weights of mean and variance of the first Sigma point are:

$$\omega_0^m = \frac{\lambda}{n + \lambda} \quad (5)$$

$$\omega_0^c = \frac{\lambda}{n + \lambda} + 1 - \alpha^2 + \beta \quad (6)$$

The superscript m denotes the mean and the superscript c denotes the variance. β is also a constant. For the remaining $2n$ Sigma points, the expected weights and variances are the same:

$$\omega_i^m = \omega_i^c = \frac{1}{2(n + \lambda)} \quad (7)$$

In the second step, the Sigma point obtained in the first step is introduced into the model to calculate the state of the Sigma point at k :

$$\chi_i(k) = A\chi_i(k - 1) \quad (8)$$

Then, one-step prediction and covariance matrix of system state variables are calculated:

$$\hat{X}(k|k - 1) = \sum_{i=0}^{2n} \omega_i \chi_i(k) \quad (9)$$

$$P(k|k - 1) = \sum_{i=0}^{2n} \omega_i [\chi_i(k) - \hat{X}(k|k - 1)][\chi_i(k) - \hat{X}(k|k - 1)]^T + Q \quad (10)$$

In the third step, the observed values of Sigma point are calculated:

$$Z_i(k) = H\chi_i(k) \quad (11)$$

The predictive mean and covariance of the system are obtained by weighted summation:

$$\hat{Z}(k|k - 1) = \sum_{i=0}^{2n} \omega_i Z_i(k) \quad (12)$$

$$P_{zz} = \sum_{i=0}^{2n} \omega_i [Z_i(k) - \hat{Z}(k|k - 1)][Z_i(k) - \hat{Z}(k|k - 1)]^T + R \quad (13)$$

$$P_{xz} = \sum_{i=0}^{2n} \omega_i [\chi_i(k) - \hat{X}(k|k - 1)][Z_i(k) - \hat{Z}(k|k - 1)]^T \quad (14)$$

The fourth step is to calculate Kalman gain $K(k)$ and update the state and variance.

$$K(k) = P_{xz} P_{zz}^{-1} \quad (15)$$

Fifth, state and variance updates:

$$\hat{\mathbf{X}}(k|k) = \hat{\mathbf{X}}(k|k-1) + \mathbf{K}(k)[\mathbf{Z}(k) - \hat{\mathbf{Z}}(k|k-1)] \quad (16)$$

$$\mathbf{P}(k|k) = \mathbf{P}(k|k-1) - \mathbf{K}(k)\mathbf{P}_{zz}\mathbf{K}^T(k) \quad (17)$$

Through the above five steps, the target state can be obtained after the UKF.

2.3 The Fitting of Ellipse in 3D Space

Through the above steps, and after accumulating enough points, the ellipse can be fitted. The reason why enough points are needed is that there will be failure in ellipsoid fitting. The fitting of ellipse in 3D space is divided into two steps. The first step is ellipsoid fitting.

Assuming that S 3D points are obtained, N of them are randomly selected for ellipsoid fitting, and if the fitting is unsuccessful, they are reselected. If S target points are directly used for ellipsoid fitting, the fitting failure will easily occur. Let the standard equation of the ellipsoid be:

$$\left(\frac{x-x_0}{A}\right)^2 + \left(\frac{y-y_0}{B}\right)^2 + \left(\frac{z-z_0}{C}\right)^2 = 1 \quad (18)$$

There are six unknown parameters: x_0, y_0, z_0, A, B, C . Write as follows:

$$x^2 + ay^2 + bz^2 + cx + dy + ez + f = 0 \quad (19)$$

where

$$a = \frac{A^2}{B^2}, b = \frac{A^2}{C^2}, c = -2x_0, d = -2y_0 \frac{A^2}{B^2}, e = -2z_0 \frac{A^2}{C^2}, f = x_0^2 + \frac{A^2}{B^2}y_0^2 + \frac{A^2}{C^2}z_0^2 - A^2$$

When the point number of ellipsoid fitting is N , we first estimate the parameters (a, b, c, d, e, f) and then we can get (x_0, y_0, z_0, A, B, C) according to the above relationship. The ellipsoid fitting based on least squares can easily calculate the six parameters (a, b, c, d, e, f). Firstly, for each sample point (x_i, y_i, z_i), the error is set as follows:

$$e_i = x_i^2 + ay_i^2 + bz_i^2 + cx_i + dy_i + ez_i + f \quad (20)$$

Then the sum of squares of errors is:

$$E = \sum_{i=1}^N e_i^2 \quad (21)$$

Let E be equal to 0 for the first order partial derivatives of (a, b, c, d, e, f) . The estimates of (a, b, c, d, e, f) can be obtained by solving a six-variable equations. Then we get (x_0, y_0, z_0, A, B, C) .

The second step is Plane Fitting based on RANSAC.

In plane fitting, there are some points with large deviation in the selected points. Direct plane fitting will make the plane deviate from the real one. Therefore, the robustness of fitting can be improved by using the RANSAC method. Firstly, three points are randomly extracted from all points for plane fitting. The coordinates of three points are known as $\mathbf{P}_1(x_1, y_1, z_1)$, $\mathbf{P}_2(x_2, y_2, z_2)$, $\mathbf{P}_3(x_3, y_3, z_3)$. So we can set the equation as follows:

$$A'(x - x_1) + B'(y - y_1) + C'(z - z_1) = 0 \quad (22)$$

Thus:

$$A' = (y_3 - y_1)(z_3 - z_1) - (z_2 - z_1)(y_3 - y_1) \quad (23)$$

$$B' = (x_3 - x_1)(z_2 - z_1) - (x_2 - x_1)(z_3 - z_1) \quad (24)$$

$$C' = (x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1) \quad (25)$$

Then the distance d_i from other points to the plane is calculated. When d_i is less than ε , it is set as an interior point in the plane. Otherwise it is an exterior point. Iteration I times, so that the plane with the most interior points is the fitting plane. The intersection of the plane and the ellipsoid is the ellipse.

$$\begin{cases} A'(x - x_1) + B'(y - y_1) + C'(z - z_1) = 0 \\ \left(\frac{x-x_0}{A}\right)^2 + \left(\frac{y-y_0}{B}\right)^2 + \left(\frac{z-z_0}{C}\right)^2 = 1 \end{cases} \quad (26)$$

The parametric equation of ellipse can be obtained by using Mathematica.

3 Experiments

Based on the method proposed in this paper, we use a binocular acquisition system to carry out experiments. Firstly, the images of left and right cameras are captured by binocular vision, and then the target positions of left and right cameras are captured by SSD respectively. ORB feature points are extracted and matched in the target region. The experimental results are shown in Fig. 3. The 3D coordinates of feature points are calculated by binocular parameters. The 3D coordinates corresponding to the z-coordinate median in the 3D coordinates are used as the 3D coordinates of the target.

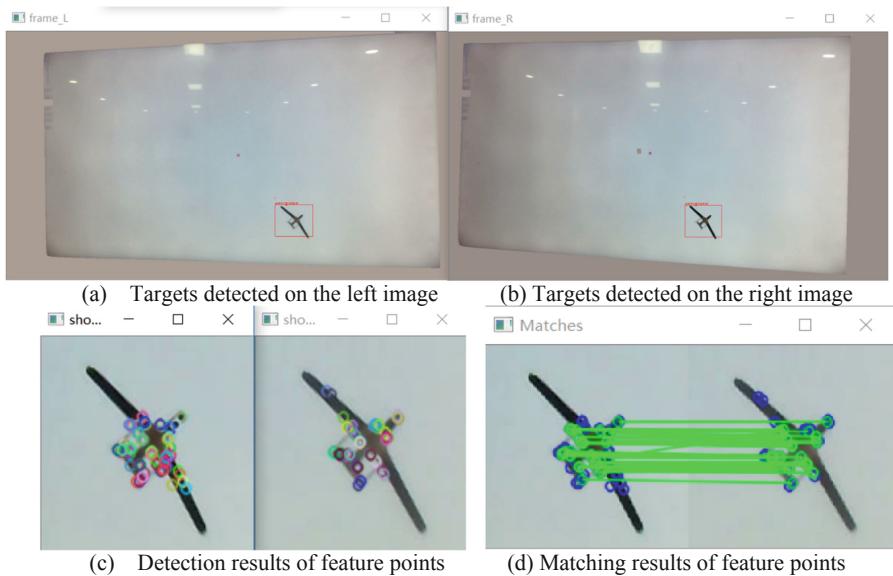


Fig. 3. Intermediate results of binocular stereovision

Using the above method, a total of 901 3D coordinates of the target as shown in Fig. 4(a) are obtained. Figure 4(b) is the result of UKF. In UKF, the sampling time T is 0.1 s. \mathbf{Q} , \mathbf{R} and \mathbf{P} are set as follows:

$$\mathbf{Q} = \text{diag}([1 \ 1 \ 1 \ 0.1 \ 0.1 \ 0.1 \ 0.01 \ 0.01 \ 0.01])$$

$$\mathbf{R} = \text{diag}([40 \ 40 \ 40])$$

$$\mathbf{P} = \text{diag}([2 \ 2 \ 2 \ 0.2 \ 0.2 \ 0.2 \ 0.02 \ 0.02 \ 0.02])$$

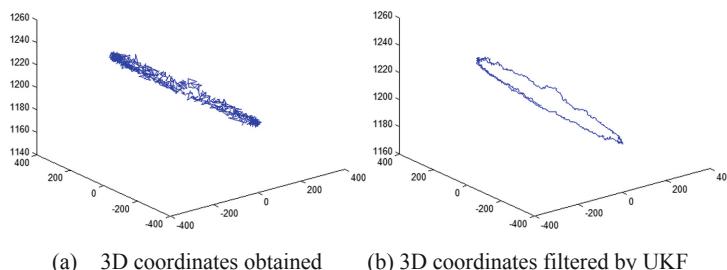


Fig. 4. UKF results

Figure 5 is the result of ellipse fitting using 400 points. When the number of points is less than 400, the fitting is easy to fail. In ellipsoid fitting, 40 points are selected for least-square ellipsoid fitting. In plane fitting, the number of RANSAC is set to 10000.

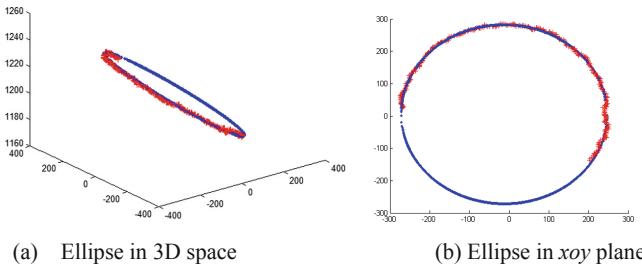


Fig. 5. Ellipse fitting results

4 Conclusion

This paper presents a position tracking and motion trajectory fitting method for UAV. This method first obtains the target's 3D coordinates by binocular method, and then filters the target's 3D coordinates by using UKF. After accumulating to a certain historical position, the ellipse in 3D space is fitted according to the track points of the target, and then the trajectory of the target is determined. When fitting ellipse, we first fit ellipsoid, and then use the RANSAC method to fit plane. This method can help users detect the target, and then track the target in 3D space. And the intention of the target can be predicted according to the trajectory of the target. This method guarantees both speed and effectiveness. However, the current method still has some shortcomings, such as only fitting elliptic trajectory, only applicable to single target tracking and so on. Later, we will improve on these shortcomings.

References

1. Wilson DB, Göktoğan AH, Sukkarieh S (2015) Experimental validation of a drogue estimation algorithm for autonomous aerial refueling. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp 5318–5323
2. Portmann J, Lynen S, Chli M et al (2014) People detection and tracking from aerial thermal views. In: IEEE international conference on robotics and automation (ICRA), Hong Kong, China, pp 1794–1800
3. Junmin L, Jing W, Wentian Z et al (2013) Robot pose estimation and accuracy analysis based on stereo vision. In: IEEE international conference on mobile ad-hoc and sensor networks, pp 555–559
4. Hui TW, Chung R (2013) Determining motion directly from normal flows upon the use of a spherical eye platform. In: IEEE International Conference on Computer Vision and Pattern Recognition, pp 2267–2274

5. Auger F, Hilairet M, Guerrero JM et al (2013) Industrial applications of the Kalman filter: a review. *IEEE Trans Ind Electron* 60(12):5458
6. Cheng S-Y, Zhang J-Y (2008) Review on particle filters. *J Astronaut* 29(4):1099–1111
7. Julier SJ, Uhlmann JK (1996) A general method for approximating nonlinear transformations of probability distributions. Technical report, Robotics Research Group, Department of Engineering Science, University of Oxford
8. Xu L, Oja E, Kultanen P (1990) A new curve detection method: randomized Hough transform (RHT). *Pattern Recogn Lett* 11(5):331–338
9. Heinz DC, Chang CI (2001) Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans Geosci Remote Sensing* 39(3):529–545
10. Li Q, Griffiths JG (2004) Least squares ellipsoid specific fitting. In: IEEE geometric modeling and processing, pp 335–340
11. Gallo O, Manduchi R, Rafii A (2011) CC-RANSAC: fitting planes in the presence of multiple surfaces in range data. *Pattern Recogn Lett* 32(3):403–410



A Novel Stability Analysis Method for Large-Scale Photovoltaic System

Xiaofan Gao^(✉), Shu Tian, Mengtian Zhang, and Yunlong Li

Henan Polytechnic University, Jiaozuo 454000, China
783712620@qq.com

Abstract. According to the interaction between the inverter and the grid in the large-scale photovoltaic grid-connected system, an auxiliary system based on grid connection of a single inverter is introduced. Research shows that the stability of the inverter output current in large-scale photovoltaic grid-connected systems depends on the stability of the auxiliary system and the stability of the inverter itself. But the stability of the total grid-connected current depends only on the stability of the auxiliary system. Therefore, a method for stability analysis of large-scale photovoltaic grid-connected systems is proposed. For a large-scale photovoltaic grid-connected system that is connected to the ideal grid through the grid impedance Z_g after the inverters of the same specification and model are connected in parallel, the stable determination conditions are: (1) The inverter itself is stable, that is, a single inverter is stable when it is directly connected to the ideal power grid; (2) A single inverter in a large-scale PV grid-connected system is stable through the grid impedance nZ_g to the ideal grid. The stability analysis method can transform the research on the stability of large-scale photovoltaic grid-connected systems into the stability of the grid-connected system of a single inverter. For the multi-inverter parallel grid-connected system with uniform specifications and models, the stability analysis method is applicable regardless of whether all inverter current loop input reference currents are identical. The simulation results verify the correctness of the theoretical analysis.

Keywords: Auxiliary system · Grid impedance · Stability analysis

1 Introduction

China's photovoltaic power generation is dominated by large-scale photovoltaic power plants, which account for 84% of the total installed capacity of photovoltaic in the country. Large-scale photovoltaic power plants usually need to pass high-voltage long-distance transmission lines to transfer electric energy to the large power grid [1, 2]. The grid impedance introduced by the transmission part causes interaction between different inverters in the system and between the photovoltaic power plants and the grid, which may cause resonance or even instability, threatening the safe and stable operation of large-scale photovoltaic power plants [3, 4]. Studying the stability of large-scale photovoltaic system have great significance for exploring the methods to solve resonance problems.

Compared with single in addition to the coupling relationship between the inverter and the grid in the multi-inverter parallel system, there are coupling relationships between different inverters [5, 6]. Therefore, the research on its stability becomes more complex. It is not possible to directly consider the grid impedance as part of the inverter side filter inductance like in the case of a single inverter system [7, 8]. It is also impossible to directly analyze the stability of large-scale photovoltaic systems by using the impedance stability criterion in a single inverter system [9, 10].

Some documents simplify the multi-inverter parallel grid-connected system through some assumptions, and then study the simplified system. Literature [11], based on the assumption that all inverter current loop input reference currents are identical, by equivalently changing the passive LCL parallel network, it is concluded that the equivalent grid impedance is equal to the product of the number of inverters and the actual grid impedance. Then the influence of grid impedance on multi-inverters parallel grid-connected system is transformed into the impact on single inverters. Literature [12] uses a similar method to study the stability of large-scale photovoltaic grid-connected systems. However, the actual input current of the inverter current loop in the actual project is generated by the power loop of each inverter, which is affected by the maximum power tracking. Therefore, all inverter current loop input reference currents may not be identical, and the conclusions drawn from this research method are not generally applicable.

In order to derive conclusions that are more general applicability, some literatures derive the controlled current expressions of all inverters according to the equivalent mathematical model of the multi-inverter parallel grid-connected system without using the above-mentioned ideal, and then study the stability of all controlled currents. In [13], the output current of each inverter is derived according to the Norton equivalent model of the multi-inverter parallel grid-connected system, and then the stability analysis of the coefficients of each current component contained in the output current is performed. Literature [12, 14] used a similar method to study the stability of large-scale photovoltaic grid-connected systems. Although the analysis methods in these documents can study the stability of multi-inverter parallel grid-connected systems, the process of deriving the output current of all inverters according to the whole system is complicated, and the output current of each inverter is higher order systems, the calculation is more complicated. In addition, the focus of the above literature is mainly on the influence of grid impedance and the number of parallel inverters on the stability of the inverter output current. The stability between the inverter output current and the total grid-connected current is not studied in depth relationship.

Based on the above analysis, this paper introduces an auxiliary system based on single inverter based on the interaction between inverter and grid in large-scale photovoltaic grid-connected system. The stability problem of large-scale photovoltaic grid-connected system is transformed into the auxiliary system and the stability problem of the inverter itself. Then a simple and clear physical analysis method for the stability of large-scale photovoltaic grid-connected system is proposed. The method can convert the stability problem of the multi-inverter parallel grid-connected system into the stability problem of the single-inverter grid-connected. Finally, the theoretical analysis of this paper is verified by simulation.

2 Mathematical Model of Large-Scale Photovoltaic Grid-Connected System

Inverter in large-scale photovoltaic power plants are usually arranged as current sources to be connected to the grid, and the current loops generally adopt a classical double-loop control structure. The inner loop is capacitive current feedback, which realizes active damping to suppress the natural resonance of the LCL filter. The outer loop is the current of the grid connection point, and is fed back to the current controller to realize the control of the grid-connected current to the reference current. In large PV grid-connected power plants, inverters are usually connected in parallel and then connected to the grid through a certain grid impedance. Therefore, the Norton equivalent model of the large-scale photovoltaic grid-connected system is obtained as shown in Fig. 1 [4]. In the figure, $Y_g(s) = 1/Z_g(s)$, which is the admittance corresponding to the grid impedance, referred to as the grid admittance; Z_g is the equivalent grid impedance and is usually inductive.

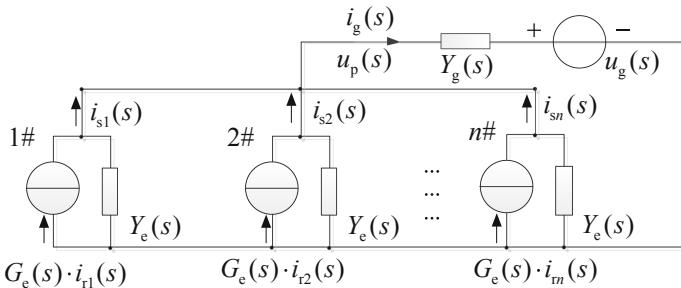


Fig. 1. Equivalent Norton model of LSPV plants

3 Large-Scale Photovoltaic Grid-Connected System Stability Analysis Method

3.1 Inverter Interaction with the Grid

The net current injected into the grid by each inverter is the difference between the current flowing into the grid only under the excitation of the inverter reference current and the current flowing into the inverter only under the excitation of the grid voltage. Therefore, the net current injected into the grid by each inverter can be obtained as shown in Eq. (1).

$$\begin{bmatrix} i_{1g} \\ i_{2g} \\ i_{3g} \\ i_{4g} \end{bmatrix} = \begin{bmatrix} G_{t1} & 0 & \cdots & 0 \\ 0 & G_{t1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & G_{t1} \end{bmatrix} \begin{bmatrix} i_{r1} \\ i_{r2} \\ i_{r3} \\ i_{r4} \end{bmatrix} - \begin{bmatrix} Y_{t1} \\ Y_{t2} \\ Y_{t3} \\ Y_{t4} \end{bmatrix} u_g \quad (1)$$

Where G_{tj} is the action coefficient of the j # inverter to the grid, which can be obtained by calculating the current flowing through the grid impedance only under the reference current excitation of the j # inverter; Y_{tj} represents the coefficient of action of the grid voltage on the j # inverter, which can be obtained by calculating the current flowing to the j # inverter only under the excitation of the grid voltage. Considering that all inverter specifications and models in the system are same, the current coefficient in the equation can be expressed as:

$$G_{t1} = G_{t2} = \dots = G_m = G_t(s) = \frac{Y_g(s)G_e(s)}{mY_e(s) + Y_g(s)} \quad (2)$$

$$Y_{t1} = Y_{t2} = \dots = Y_m = Y_t(s) = \frac{Y_e(s)Y_g(S)}{nY_e(s) + Y_g(S)} \quad (3)$$

Therefore, the net current injected into the grid by the j # inverter can be obtained as shown in the equation.

$$i_{tg}(s) = G_t(s)i_{rj} - Y_t(s)u_g(s) \quad (4)$$

Bring the grid impedance, the inverter closed-loop transfer function $G_e(s)$ and the equivalent output admittance $Y_e(s)$ into the current coefficients $G_t(s)$ and $Y_t(s)$, the following equation can be obtained:

$$G_e = k_{pwm}G_c(s)/M_t \quad (5)$$

$$Y_e = (s_2L_1C + sk_{pwm}k_cC + 1)/M_t \quad (6)$$

Where $M_t = s^3L_1L_2C + sk_{pwm}k_cL_tC + s(L_1 + L_t) + k_{pwm}G_c(s)$, $L_t = L_2 + nL_g$.

Defining the auxiliary system: A single inverter in a large-scale PV grid-connected system is connected to the ideal grid through the grid impedance nZ_g . The Norton equivalent model is shown in Fig. 2.

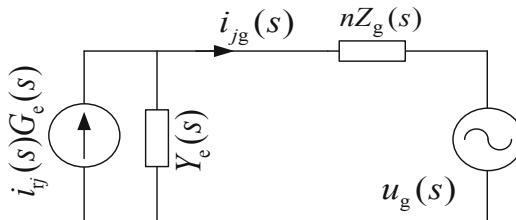


Fig. 2. Auxiliary system equivalent Norton model

For auxiliary systems, the grid impedance nZ_g can be considered as part of the filter side of the inverter network side. Then the equivalent mathematical model of the inverter current loop in s-domain in the auxiliary system can be obtained as shown in

Fig. 3. It can be seen from Eq. (4) and Fig. 3, that the current coefficient $G_t(s)$ is a closed-loop transfer function from the reference current $i_{rj}(s)$ to the output current $i_{rg}(s)$, and $Y_t(s)$ is the equivalent admittance of the inverter from the grid voltage to the output current $i_{rg}(s)$. Therefore, the net current model of each inverter flowing into the grid is exactly the same as the defined auxiliary system. The stability of the current coefficients $G_t(s)$ and $Y_t(s)$ is the same as that of the auxiliary system.

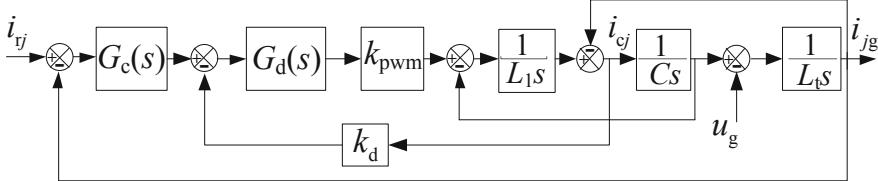


Fig. 3. Inverter current-loop model in auxiliary system

According to the grid-connected current equal to the sum of the net currents of all the inverters injected into the grid, $i_g(s)$ is obtained as shown in the Eq. (7). It can be seen that the stability of the grid-connected current is the same as the stability of the net current injected into the grid by each inverter, that is, depending on the auxiliary system, if the auxiliary system is stable, the grid-connected current is stable, and vice versa.

$$i_g(s) = \sum_{j=1}^n i_{jg}(s) = G_t(s) \sum_{j=1}^n i_{rj}(s) - n Y_t u_g \quad (7)$$

3.2 Interaction Between Different Inverters

According to the Fig. 1, the output current of each inverter can be obtained as shown in Eq. (8), where the non-diagonal elements represent the interaction coefficients between different inverters.

$$\begin{bmatrix} i_{s1} \\ i_{s2} \\ i_{s3} \\ i_{s4} \end{bmatrix} = \begin{bmatrix} G_{11} & -G_{12} & \cdots & -G_{1n} \\ -G_{21} & G_{22} & \cdots & -G_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -G_{n1} & -G_{n2} & \cdots & G_{nn} \end{bmatrix} \begin{bmatrix} i_{r1} \\ i_{r2} \\ i_{r3} \\ i_{r4} \end{bmatrix} - \begin{bmatrix} Y_t \\ Y_t \\ Y_t \\ Y_t \end{bmatrix} u_g \quad (8)$$

Where G_{jm} is the effect coefficient m# inverter on the j# inverter. It can be obtained by calculating the current flowing to the j# inverter only under the m# inverter reference current excitation. Since all inverter specifications are identical, the interaction coefficient between any two inverters is the same, that is, all the non-diagonal elements in the Eq. (8) are the same, which can be expressed as:

$$G_{12} = G_{13} = \cdots = G_{ij} = G_f(s) = \frac{Y_e(s) G_e(s)}{n Y_e(s) + Y_g(s)} \quad (9)$$

Equivalent transformation of $G_f(s)$ can be obtained:

$$G_f(s) = [G_e(s) - G_t(s)]/n \quad (10)$$

Therefore, the net current flowing from $j\#$ inverter to $m\#$ inverter is:

$$\begin{aligned} i_{sjm}(s) &= G_f(s)[i_{rj}(s) - i_{rm}(s)] \\ &= \frac{1}{n}G_e(s)[i_{rj}(s) - i_{rm}(s)] - \frac{1}{n}G_t(s)[i_{rj}(s) - i_{rm}(s)] \end{aligned} \quad (11)$$

It can be seen that the current interaction between different inverters can be decomposed into two parts of the current component that are affected by the inverter itself and affected by the auxiliary system. The currents that interact between different inverters are stable if and only if both currents are stable. If the grid impedance of the system is zero and $G_e(s) = G_t(s)$, all inverters in large PV grid-connected systems are directly connected to the ideal grid, so there is no interaction between different inverters. In addition, if the reference currents of the two inverters are the same, the interaction current between them is zero. Only when there is a difference in the reference currents of the two inverters can there be an interaction between the two inverters.

3.3 Inverter Output Current

The diagonal element in the Eq. (8) represents the influence coefficient of the inverter's own reference current on the inverter output current. For the $j\#$ inverter's own reference current influence coefficient $G_{jj}(s)$, it can be obtained by calculation the current flowing out of the $j\#$ inverter only under the $j\#$ inverter reference excitation. Since all inverter are identical in size and specification, all diagonal elements are the same, which can be derived:

$$G_{11} = \dots = G_{nn} = G_a(s) = G_e(s) - G_f(s) \quad (12)$$

Therefore, the $j\#$ inverter output current can be expressed as:

$$\begin{aligned} i_{sj}(s) &= G_a(s)i_{rj} - G_f(s) \sum_{m=1, j \neq m}^n i_{rm} - Y_t u_g \\ &= \frac{1}{n}[G_e(s) - G_t(s)] \sum_{m=1, j \neq m}^n (i_{rj} - i_{rm}) + G_t(s)i_{rj} - Y_t u_g \end{aligned} \quad (13)$$

It can be seen that the output current of each inverter can be decomposed into two parts: the current that interacts with other inverters and the current that interacts with the grid. Its stability depends on both the stability of the inverter itself and the stability of the auxiliary system. The inverter output current will only show stability when both the inverter itself and the auxiliary system are stable. If the input reference currents of

all inverter current loops are the same and the current in the inverter output current interaction with other inverters is zero, there is only interaction with the grid.

3.4 Large-Scale Photovoltaic Grid-Connected Stability Determination Conditions

In summary, for a large-scale photovoltaic grid-connected system composed of n identical specifications and identical inverters connected in parallel and connected to the ideal grid through the grid impedance Z_g , the stable determination conditions are:

- (1) The inverter itself is stable, a single inverter in large-scale photovoltaic grid-connected system is stable when directly connected to an ideal power grid.
- (2) The auxiliary system is stable, that is, a single inverter in the system is stable when it is connected to the ideal power grid through the resistor Z_g .

According to the stability determination conditions described above, the stability of the large-scale photovoltaic grid-connected system can be studied by analyzing only the stability of the inverter itself and the stability of the auxiliary system in the large-scale photovoltaic grid-connected system. In this way, the research on the stability of the multi-inverter parallel grid-connected system can be transformed into the study of the stability of the grid-connected systems, the increase of grid impedance, the increase of the number of inverters or the decrease of the stability of the inverter itself or the auxiliary system may cause the resonance phenomenon of the system. According to the above analysis, the resonance of the large-scale photovoltaic grid-connected system can be classified into the following three cases:

Case 1: When condition (2) is not satisfied, regardless of whether condition (1) is satisfied, the grid-connected current and the inverter output current will exhibit resonance phenomenon.

Case 2: When condition (2) is satisfied, but condition (1) is not satisfied, the grid-connected current will not show resonance. If all inverters have the same reference current, the output current of all inverters will not show resonance; if all inverters have different reference current, the output current of all inverters will show resonance.

Case 3: When both conditions (1) and (2) are satisfied, the grid-connected current and the inverter output current do not exhibit resonance, regardless of whether the reference currents of all the inverters are identical.

4 Experimental Verification

In order to further verify the theoretical analysis, and due to the limitations of experimental conditions, only two identical 6 kW single-phase inverters are used as an example. The grid impedance is simulated by the winding inductance, $L_g = 0.5 \text{ mH}$.

During the experiment, the 1# inverter reference current is set to 40 A, and the 2# inverter reference current is set to 30 A.

The pole of inverter's own current loop closed-loop transfer function $G_e(s)$ and the auxiliary system current loop closed-loop transfer function $G_f(s)$ varies with k_p as shown in Fig. 4.

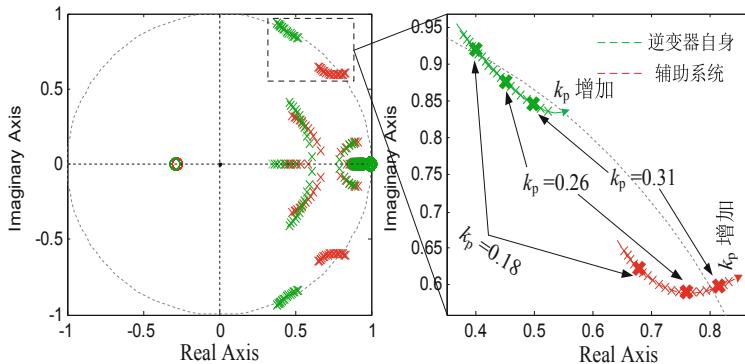


Fig. 4. Pole variation of $G_e(s)$ and $G_f(s)$ with changing

When $k_p = 0.18$, condition (1) is not satisfied, and condition (2) is satisfied, and the experimental waveform is shown in Fig. 5. It can be seen that the output currents of the two inverters all contain a large number of harmonics, and the grid-connected point voltage waveform are good and have no resonance.

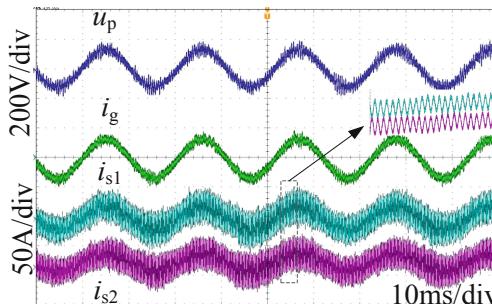


Fig. 5. Experiment waveforms when $k_p = 0.18$

When $k_p = 0.26$, condition (1) and condition (2) are satisfied, and the experimental waveform is shown in Fig. 6. It can be seen that the output current waveforms, grid-connected current waveforms and grid-connected voltage waveforms of the two inverters are all smooth and have no resonance.

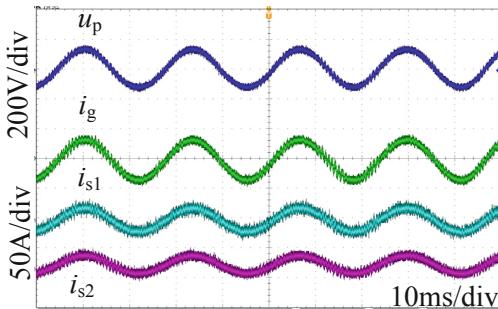


Fig. 6. Experimental waveforms when $k_p = 0.26$

When $k_p = 0.31$, condition (1) is satisfied, and condition (2) is not satisfied, and the experimental waveform is shown in Fig. 7. It can be seen that the output current of the two inverters, the grid-connected current point voltage all contain a large number of resonant harmonics.

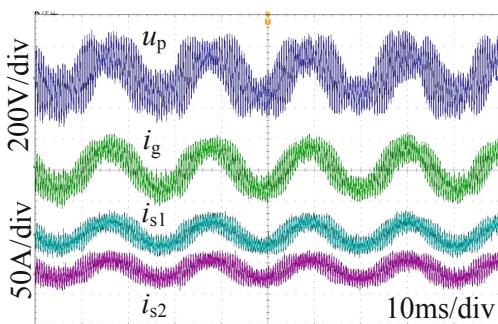


Fig. 7. Experimental waveforms when $k_p = 0.31$

References

1. Li X, Fan J, Cao L et al (2018) Analysis on the adaptability of large-scale grid-connected PV station. Power Syst Prot Control 46(8):164–169
2. Li C, Yang X (2011) Grid-connected technology of solar and wind power, 1st edn. China Water & Power Press, Beijing
3. Yang M, Zhou L, Zhang D et al (2013) Stability analysis of large-scale photovoltaic power plants for the effect of grid impedance. Trans China Electrotech Soc 28(9):214–223
4. Zheng C, Zhou L, Guo K et al (2016) Stability study of large-scale photovoltaic plant containing polytype inverters. In: 2016 IEEE international power electronics and motion control conference (PEMC), Varna, Bulgaria. IEEE, pp 470–475
5. Enslin JHR, Heskes PJM (2004) Harmonic interaction between a large number of distributed power inverters and the distributed networks. IEEE Trans Power Electron 19(6):1586–1593

6. Zhou L, Zhang M (2013) Modeling and stability of the large-scale PV plants due to grid impedance. In: 39th annual conference of the IEEE industrial electronics society, Vienna, Austria. IEEE, pp 1025–1030
7. Liserre M, Teodorescu R, Blaabjerg F (2006) Stability of photovoltaic and wind turbine grid-connected inverters for a large set of grid impedance values. *IEEE Trans Power Electron* 21(1):263–272
8. Zheng Z, Xu H, Ming Y et al (2018) Grid voltage feedforward control strategy for weak grid based on complex filter. *Power Prot Control* 46(24):70–75
9. Sun J (2011) Impedance-based stability criterion for grid-connected inverters. *IEEE Trans Power Electron* 26(11):3075–3078
10. Yang D, Ruan X, Wu H (2014) Impedance shaping of the grid-connected inverter with LCL filter to improve its adaptability to the weak grid condition. *IEEE Trans Power Electron* 29(11):5795–5805
11. Agorreta JL, Borrega M, Lopez J et al (2011) Modeling and control of N-paralleled grid-connected inverters with LCL filter coupled due to grid impedance in PV plants. *IEEE Trans Power Electron* 26(3):770–785
12. Zhou L, Zhang M (2013) Modeling and stability of large-scale PV plants due to grid impedance. In: 39th annual conference of the IEEE industrial electronics society, Vienna, Austria. IEEE, pp 1025–1030
13. He J, Li YW, Bosnjak D et al (2013) Investigation and active damping of multiple resonances in a parallel-inverter-based microgrid. *IEEE Trans Power Electron* 28(1):234–246
14. Zhou L, Zhang M, Ju X et al (2013) Stability analysis of large-scale photovoltaic plants due to grid impedances. *Proc CSEE* 33(34):34–41



Simulation Analysis of Nonlinear Characteristics in Complex System

Jiaomei Zhou¹, Fengshun Lu², Sumei Xiao¹, Long Qi^{2(✉)},
and Yufei Pang²

¹ School of Manufacturing Science and Engineering, Southwest University of Science and Technology, Mianyang 621010, China

² Computational Aerodynamics Institute, China Aerodynamics Research and Development Center, Mianyang 621010, China
qilong1226@163.com

Abstract. In order to investigate the nonlinear systems, we introduce different nonlinear characteristics into a given linear unit feedback system and utilize MATLAB/Simulink toolbox to simulate the output response of the step signal of the system. Simulation results show that the introduction of nonlinear links affects the stability of the control system. The dead zone nonlinear link decreases the equivalent gain of the system under the action of large signal, and the system loses the closed-loop control under the action of deep saturation. Saturation nonlinear link can make the system tend to be stable, but it reduces the amplification factor and the steady-state accuracy. Saturation nonlinear link can make the system tend to be stable, but it reduces the amplification factor and the steady-state accuracy. These nonlinear characteristics can also be used in the analyzing of the nonlinear characteristics in the multidisciplinary coupled software system or other complex systems.

Keywords: Nonlinear characteristics · MATLAB/Simulink simulation · Dead zone · Saturation · Backlash · Multidisciplinary coupled software

1 Introduction

With the continuous development of science and technology, the analysis of actual production process becomes more and more precise. The relevant experimental results demonstrate that arbitrary actual physical system is nonlinear. Nonlinearity is an inherent property of physical systems, and all control systems have some degree of nonlinearity [1, 2]. More and more attention has been paid to the nonlinear system and its control. The nonlinear system has become one of the research hotspots in the field of control engineering. It is well acknowledged that there is none general method to analyze and study the nonlinear systems. Each method can only be applied to a certain kind of problems. Besides, some linear methods are adopted to simplify the analysis process for certain control systems. However, when the system contains essentially nonlinear characteristics, it cannot be simplified by linearization [3]. Therefore, it is necessary to analyze and investigate the nonlinear characteristics for the nonlinear system with essential nonlinear characteristics.

In this paper, the typical nonlinear characteristics of nonlinear systems are simulated and analyzed. And stating the function of it in the mathematical modeling of Multidisciplinary coupled computing software systems or other complex systems.

2 Typical Nonlinear Characteristics of Nonlinear Systems

A system with nonlinear elements or links is called a nonlinear system. There many are types of nonlinear characteristics, such as saturation, dead zone, hit crossing, backlash, relay, etc. In this section, we give a brief introduction to several typical static characteristics.

2.1 Saturation Nonlinear Characteristic

Figure 1 presents the saturation nonlinearity that is caused by saturation. When the input increases to a certain value and the output does not change, this phenomenon is called saturation. Considering the inherent property of the amplifier element or the actuator element, the output of the system cannot be infinitely increased and arbitrary actual control system is subject to the saturation nonlinearity [4, 5].

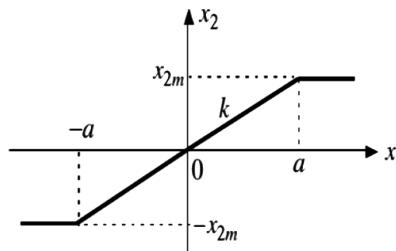


Fig. 1. Saturation nonlinear characteristic

The ideal saturated nonlinear characteristic can be described by mathematical model (1) as

$$x_2 = \begin{cases} -ka & x_1 \leq -a \\ kx_1 & |x_1| \leq a \\ ka & x_1 \geq a \end{cases} \quad (1)$$

where x_1 is the input, x_2 is the output, and k is the proportionality coefficient.

Saturated nonlinearity usually promotes system stability and reduced amplification factor. Sometimes the saturated nonlinearity is utilized to narrow the linear region of a property.

2.2 Dead Zone Nonlinear Characteristic

Dead zone nonlinearity is shown in Fig. 2. Dead zone nonlinearity is a common nonlinearity characteristic, which is usually caused by the insensitive range of sensors, amplifiers and actuators [6].

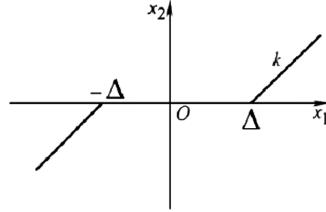


Fig. 2. Dead zone nonlinear characteristic

The ideal dead zone nonlinear characteristics can be described by mathematical model (2) as

$$x_2 = \begin{cases} 0 & |x_1| \leq \Delta \\ k(x_1 - \Delta \operatorname{sgn} x_1) & |x_1| > \Delta \end{cases} \quad (2)$$

where sgn is a symbol function and $\operatorname{sgn} x_1 = \begin{cases} 1 & x_1 > 0 \\ -1 & x_1 < 0 \end{cases}$.

When the absolute value of the input is less than a certain value, the output is zero. At this time, the system is in an open loop state and loses its regulating function. Consequently, the control sensitivity decreases and the steady-state error increases, which brings adverse effects to the control system. When there is a small signal disturbance at the input end of the system, the dead zone can reduce the disturbance signal and bring beneficial influence to the control system. Some systems employ the dead zone nonlinearity so as to improve their anti-jamming capability.

2.3 Backlash Nonlinear Characteristic

Backlash is shown in Fig. 3. The reason of backlash nonlinearity is usually caused by hysteresis, which is similar to the hysteresis of linear system. Gear transmission is a typical backlash nonlinearity [7].

The nonlinear characteristics of backlash can be described by mathematical model (3) as

$$x_2 = \begin{cases} k(x_1 - a \operatorname{sgn} \dot{x}_1) & \dot{x}_1 \neq 0, \dot{x}_2 \neq 0 \\ x_{2m} \operatorname{sgn} x_1 & \dot{x}_2 = 0 \end{cases} \quad (3)$$

Backlash is always harmful for control systems, since it causes instability or self-excited oscillation of the system, which is not good for the stability of the system. Therefore, its influence should be eliminated or weakened.

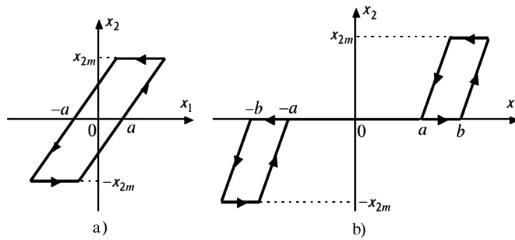


Fig. 3. Backlash nonlinear characteristic

3 Simulations of Typical Nonlinear Characteristics

In the unit negative feedback system as shown in Fig. 4, different nonlinear links are introduced to carry out MATLAB/Simulink simulations [8, 9]. A step response of the system can be observed, and the effects of different nonlinear links on the system performance are analyzed and compared. The unit step signal serves as the input signal to the whole system.

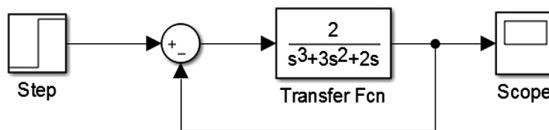


Fig. 4. Given unit negative feedback system

The step response curve of the unit feedback system is shown in Fig. 5.

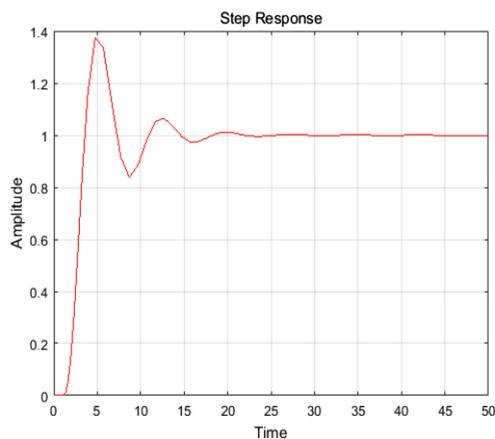


Fig. 5. The system step response curve without nonlinear element

3.1 Saturated Nonlinear Simulation Analysis

The saturated nonlinear element is introduced into the aforementioned unit feedback system, whose block diagram is shown in Fig. 6. The step response curve after adding saturated nonlinear element is shown in Fig. 7. The saturation range is from -0.1 to 0.1 .

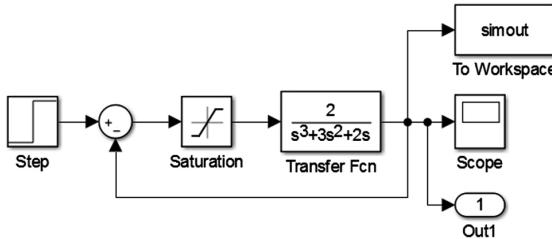


Fig. 6. System block diagram after adding saturation

From Figs. 6 and 7 we can observe that after the addition of saturated nonlinear link, the rise time of the system increases obviously, while the adjustment time has no obvious change. Note that the overdose and the peak time both decrease [10, 11].

Now the upper and lower output limits of different saturated nonlinear links are set to compare and analyze the performance impact of different output ranges on the system. Figure 8 presents the step response curves of the system at different saturation intervals, namely 0.05, 0.1, 0.2, and 0.5.

When the output range of the saturated nonlinear element is narrow, the step response speed of the system is slow and the rise time is long. The overshoot is small and the oscillation is not obvious. With the increase of the output range, the response speed of the system increases, the rise time decreases obviously, and the oscillation is also obvious. This is because the saturated nonlinear element will limit the signal. It can also be observed that the stronger the limiting effect is, the less likely the output of the system will produce overshoot.

3.2 Dead Zone Nonlinear Simulation Analysis

A dead zone nonlinear element is introduced into a given unit feedback system [6, 9], and the block diagram of the system is shown in Fig. 9. The step response curve after adding the dead zone nonlinear element is shown in Fig. 10.

By comparing Figs. 6 and 11, we can see that when the dead zone range is set to be $[-0.1, 0.1]$, there is no significant change in the dynamic performance index of the system when it is not added. Figure 11 demonstrates step response curves when different dead zone intervals are added to the system, namely 0.2, 0.5, 1.0, and 2.0.

Figure 11 shows that as the dead zone range increases gradually, the time for the system to start responding to the step input signal is also gradually delayed. This is because the dead zone link will automatically ignore the input in the dead zone, so that the response of the system becomes slow.

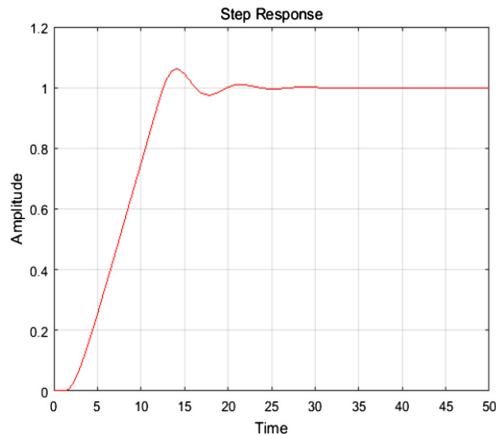


Fig. 7. Step response curve after adding saturated nonlinear element

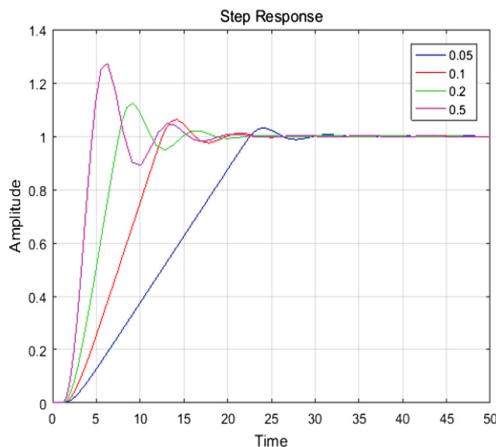


Fig. 8. Step response curves of the system when different saturation are added

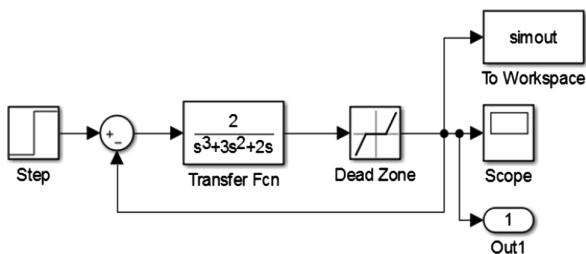


Fig. 9. System block diagram after adding dead zone

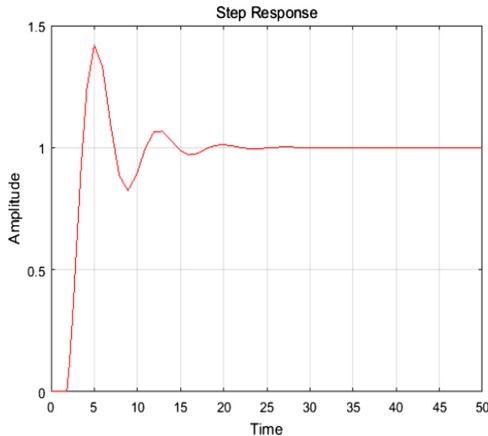


Fig. 10. Step response curve after adding dead zone nonlinear element

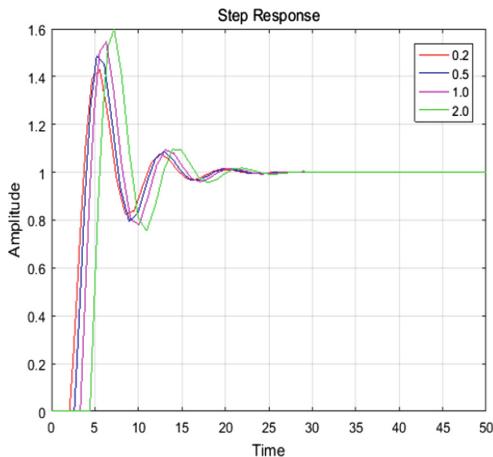


Fig. 11. Step response curves of the system when different dead zone are added

3.3 Backlash Simulation Analysis

A backlash link is introduced into a given unit feedback system, and the system block diagram is shown in Fig. 12. Changing the hysteresis width of the backlash nonlinear link, the relevant step response curve of the system is shown in Fig. 13. When the hysteresis width is zero, this is equivalent to Figs. 5. In Fig. 13, the numbers in the legend represent different hysteresis width [9].

From Fig. 13, it can be seen that with the increase of hysteresis ring width, the oscillation of the system intensifies and the system becomes more and more unstable. Therefore, the introduction of the backlash nonlinear link will cause the violent oscillation of the system [12].

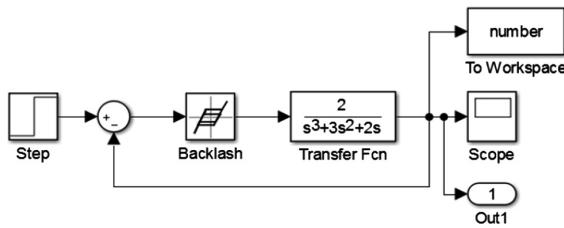


Fig. 12. System block diagram after adding backlash

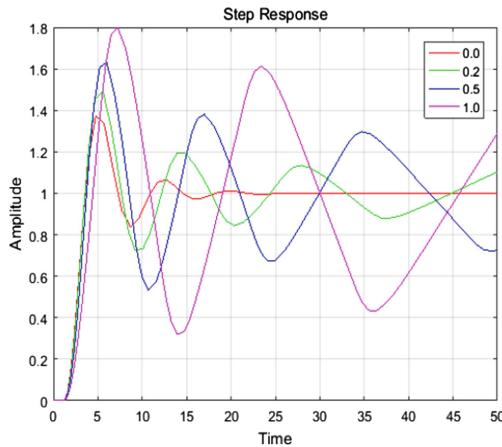


Fig. 13. Step response curves of the system when different backlash are added

4 Conclusion

By analyzing the experimental results from the aforementioned simulations, we observe that the nonlinear link has the following influence on the stability performance of the control system. First, when there are saturated nonlinear links in the system, the response of the system is slow, but the overshoot of the system decreases. Then, the dead zone nonlinear link has no response to the input signal in a small range near zero, but the output is linear to the input when the input exceeds this region. Finally, the backlash nonlinear link will cause the oscillation of the system and make the system become unstable.

For multidisciplinary coupled computing software systems or other complex practical systems [13], it is very difficult to establish the mathematical model according to the physical mechanism and then analyze the restrictive conditions from the mathematical perspective to ensure the stability of the numerical solution process. We first treat the entire system as a black box model, then use the nonlinear characteristics described above to simulate and approximate the nonlinear characteristics existing in multidisciplinary coupled software or other complex systems, and finally transform the black box model to the gray box model gradually. After that, combined with the

simulation analysis of other characteristics of the multidisciplinary coupled software system, the gray box model can be further transformed into a white box model. We use the feedback control on the obtained model so that the floating range of the control variable, advance time step [14], can always ensure that the numerical solution process is carried out stable.

The future work includes the application of the aforementioned nonlinear characteristics on the modeling of multidisciplinary coupled software.

References

1. Franklin GF, Powell JD, Emami-Naeini A (2007) Feedback control of dynamic systems. Posts & Telecom Press, Beijing
2. Dorf RC (2015) Modern control system. Posts & Telecom Press, Beijing
3. Yang G (2008) Research of nonlinear control and controlling strategies. *J Huainan Normal Univ* 10(05):37–38, 145
4. Costa MH (2017) Theoretical transient analysis of a hearing aid feedback canceller with a saturation type nonlinearity in the direct path. *Comput Biol Med* 8(91):243–259
5. Wang N (2014) Research on saturated nonlinear system control. South China University of Technology
6. Xie Y, Lin D, Bian X (2016) Uncertain chaotic system control with dead zone input. *J Chongqing Univ Technol (Nat Sci)* 30(04):120–126
7. Chen H, Yan S, Zhi Y, Yu F (2017) Simulation and dynamic response analysis of gear nonlinear system. *J Mech Transm* 41(08):19–22
8. Fei M, Lin H, Xing W, Li Y (2017) Simulation study on stability of linear control system based on MATLAB. *Sci Technol Inf* 15(30):34–36
9. Wang Z, Lin W (2017) MATLAB/Simulink and control system simulation. Publishing House of Electronics Industry
10. Guangjie X (1987) Application of law criterion in nonlinear system analysis. *Autom Refin Chem Ind* 13(03):17–20
11. Cheng J, Xie J, Wei W (2016) Calculation and simulation of multi-parameter nonlinear mechanical power system. *Modern Manuf Eng* 53(01):20–26
12. Yan J, Liu C, Zhang Y (2011) Design and application of nonlinear vibration system analysis software. *Mech Electr Inf* 10(9):104–105
13. Hu X, He X, Tao J (2008) Cognitive simulation: a new approach to modeling complex systems? *Sci Technol Guide* 36(12):46–54
14. Anderson JD (2012) Computational fluid dynamics. China Machine Press, Beijing



Scene Classification of Remote Sensing Images Based on ResNet

Lin Li^(✉) and Da Li

School of Optoelectronic Information and Computer Engineering,
University of Shanghai for Science and Technology, Shanghai 200093, China
lilin0211@163.com

Abstract. To improve the classification accuracy in remote sensing image processing, this paper proposes a remote sensing image classification method based on depth residual neural network. Because the deep learning model itself needs a lot of training data, the traditional method of rotation and clipping is needed to expand the image data. A large number of manual operations, the efficiency improvement is not obvious, so this paper draws lessons from the idea of migration learning, pretrains on large-scale data set ImageNet, gets the inverted initial model, then trains on UCM_LandUse_21 with the pretrained model, and optimizes the training strategy to obtain the ideal model. The experimental results show that compared with the traditional classification method, the proposed method improves the classification accuracy more obviously, reaching 97.87%, thus verifying the effectiveness of the method.

Keywords: Remote sensing image · Scene classification · Residual network · Transfer learning

1 Introduction

With the rapid development of remote sensing satellite and aircraft earth observation technology, the acquisition of remote sensing image data has become more and more easy. At the same time, remote sensing image contains abundant spatial and semantic information, such as hyperspectral remote sensing image and synthetic aperture radar image [1–3]. Therefore, how to effectively and efficiently process these images becomes more important. Remote sensing image scene classification is one of the most important tasks. Since the accuracy of classification largely depends on the effectiveness of feature extraction, it is a key problem to extract high-quality features for high-resolution remote sensing images. In recent years, the method of deep learning [4] has developed rapidly, and its effect is obvious. It has been introduced into the field of image processing, such as image processing, Classification [5], target recognition [6], and so on. Deep learning convolution neural network, because of its deep non-linear structure, determines its modeling ability in complex classification problems. The reason why it adopts hierarchical structure is inspired by the information processing mechanism of visual system, which was proposed by Deng et al. [7]. At present, many researchers have introduced the deep learning method into the classification of remote sensing images. For example, Wang et al. [8] proposed an improved remote sensing

image classification method based on deep learning, which can improve the classification accuracy through multi-feature fusion. In addition, Zhang et al. [9] proposed a classification method based on the combination of deep learning features and support vector machines, which improved the classification to a certain extent. Precision, but the sample size is small, only CNN network with fewer layers can be used. Luo et al. [11] designed a method of combining PCA and DCNN, But lack of innovation. In addition, Liu et al. [10] made some progress in integrating traditional image processing methods into CNN. But the pretreatment process is complex and the generalization ability is limited.

The reason why the ResNet model trained by ImageNet data set can successfully migrate to the scene classification of high resolution remote sensing image is that the daily visible image is similar to the remote sensing image itself in the basic visual module (edge, angle, etc.), so ResNet has a certain generalization ability in the migration process, so this paper uses the residual of pre-training on the ImageNet data set. The network (ResNet-152) migrates the model to UCMerecd_LandUse21 data set. By adjusting and optimizing the network parameters, the classification accuracy can be improved.

2 Convolutional Neural Network

Convolutional neural network (CNN) is a multi-layer neural network and a classical deep learning model. It is proposed by Lecun [12]. Its basic components include input layer, convolution layer, pooling layer, full connection layer and output layer. The network learns the high-level features of the input image by extracting features layer by layer, and then classifies them by input classifier. CNN reduces parameters by sharing weights. Figure 1 is a basic structure model of CNN.

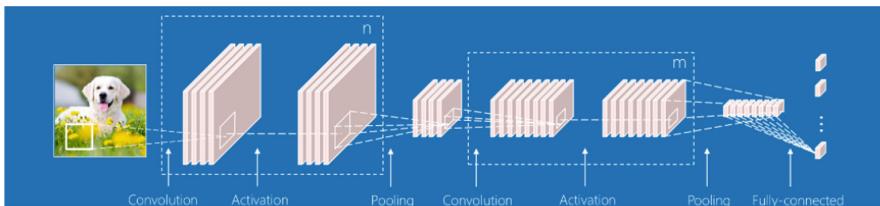
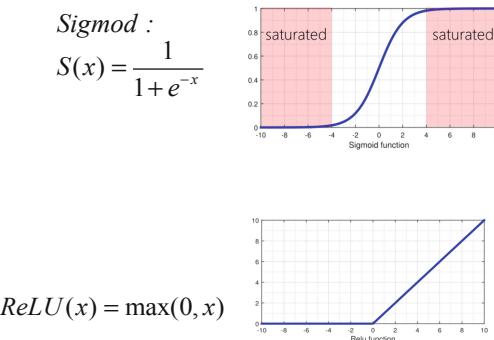


Fig. 1. A schematic diagram of CNN

- (1) Convolution layer: Convolution layer is the key of network. Its function is to extract features from the input of the front layer. The relevant formulas are as follows:

$$y_i = b_i + \sum_i W_{ij} \otimes x_i$$

The commonly used activation functions are Sigmoid, ReLu, etc. as shown in the following Fig. 2.

**Fig. 2.** Activation function

- (2) Pooling layer: Generally, there are two ways of pooling, maximum pooling and average pooling.

Maxpooling

$$\text{max-pooling}(f[i-1], f[i], f[i+1]) = \max(f[i-1], f[i], f[i+1])$$

Averagepooling

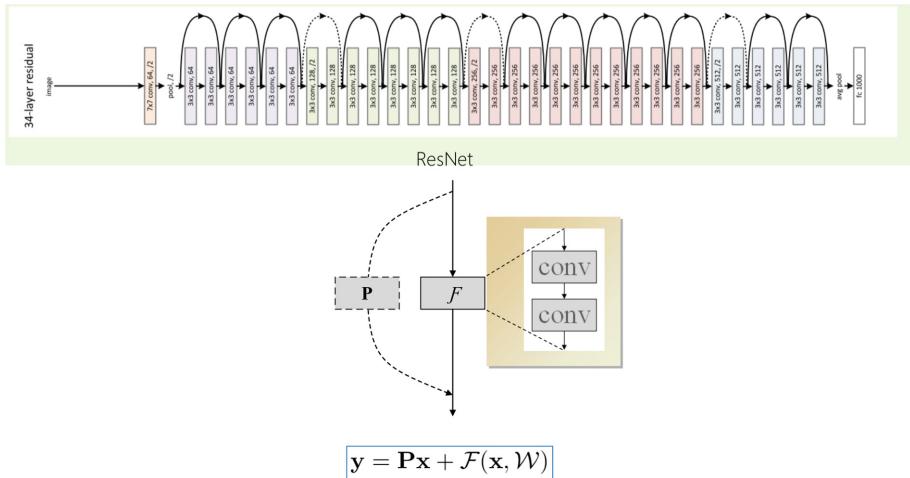
$$\text{Ave-pooling}(f[i-1], f[i], f[i+1]) = \frac{1}{3}(f[i-1], f[i], f[i+1])$$

The pooling layer is mainly used to solve the problem of data redundancy. Using the pooling layer to downsample the feature map is also conducive to improving the network's computing ability.

The training process of convolution neural network can be summarized as follows: training a neural network is actually training the convolution core of each convolution layer so that the convolution core can detect the specific features of the image. In addition, training CNN has two main processes: forward propagation and backward propagation. See Wang [8].

3 Remote Sensing Image Classification Method Based on ResNet

ResNet was proposed by four Chinese such as He [13] of Microsoft Research Institute. Through the use of ResNet Unit, the 152-layer neural network was successfully trained and won the championship on ILSVRC-2015. Its parameter quantity is lower than that of VGGNet. Prominently, the structure of ResNet can speed up the training of neural networks, and its promotion performance is also good. The following figure is a network model and learning module of ResNet (Fig. 3):

**Fig. 3.** Network model

It can be seen from the figure that input \mathbf{x} can be input across layers into the next layer and fused with $\mathcal{F}(\mathbf{x})$ as input of the next residual module. This design idea is mainly aimed at the loss and loss of information in traditional convolutional network or fully connected network when transmitting information. At the same time, there are problems of gradient disappearance and explosion, which will lead to the failure of deep network training. ResNet solves this problem to a certain extent. It protects the integrity of information by directly bypassing information to output. The whole network only needs to learn the difference between input and output, which simplifies the learning goal and difficulty.

3.1 Network Structure

Two kinds of residual modules are used in ResNet network structure. One is to connect two $3 * 3$ convolutions in series as a residual module, the other is to connect three convolution networks of $1 * 1$, $3 * 3$ and $1 * 1$ in series as a residual module. As shown in Fig. 4.

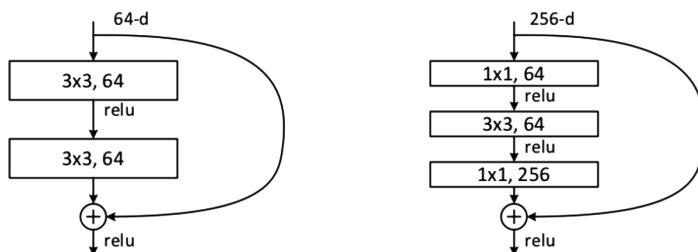
**Fig. 4.** Network structure

Figure 4 ResNet Residual Learning Module with Two Layers and Three Layers In addition, ResNet has different network layers, commonly used are 50-layer, 101-layer and 152-layer, among which Li [14] adopts the first model. Figure 5 is the parameter information of the three models.

| layer name | output size | 18-layer | 34-layer | 50-layer | 101-layer | 152-layer |
|------------|-------------|---|---|---|--|--|
| conv1 | 112×112 | | | 7×7, 64, stride 2 | | |
| conv2.x | 56×56 | $\left[\begin{array}{l} 3 \times 3, 64 \\ 3 \times 3, 64 \end{array} \right] \times 2$ | $\left[\begin{array}{l} 3 \times 3, 64 \\ 3 \times 3, 64 \end{array} \right] \times 3$ | $\left[\begin{array}{l} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{array} \right] \times 3$ | $\left[\begin{array}{l} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{array} \right] \times 3$ | $\left[\begin{array}{l} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{array} \right] \times 3$ |
| conv3.x | 28×28 | $\left[\begin{array}{l} 3 \times 3, 128 \\ 3 \times 3, 128 \end{array} \right] \times 2$ | $\left[\begin{array}{l} 3 \times 3, 128 \\ 3 \times 3, 128 \end{array} \right] \times 4$ | $\left[\begin{array}{l} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{array} \right] \times 4$ | $\left[\begin{array}{l} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{array} \right] \times 4$ | $\left[\begin{array}{l} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{array} \right] \times 8$ |
| conv4.x | 14×14 | $\left[\begin{array}{l} 3 \times 3, 256 \\ 3 \times 3, 256 \end{array} \right] \times 2$ | $\left[\begin{array}{l} 3 \times 3, 256 \\ 3 \times 3, 256 \end{array} \right] \times 6$ | $\left[\begin{array}{l} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{array} \right] \times 6$ | $\left[\begin{array}{l} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{array} \right] \times 23$ | $\left[\begin{array}{l} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{array} \right] \times 36$ |
| conv5.x | 7×7 | $\left[\begin{array}{l} 3 \times 3, 512 \\ 3 \times 3, 512 \end{array} \right] \times 2$ | $\left[\begin{array}{l} 3 \times 3, 512 \\ 3 \times 3, 512 \end{array} \right] \times 3$ | $\left[\begin{array}{l} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{array} \right] \times 3$ | $\left[\begin{array}{l} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{array} \right] \times 3$ | $\left[\begin{array}{l} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{array} \right] \times 3$ |
| | 1×1 | | | average pool, 1000-d fc, softmax | | |
| FLOPs | | 1.8×10^9 | 3.6×10^9 | 3.8×10^9 | 7.6×10^9 | 11.3×10^9 |

Fig. 5. Several common models of ResNet

4 Experiment

4.1 Experimental Data Set

The UCM_LandUse_21 data set [15] is a U.S. land use classification data set. The data set contains 21 types of land use scenarios, each of which has a total of 100 scene maps of $256 \times 256 \times 3$ size. We randomly extract 80 images from each type of remote sensing image as a training set, the remaining 20 as a test set. Some examples are given in Fig. 6.

4.2 Experimental Results and Analysis

Some parameters of the model obtained by the experiment are as follows:

The loss function is based on cross-entropy.

losscriterion = nn.CrossEntropyLoss().cuda()

Optimizer adoption:adam

optimizer = torch.optim.Adam(model.parameters(), lr = lr, betas = (0.9, 0.999), weight_decay = 0.0002)

Using soft Max classifier

In addition, we add two full-connection layers to the existing RESNET network to solve the problem of this experiment.

x = self.fc1(x)

x = self.fc2(x)



Fig. 6. 21 scene graph

Figure 7 shows the correct rate curve and the confusion matrix diagram of the classification.

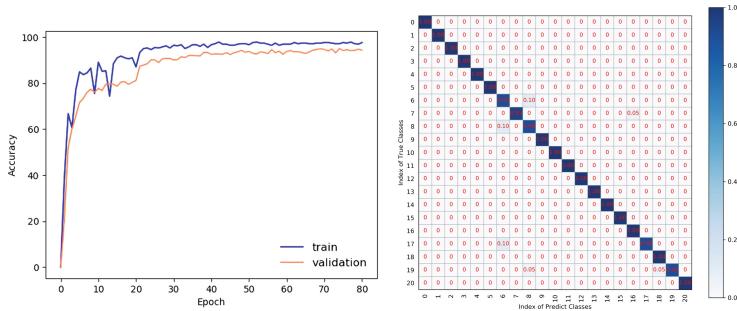


Fig. 7. Accuracy curves and confusion matrices

In this paper, confusion matrix is used to show the accuracy of single-category classification of the model. The horizontal axis represents the label value of each type of remote sensing image, the vertical axis represents the true label value of each type of remote sensing image, and the deeper the color, the higher the accuracy, and the other positions are the corresponding misclassification rate. From the image, we can see the general accuracy. They are all very high, with a low error rate and an average accuracy rate of over 97%. The effect is very obvious.

Table 1 the classification accuracy comparison of different models is given.

Table 1. Average classification accuracy of UCM_LandUse_21 data set by different methods

| Method | Accuracy/% |
|--------------|------------|
| SVM_LDA [14] | 80.33 |
| MS_DCNN [16] | 91.34 |
| GDF_CNN [10] | 94.76 |
| CNN_SVM [8] | 96.43 |
| Desnet [17] | 96.90 |
| Our method | 97.87 |

Table 1 shows that compared with the traditional classification methods, the convolution-based neural network method can improve the classification accuracy, but because of the different network structure, there are also great differences between the deep network models based on convolution-based neural network. For example, the difference between MS_DCNN and this paper is 6%, compared with the paper [17] this paper has a 1% improvement, indicating that the network structure is deeper. It brings about the improvement of accuracy, but also faces the difficulty of further improvement.

5 Conclusion

This paper presents a remote sensing image classification method based on depth residual neural network, which is verified on UCM_Landuse 21 standard data set. The advantages of this paper are: (1) Make full use of the advantages of depth neural network in feature extraction, and use the idea of migration learning to solve the problem of insufficient sample size. (2) Implementing the effective improvement of classification accuracy, (3) Providing some reference for further research. It provides a new idea for remote sensing image scene classification.

References

1. Zhao J (2012) Research on target recognition of remote sensing images with matching technology. University of the Chinese Academy of Sciences, Beijing
2. Ren C (2016) Forest types precise classification and forest resources change monitoring based on medium and high spatial resolution remote sensing images. Chinese Academy of Forestry, Beijing
3. Wang X, Xiong X, Ning C et al (2018) Integration of heterogeneous features for remote sensing scene classification. *J Appl Remote Sensing* 12(1):15–23. <https://doi.org/10.1117/1.jrs.12.015023>
4. Dong L, Zhang J (2018) Research on age and gender classification of face based on in-depth learning and random forest 44(5):246–251
5. Feng Z (2016) Research and application of image feature learning and classification methods based on deep learning. Ph.D. dissertation, South China University of Technology
6. Lin M (2013) Research on face recognition based on deep learning. Master dissertation, Dalian University of Technology
7. Deng J, Dong W, Socher R, et al (2009) ImageNet: a large-scale hierarchical image database. In: Proceedings of the 2009 IEEE conference on computer vision and pattern recognition, Piscataway, NJ, USA. IEEE, pp 248–255
8. Wang X, Li K, Chen N, Huang F. Remote sensing image classification method based on deep convolution neural network and multi-core learning. *J Electron Inf Sci*
9. Ye Q, Xu B, Zhang D (2019) Remote sensing image classification based on in-depth learning feature and support vector machine. *J For Eng* 02:119–125. <https://doi.org/10.13360/j.issn.2096-1359.2019.02.019>
10. Liu Y, Chen L, Cui X, Yuan Z, Xing X (2019) Scene classification of remote sensing images based on spatial feature re-calibration network. *Comput Eng* 1–6
11. Luo C, Wang J, Wang S, Stone, Ren W (2018) High-resolution remote sensing scene classification based on generalized depth migration characteristics. *Syst Eng Electron Technol* 40(03):682–691
12. Lecun Y, Bottou L, Bengio Y et al (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
13. He K, Zhang X, Ren S, et al (2016) Deep residual learning for image recognition. In: Computer vision and pattern recognition. IEEE, pp 770–778
14. Li H, Hu X, Guohang, Wu W (2018) Land use/land cover classification supported by migration learning. *Surv Mapp Bull* 2018(09):50–54, 73

15. Yang Y, Newsam S (2010) Bag-of-visual-words and spatial extensions for land-use classification. In: Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems, GIS 2010. ACM, New York, pp 270–279
16. Xu SH, Mu XD, Zhao P, Ma J et al (2016) Scene classification of remote sensing image based on multi-scale feature and deep neural network. *Acta Geodaetica Cartogr Sin* 45 (7):834–840
17. Li D, Li L, Li X (2018) Remote sensing image classification based on dense nets (10):60–63



A Method of Ball Drop Point Prediction Based on Racket Pose

Fei Yan^(✉) and Chaoli Wang

University of Shanghai for Science and Technology, Shanghai 200000, China
18817901035@163.com

Abstract. The judgment of table tennis falling point has a great influence on the hitting effect. Many athletes judge the falling point of table tennis by the position of the racket (the position and posture of the racket and the posture below) when the opponent hits the ball. This paper proposes to verify the influence of the pose of the racket on the ping-pong track and the drop point based on the visual. First, the ball is subjected to force analysis, a sports model of table tennis is established, and the predicted trajectory and landing point of the ball are obtained. Then, we use the threshold cut method to set multiple thresholds to identify the pose of the racket when hitting the ball. Finally, the binocular vision system is used to obtain the true trajectory and landing point of the ball after hitting the ball. Successfully verified the correctness of our motion model. And through the falling point of the ball, we confirmed that the pose of the racket has an influence on the ball trajectory and the falling point when hitting the ball.

Keywords: Table tennis · Trajectory · Drop point · Binocular vision

1 Introduction

Table tennis robot is an important application of humanoid robot. A conventional table tennis robot system [1, 2] consists of ball position estimation, ball trajectory prediction, hitting point determination, inverse kinematics and robot trajectory generation. In general, the process of playing table tennis in most existing robot systems is as follows. First, the robot tracks the ball in motion according to its vision system. After predicting the trajectory of the ball, the robot calculates the position of the shot. Finally, the ball is returned to the opponent. In this process, it is especially important to accurately obtain the position of the ball. The three-dimensional position of the table tennis is detected in the reference coordinates [3, 4], and then the state of the ball (position and velocity) is accurately estimated by polynomial fitting [5] or by Kalman filtering [4]. According to the current state, the ball trajectory and the ball's landing point (the ball's landing point at the opponent's table) can be predicted.

Compared with the monocular vision system, the multi-vision system is also a good choice for robots and has been widely used in target estimation [6].

In recent years, researchers have focused on how to more intelligently hit the ball, including the rotation of the ball [6, 7], the accurate rebound model [8], and the recognition of the racket pose [9]. The existing rebound model relies mainly on the

speed before and after the shot, and these models are not very accurate. In fact, the posture of the racket also has an important influence on the rebound model. Because in the human table tennis competition, the observation of the opponent's behavior is the key to winning the game. Experienced athletes can react in advance by capturing the opponent's batting behavior. Chen et al. proposed a multi-condition threshold segmentation method widely applicable to racquet trajectory tracking. One limitation of this approach is the choice of parameter thresholds. Due to different lighting conditions, these thresholds need to be redefined and the environmental adaptability is weak. Wang et al. [9] proposed an interaction between the frame-observable Markov decision process of a human-robot expected behavioral choice, which is applicable to table tennis. The trajectory of the opponent's racquet is acquired by three cameras. The predictive model is a three-relational hit back action (forehand, default, backhand) that establishes the relationship with the opponent's action through the form of a black box. Only these can not get the trajectory of the opponent's racket. According to the number of cameras used, it can be divided into monocular vision system, binocular vision system and multi-vision system. This article uses a binocular vision system.

The main contributions of this paper are as follows. Firstly, the kinematics model of the ball trajectory is established, and the predicted trajectory of the ball is obtained according to the kinematic model and two high-speed cameras to find the predicted falling point. Subsequently, the trajectory of the racquet and the position of the racquet at the time of the hitting are obtained by using the thresholding method. Attitude, the true trajectory of the ball is obtained through the binocular vision system; the final analysis results show that the different racquet poses at the moment of hitting have an effect on the trajectory and the point of the ball.

2 Establish a Coordinate System

The world coordinate system we established is based on the center of the table tennis table. The Y axis is parallel to the long side of the table. The X axis is parallel to the short side of the table. The Z axis is perpendicular to the plane of the table. The coordinates are all in this coordinate system. The units in this coordinate system are all centimeters.

3 Track Prediction Part

In this part, we will complete the part of the trajectory prediction, including the prediction of the ball trajectory and the prediction of the landing point.

3.1 Ball Trajectory Prediction

First, we establish a kinematic model of table tennis when it is moving in the air. We use standard table tennis, quality $m = 0.0027 \text{ kg}$, diameter $d = 40 \text{ mm}$. Through mechanical analysis, there are three forces in the table tennis game, gravity G , air resistance F_f , Magnus force F_m .

- (a) gravity $G = mg$. g is the acceleration of gravity, the direction is vertical downwards, $g = 9.8 \text{ m/s}^2$.
- (b) air resistance F_f . It is opposite to the direction of velocity v and is calculated by fluid dynamics:

$$F_f = \frac{1}{2} \rho C_f S v^2 \quad (1)$$

Among them, ρ is the air density, $S = \frac{\pi d^2}{4}$ is the windward area of the ball in the air, v is the speed of the ball, and the drag coefficient 4 must be measured experimentally. Its expression is:

$$C_f = 0.508 + \left(\frac{1}{22.053 + 4.196 \left(\frac{v}{\omega} \right)^{5/2}} \right)^{2/5} \quad (2)$$

From the above formula, we can see that C_f depends on $\frac{v}{\omega}$, v is the speed of the ball movement, and ω is the angular velocity of the ball rotation.

- (c) Magnus force F_m . This force is perpendicular to v and ω . The formula for Magnus force is:

$$F_m = \frac{1}{2} \rho C_m S v^2 \frac{\vec{\omega}}{\omega} \times \vec{v} \quad (3)$$

C_m is the Magnus force lift coefficient, similar to the drag coefficient, which must be measured experimentally. The expression is:

$$C_m = \left(\frac{1}{2.022 + 0.981 \left(\frac{v}{\omega} \right)} \right) \quad (4)$$

The formation of Magnus force, such as when the ball is the bottom spin, the right hand rule can know that the angular velocity of the ball is away from you, Magnus force F_m direction upwards. The Magnus force F_m of the upper spin is down.

$$\begin{cases} \frac{dx}{dt} = v_x \\ \frac{dy}{dt} = v_y \\ \frac{dz}{dt} = v_z \\ \frac{dv_x}{dt} = \frac{1}{m} \left(-\frac{1}{2} C_f \rho S v v_x + \frac{1}{2} C_m \rho S v (v_z \sin \varphi - v_y \cos \varphi) \right) \\ \frac{dv_y}{dt} = \frac{1}{m} \left(-\frac{1}{2} C_f \rho S v v_y + \frac{1}{2} C_m \rho S v v_x \cos \varphi \right) \\ \frac{dv_z}{dt} = \frac{1}{m} \left(-\frac{1}{2} C_f \rho S v v_z + \frac{1}{2} C_m \rho S v v_y \sin \varphi \right) - g \end{cases} \quad (5)$$

Assume that the initial velocity of the ball when it leaves the racket at the time of hitting is v_0 , its angle with the XY plane is θ , and its angle on the XY plane and the X-axis is β , $\vec{\omega} = \omega(0, \sin, \cos)$, φ is the Angle between the ball's axis of rotation and the

z -axis, ω is the axis of rotation of the ball. Thus, we get the kinematics equation of the ball.

For a standard table tennis table, the length is 274 cm and the width is 152.5 cm. The center of the table tennis bat on the robot arm is 96 cm above the ground. At the same time, we are clicking the ball, so the initial position of the ball (190, 20, 20).

Figure 1 clearly reflects the difference between the three types of balls. The trajectories of the three balls are topspin, non-spin and backspin respectively, and the simulation results well reflect the predicted results. Because of the magnus force, the force on the topspin ball is downward, the force on the backspin ball is upward, and the unrotated ball is halfway between the two trajectories. The drop point we get for the top spin is $(-10.873, 66.576, 0)$, and the drop point for the lower spin is $(-20.269, 145.057, 0)$.

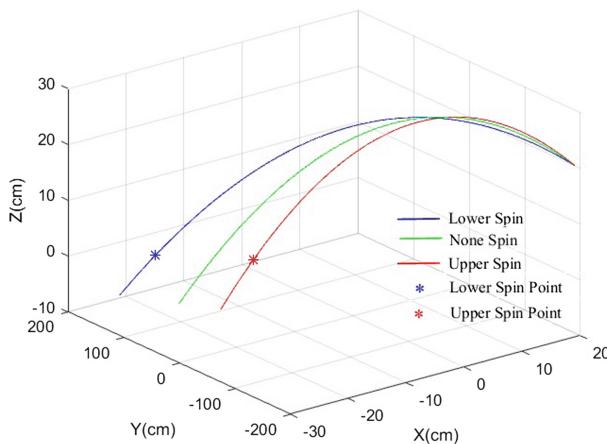


Fig. 1. Comparison of predicted trajectories of upper and lower spins

4 Experiment

In this part, we identify the trajectory of the racket, the trajectory of the ball, and the falling point of the ball through the binocular vision system. The hockey is compared with the predicted results to verify the accuracy of the predicted results.

For table tennis, in the initial state of the ball and the external environment, the pose of the racket becomes an important factor affecting the ball trajectory and the drop point. In order to better recognize the pose of the racket, we select the red side of the racket for recognition and do some processing on the racket (Fig. 2). We draw four such black lines on the red surface of the racket. Because the recognition of the racket requires simple calculation, high computational efficiency, and high speed, we use the threshold cut method to identify the black square on the racket.

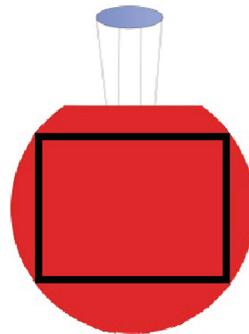


Fig. 2. Labeled racket

4.1 Determine the Racket Position

- (a) Extract the red area of the racket: First, we convert the image of the RGB space into the HSV space because the HSV space is more sensitive to color. In the HSV space, the minimum value of the red H component is 0, and the maximum value of the S component is 255, so we use the following formula to extract the red area of the racket.

$$I_R = \{(i, j) | H(i, j) < R_H, S(i, j) > R_S\} \quad (6)$$

Where (i, j) represents the coordinates of the point on the racket, $H(i, j)$ represents the H component of the point in the HSV space, $S(i, j)$ represents the S component of the point in the HSV space, R_H, R_S is the threshold on the H component and the S component, and I_R represents the extracted the red area extracted from the red area.

- (b) Identify black lines: Find the position of the black line on the racket, identify the black rectangle marked in the red area I_R , determine the position of the black box, and use the threshold cutting method. The formula is as follows.

$$I_B = \{(i, j) | H(i, j) < B_H, V(i, j) > B_V, (i, j) \in I_R\} \quad (7)$$

Through the above method, the black border on the racket can be obtained. At the same time, at the moment of hitting the ball, the pose of the racket is recognized. This will identify whether it is the upper or lower spin.

4.2 Get the Drop Point

In a real robotic system, we used a seven-degree-of-freedom robotic arm to hit the ball. At the same time, the ball is initially fixed with a highly stable ball-out mechanism, and the initial position of the ball is $(190, 20, 20)$. We let the robotic arm hit the ball in both the above-mentioned spin and the lower spin, and each swing 100 times, and record the position of the ball.

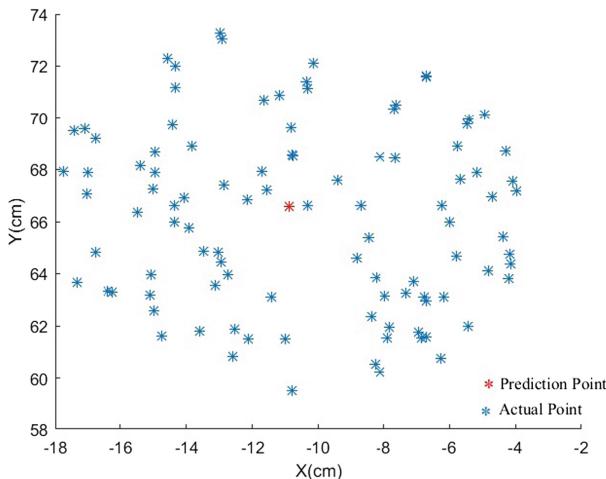


Fig. 3. Actual drop point of the upper spin

From Fig. 3, we can see the points and predictions obtained during the actual hitting process. The distance between the falling points is within 10 cm, which shows that the predicted trajectories and falling points have extremely high accuracy.

5 Conclusion

As shown in Fig. 4, we can clearly see the trajectory and falling point of the upper and lower spins on the ball. The impact is still very large. Under the influence of Magnus force, the falling point of upper ball is closer, and the lower ball point is relatively far.

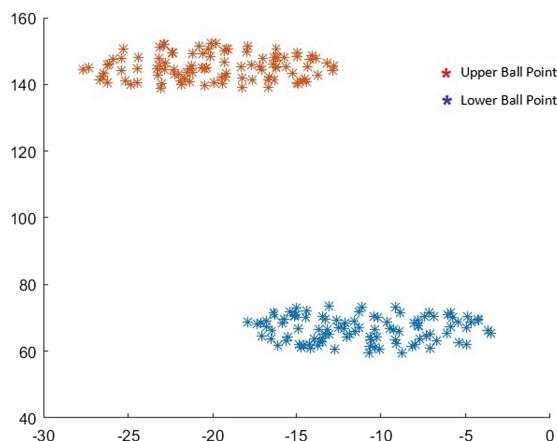


Fig. 4. Comparison of the actual placement of the upper and lower spins

In the table tennis system, not only the posture of the racquet, but also the position and strength of the ball and the initial speed of the ball before the ball are affected. This paper is to study the pose of the racquet for the trajectory. The effect of the points, so these are set to a fixed value. In the following work, we will add these factors, the judgment of the trajectory and the drop point will be more improved, and it will have a better hitting effect.

References

1. Mülling K, Kober J, Krömer O, Peters J (2013) Learning to select and generalize striking movements in robot table tennis. *Int J Robot Res* 32(3):263–279
2. Yu Z, Liu Y, Huang Q, Chen X, Zhang W, Li J et al (2013) Design of a humanoid ping-pong player robot with redundant joints. In: Proceedings of international conference on robotics and biomimetics, Shenzhen, China, pp 911–916
3. Andersson RL (1988) A robot ping-pong player: experiment in real-time intelligent control. MIT Press, Cambridge
4. Zhang Y, Wei W, Yuan D, Zhong C (2011) A tracking and predicting scheme for ping pong robot. *J Zhejiang Univ Sci C* 12(2):110–115
5. Zhang Z, Xu D, Tan M (2010) Visual measurement and prediction of ball trajectory for table tennis robot. *IEEE Trans Instrum Meas* 59(12):3195–3205
6. Zhang Y, Xiong R, Zhao Y, Wang J (2015) Real-time spin estimation of ping-pong ball using its natural brand. *IEEE Trans Instrum Meas* 64(8):2280–2290
7. Su H, Fang Z, Xu D, Tan M (2013) Trajectory prediction of spinning ball based on fuzzy filtering and local modeling for robotic ping-pong player. *IEEE Trans Instrum Meas* 62 (11):2890–2900
8. Nakashima A, Takayanagi K, Hayakawa Y (2014) A learning method for returning ball in robotic table tennis. In: Proceedings of IEEE international conference on multisensor fusion and integration for intelligent systems (MFI), September, pp 1–6
9. Wang Z, Boulias A, Mülling K, Schölkopf B, Peters J (2017) Anticipatory action selection for human-robot table tennis. *Artif Intell* 247:399–414



Rigid-Flexible Dynamics Analysis of a Large Deformation Flexible Beam Based on Finite Volume Beams

Wei Han, Long Qi, and Zhongyun Xiao^(✉)

China Aerodynamics Research and Development Center,
Mianyang 621000, China
scxiaozy@sina.cn

Abstract. The accurately calculating of dynamic response of large deformation of rigid-flexible coupling beam system is important for vibration control and safety and reliability design of flexible beam components, such as rotor and robot arm and so on. Although the traditional geometric accurate beam model based on finite elements can describe geometrically nonlinear effects, however, the shear self-locking phenomenon is easy to occur, which leads to big errors. The modeling methods based on finite volume beams in this paper provides a new method to solve this problem, because the finite volume method can avoid the locking phenomenon, and the large deformation model of the pure bending cantilever beam and high speed rotating flexible beam was established, then the simulation of the models was completed. The comparison of calculation results with reference values shows that the rigid-flexible large deformation model established in this paper proves to be free from shear locking, the accurate results can be obtained by using a few finite volume beam elements.

Keywords: Large deformation · Shear self-locking · Finite volume beams · Rigid-flexible coupling

1 Introduction

The cantilever beam components which was used in helicopter, robot, satellite and space station becomes high speed, heavy load, light weight and precise orientation, the rigid-flexible coupling dynamics analysis of the beam has become a research hot spot in recent years. The motion of flexible beam includes the coupling of rigid body motion and elastic deformation motion in a wide range, the coupling mechanisms are very complex, which leads to a more complex rigid-flexible coupling dynamics modeling and solution of large-scale moving beams.

The traditional methods of modeling large range of motion of rigid-flexible beam are divided into: mixed coordinates method, the absolute coordinate method and geometry precision beam method. Among them, the hybrid coordinate method is the earliest rigid-flexible dynamics modeling method. Kane and Eke [1, 2] in 1987 for the first time, puts forward the concept of “dynamic hardening effect”, when researching on the high-speed rotation problem of cantilever beam. On the basis of Kane, many scholars constantly have improved and corrected it, several methods of capturing

hardening dynamic term had been putted forward. For example, geometric stiffness method [3], initial stress method [4], geometric deformation constraint method [5] etc. But this method is not suitable for geometrically nonlinear beams. In order to fully consider the geometrical nonlinear effect of flexible beam, two different beam models are proposed. One method is the absolute node coordinate method, which was first proposed by Shabana [6]. Subsequently, many scholars have conducted a lot of research [7–10], and have also obtained rich research results. The other large deformation beam model is the geometrically accurate beam model, which preserves the assumption of rigid section and derives the objective strain variable based on the finite rotation theory. It was proposed by Ressner [11] in his study on the statics of beams, then Simo and Vu-Quoc [12, 13] extended it to the dynamic model of beam. In the last twenty years, many scholars have proposed a number of new geometrically nonlinear beam elements based on the geometrically accurate beam theory [14, 15]. They mostly use finite element methods to list equilibrium equations, which may occur “shear locking” in the calculation process, while the finite volume beam element based on finite volume method can avoid shear self-locking [16–18].

In this paper, the rigid-flexible coupling beam model is established based on geometrically accurate beam theory. A type of three-node finite volume beam element is used to discrete the beam. Two typical models of pure bending and high rotation flexible cantilever beam with large deformation are modeled and analyzed respectively. The calculated results are consistent with the experimental data and the reference data, which verifies the effectiveness of the proposed method, which is suitable for the rigid-flexible coupling dynamics analysis of large deformation and high rotation beam.

2 Basic Theory of Finite Volume Beam Element

Taking the three-node beam element as an example, the rigid-flexible coupling dynamic modeling method based on finite volume method is introduced below. In this paper, the beam model is discretized by a three node and three-dimensional control body, based on the finite volume method, six variables (three position coordinates and three cross-section direction parameters) are stored in two control body centers, the value of the interface between control bodies can be obtained by interpolating the central variables of two adjacent control bodies. The discrete diagram of the control body is shown in Fig. 1 below. Based on the above two control center nodes, the beam can be divided into three small beams, and the force is shown in Fig. 2.

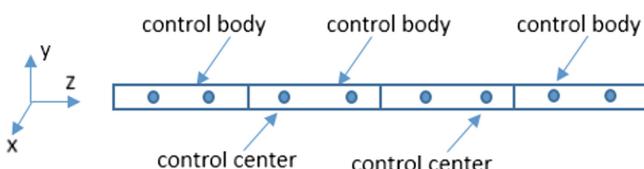


Fig. 1. Discrete control volume

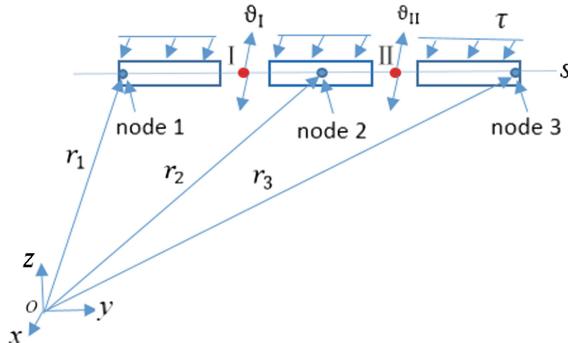


Fig. 2. The three-node finite volume beam element under internal and external loads

Take one of the control body element and use the Arabic numerals 1, 2 and 3 to represent the nodes of the beam. The points represented by the Roman numerals I and II are the two control center points of the element and the calculation points of the beam, the coordinate of these points in the reference frame is $s_I = -1/\sqrt{3}$ and $s_{II} = 1/\sqrt{3}$, there exists a quadratic polynomial combination function $N_j(s)$ as shown below:

$$N_1(s) = \frac{s(s-1)}{2}, \quad N_2(s) = 1 - s^2, \quad N_3(s) = \frac{s(s+1)}{2} \quad (1)$$

Thus, the position $\mathbf{r}(s)$ and direction parameters $\mathbf{g}(s)$ of any point on the beam element can be expressed as a linear combination of the corresponding parameters of node 1, 2 and 3. Where, the direction parameter is the Gibbs-Rodriguez parameter \mathbf{g} , whose specific expression is shown as follows:

$$\mathbf{g} = 2\tan(\theta/2)\mathbf{n} \quad (2)$$

In the above equation, \mathbf{n} is the direction vector of the rotation axis, and the amplitude θ is the overall rotation Angle.

According to Eq. (1), the position of any point on the reference line can be written as follows:

$$\mathbf{r} = N_j \mathbf{r}_j \quad (3)$$

N_j is the shape function of the Jth node, \mathbf{r}_j is the position vector of the Jth node. The above expression represents a sum operation on a pointer of j, $j = 1, 2, 3$. Similarly, the direction parameter of any point can be expressed as:

$$\mathbf{g} = N_j \mathbf{g}_j \quad (4)$$

\mathbf{g}_j is the direction vector of the Jth node, when $s = -1, 0$ and 1 correspond to the position vectors and direction parameters of the three nodes in the local coordinate system respectively.

In the Fig. 2 above, each beam should satisfy the balance formula of force and moment. The internal force is only calculated at the calculation point, which greatly simplifies the balance equation of the beam.

Therefore, the balance equation of each beam element should be written as the balance equation of three beam segments, and the moment arm matrix \mathbf{U} is defined as follows:

$$\mathbf{U}(\mathbf{r}) = \begin{bmatrix} \mathbf{I} & \mathbf{r} \times^T \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad (5)$$

In the above equation, \mathbf{r} is the relative position vector, \times represents the cross product operation of the vector. Therefore, the equilibrium equation of a single finite volume beam element can be written as follows:

$$\begin{bmatrix} \mathbf{U}(\mathbf{r}_{II} - \mathbf{r}_1)^T & 0 \\ -\mathbf{U}(\mathbf{r}_1 - \mathbf{r}_2)^T & \mathbf{U}(\mathbf{r}_{II} - \mathbf{r}_2)^T \\ 0 & -\mathbf{U}(\mathbf{r}_{II} - \mathbf{r}_3)^T \end{bmatrix} \begin{Bmatrix} \mathfrak{g}_I \\ \mathfrak{g}_{II} \end{Bmatrix} = \begin{Bmatrix} \mathcal{F}_I^I \\ \mathcal{F}_I^{II} \\ \mathcal{F}_{II}^3 \end{Bmatrix} \quad (6)$$

In the above equation, $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ respectively represent the position vectors of nodes 1, 2 and 3, \mathbf{r}_I and \mathbf{r}_{II} is the position vector of the calculated points. $\mathbf{U}(\cdot)$ is the moment arm matrix, whose definition is shown in Eq. (5). \mathfrak{g}_I and \mathfrak{g}_{II} respectively represent the internal force and moment of the calculation points and the expression is as follows:

$$\mathfrak{g}_i = \begin{Bmatrix} t_i \\ \mathbf{m}_i \end{Bmatrix} \quad (7)$$

In the above equation, i is the number of calculation points, t_i and \mathbf{m}_i are the internal force and the internal moment of the calculation points respectively. By the mechanics of materials, \mathfrak{g}_I and \mathfrak{g}_{II} can be further expressed as a function of the strain and curvature of the calculation points and, as shown below:

$$\begin{Bmatrix} \mathfrak{g}_I \\ \mathfrak{g}_{II} \end{Bmatrix} = \tilde{\mathbf{D}} \Psi \quad (8)$$

In the above equation, $\Psi = \{ \gamma_I \ \ \kappa_I \ \ \gamma_{II} \ \ \kappa_{II} \}^T$ is the generalized deformation of the beam, γ_I and γ_{II} is the translational strain of the calculation points, κ_I and κ_{II} is the curvature of calculation points, it is known from the geometrically accurate beam theory, the expression of them as shown below:

$$\tilde{\gamma} = \tilde{\mathbf{r}}'_p - \tilde{\mathbf{r}}'_0 \quad (9)$$

$$\tilde{\kappa} = \tilde{\theta}'_p \times - \tilde{\theta}'_0 \times \quad (10)$$

In the above equation, θ is angle of the vector, the derivative of the above equation is the derivative of arc length s . $\tilde{\mathbf{D}}$ is the stiffness matrix of the beam in the local cross-section coordinate system, and its expression is

$$\tilde{\mathbf{D}} = \begin{bmatrix} \tilde{\mathbf{D}}_I & 0 \\ 0 & \tilde{\mathbf{D}}_{II} \end{bmatrix} \quad (11)$$

The external force can be expressed as

$$\mathbf{F} = \begin{Bmatrix} \int_{S_I}^{S_I} \tau ds \\ \int_{S_I}^{S_{II}} \tau ds \\ \int_{S_{II}}^{S_3} \tau ds \end{Bmatrix} = \begin{Bmatrix} F_1 \\ C_1 \\ F_2 \\ C_2 \\ F_3 \\ C_3 \end{Bmatrix} \quad (12)$$

\mathbf{F} includes the external force term and the inertia term. F_i and C_i represents the external force and external torque of the i th part beam.

It can be seen from the above that for each volume beam element, there are six equilibrium equations and three equations about angular velocity, and there are a total of nine equations. After obtaining the dynamic balance equation of a beam element, the balance equation of all beam elements is assembled to obtain the balance equation of the whole beam:

$$\mathbf{M}(\mathbf{q}, t)\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, t)\dot{\mathbf{q}} + \mathbf{K}(\mathbf{q}, t)\mathbf{q} = \mathbf{Q}(\dot{\mathbf{q}}, \mathbf{q}, t) \quad (13)$$

Similar to the equilibrium formula of multi-rigid body system, in the above formula t is time, $\mathbf{M}(\mathbf{q}, t)$ is the overall mass matrix, $\mathbf{C}(\mathbf{q}, t)$ is the damping matrix, $\mathbf{K}(\mathbf{q}, t)$ is the stiffness matrix, $\mathbf{Q}(\dot{\mathbf{q}}, \mathbf{q}, t)$ is external force. \mathbf{q} is the degree of freedom vector, which can be expressed as

$$\mathbf{q} = [\mathbf{q}_1, \mathbf{q}_2 \cdots \mathbf{q}_n] \quad (14)$$

n is number of coordinates in the above formula. In addition to the above equilibrium equation, the beam system is also constrained by the motion equation, which can be expressed as

$$\phi(\mathbf{q}, t) = [\phi_1(\mathbf{q}, t), \phi_2(\mathbf{q}, t) \cdots \phi_m(\mathbf{q}, t)] = 0 \quad (15)$$

m is number of motion equations in the above formula. The constraint equation is added with an unknown reaction. Introducing the Lagrange multiplier $\lambda = [\lambda_1, \lambda_2 \cdots, \lambda_m]$, the constraint reaction can be expressed as

$$\mathbf{f} = \Phi_q^T(\mathbf{q}, t)\lambda \quad (16)$$

$$\Phi_q(\mathbf{q}, t) = \begin{bmatrix} \frac{\partial \Phi_1}{\partial q_1} & \frac{\partial \Phi_1}{\partial q_2} & \cdots & \frac{\partial \Phi_1}{\partial q_n} \\ \frac{\partial \Phi_2}{\partial q_1} & \frac{\partial \Phi_2}{\partial q_2} & \cdots & \frac{\partial \Phi_2}{\partial q_n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial \Phi_m}{\partial q_1} & \frac{\partial \Phi_m}{\partial q_2} & \cdots & \frac{\partial \Phi_m}{\partial q_n} \end{bmatrix} \quad (17)$$

Where, Φ_q^T is constrained jacobian matrix, after adding the constraint equation to the Eq. (13), the final equation system is the differential algebraic equations in the form of the first the Lagrange form, as shown below:

$$\begin{cases} \mathbf{M}(\mathbf{q}, t)\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, t)\dot{\mathbf{q}} + \mathbf{K}(\mathbf{q}, t)\mathbf{q} + \\ \Phi_q^T(\mathbf{q}, t)\boldsymbol{\lambda} - \mathbf{Q}(\mathbf{q}, \dot{\mathbf{q}}, t) = 0 \\ \phi(\mathbf{q}, t) = 0 \end{cases} \quad (18)$$

The exact solution can be obtained by solving the above differential algebraic equations DAEs. It is worth noting that this method is different from the finite method based on the traditional geometrical precise beam element, and the rigid-flexible couple multi-body equation of the beam is established based on D Alembert principle. The integral term of strain will not appear in the equation, which will not exaggerate the shear deformation and can effectively avoid the phenomenon of shear self-locking.

3 Numerical Example

Based on the finite volume beam element theory, the relevant calculation program was complied, the large deformation of cantilever beam under concentrated bending moment and high speed rotating flexible beam was calculated in this paper.

3.1 Calculation of Large Deformation of Cantilever Beam Under Concentrated Bending Moment

The calculation model is shown in Fig. 3, One end of the beam is fixed and the other end is completely free, The length of the beam $L = 10$ m, rectangular section size $A = 0.1 \times 0.1$ m², the elasticity modulus $E = 2.1 \times 10^{11}$, the density $\rho = 7800$ kg/m³, Considering that the free end of the beam is subjected to concentrated bending moment, the bending deformation of the beam and the change of the displacement at the end of the beam are considered. The axial displacement and lateral displacement of the beam are respectively expressed by u and v .



Fig. 3. Cantilever beam undergoes bending load

Considering the damping of the beam, the beam will eventually reach the state of static equilibrium under the action of load, the beam is divided by three three-node finite volume beam elements, the inertia of the measurement was simulated with six concentrated mass element. In order to make the cantilever beam reach the equilibrium state as soon as possible, the vertical damping ratio $c = 0.1$ was selected for simulation. The final deformation state of the beam under the action of different bending moments is shown in Fig. 4.

Under the action of bending moment $M = 2\pi Ez/L$, the change curve of displacement with time at the end of beam is shown in Fig. 5.

It can be seen from the above displacement time curve that the cantilever beam model reaches the final equilibrium state after two cycles under the action of large structural damping, and the result is highly consistent with the analytical solution with only 3 elements. It is shown that the large deformation beam model based on the proposed method is effective and efficient.

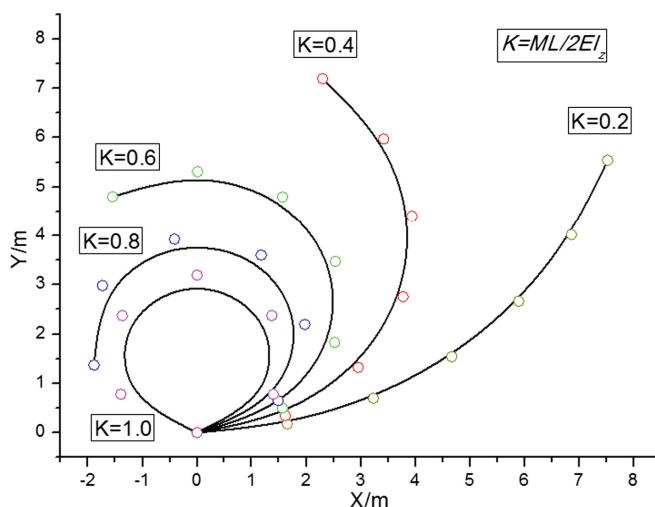


Fig. 4. Large deformation of the cantilever under bending moments

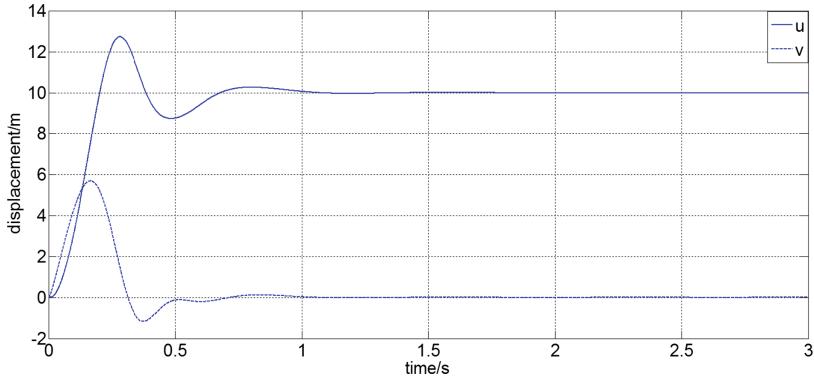


Fig. 5. The tip displacements of the cantilever beam with an end bending load

3.2 Deformation Calculation of High Speed Rotating Flexible Beam

High speed rotating flexible beam has been used by many scholars [19, 20] to investigate whether the established model has the ability to capture the dynamic hardening effect, calculation parameters of rotor are taken from literature [20], the length of the rotor $L = 10$ m, the area of section $A = 1 \text{ m}^2$, the elasticity modulus $E = 2.8 \times 10^7 \text{ Pa}$, the second moment of the section $I_z = I_y = 5 \times 10^{-4} \text{ m}^2$, the density $\rho = 1.2 \times 10^3 \text{ kg/m}^3$. The model is shown in Fig. 6.



Fig. 6. Rotating flexible beam

The rotation Angle function of the flexible beam around the rotation center is expressed as

$$\varphi(t) = \begin{cases} \frac{5}{16} \left[0.5t^2 + \frac{15}{2\pi^2} (\cos(\frac{2\pi t}{15}) - 1) \right], & 0 \leq t < 15 \\ 6t - 15, & 15 \leq t \end{cases} \quad (19)$$

Ignoring the structural damping, the flexible beam is divided into three three-node finite volume beam elements, the root of the beam is modeled by rotating hinge elements, and the quality of the flexible beam is modeled by six concentrated mass elements. The simulation time $t = 30$ s, time step $h = 0.001$ s, the displacement response curve of the free end of the flexible beam obtained by calculation is shown in Fig. 7.

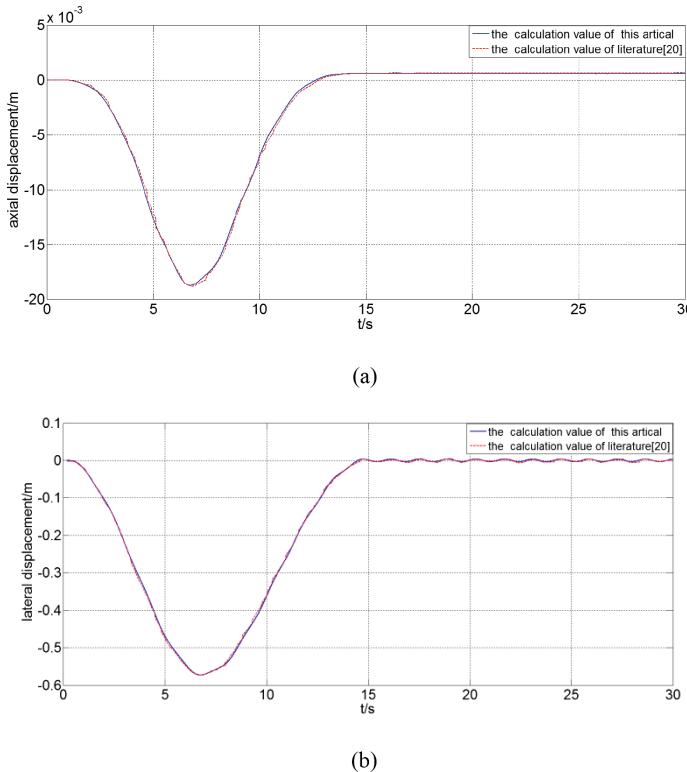


Fig. 7. Displacement curve at the free end of beam

As can be seen from Fig. 7: the free end displacement of the flexible beam increases gradually when the angular velocity increases with the time, and then decreases gradually, when the speed stabilizes at a constant value, the axial displacement gradually declines and stabilizes at 5.51×10^{-4} m, this is 7.2% different from the calculation results in literature [20], the error is in the scope of engineering error. The horizontal displacement fluctuates slightly around 0. The calculation result in this paper is very close to that in literature [20]. It is shown that the proposed method is suitable for hardening effect of high-speed rotating flexible beams.

In addition, only three finite volume beam elements are used in this paper, while five beam elements are used in literature [20], indicating that this element has well modeling efficiency.

4 Conclusion

In this paper, a rigid-flexible coupling dynamic model of large deforming beam is established based on finite volume beam elements, subsequent two typical large deformation flexible beams are simulated by writing programs, the method is proved to

be reliable by comparing the calculation result with the reference value. The method in this paper has the following characteristics: First, the rigid-flexible coupling beam with large deformation based on the finite volume elements can avoid “shear locking” phenomenon; Second, the method in this paper using less element can get accurate results, the high computational efficiency is reached. Third, the method in this paper can well consider “hardening effect” effect for the high speed rotating beam model. The method in this paper provides a theoretical basis for vibration control of beam components and improve the reliability and safety of beam component design, such as rotor and robot arm etc. and so on, the design can also be optimized according to the results of this paper.

Acknowledgments. This work was supported by the National Key Research and Development Program of China (2016YFB0200704) and the National Natural Science Foundation of China (11572341).

References

1. Kane TR, Ryan R, Banerjee AK (1987) Dynamics of a cantilever beam attached to a moving base. *J Guid Control Dyn* 10(2):139–151
2. Eke FO, Lasking RA (1987) On the inadequacies of current multi-flexible body simulation codes. In: Navigation and control conference, Monterey
3. Mayo J (1993) Geometrically nonlinear formulations of flexible multibody dynamics. University of Seville, Seville
4. Wallrapp O (1991) Linearized flexible multibody dynamics including geometric stiffening effects. *J Struct Mech* 19(3):385–409
5. Zhang DJ, Huston RL (1995) On dynamic stiffening of flexible bodies having high angular velocity. *J Struct Mech* 23(3):419–438
6. Shabana AA (1996) An absolute nodal coordinate formulation for the large rotation and deformation analysis of flexible bodies. *J Struct Mech* 21(5):453–486
7. Kerkkanen KS, Sopanen JT, Mikkola AM (2005) A linear beam finite element based on the absolute nodal coordinate formulation. *J Mech Des* 127(4):621–630
8. Garcid D, Valverde J, Domiguez J (2005) An internal damping model for the absolute nodal coordinate formulation. *Nonlinear Dyn* 42(4):347–369
9. Garcia-Vallejo D, Mikkola AM, Escalona JL (2007) A new locking-free shear deformable finite element based on absolute nodal coordinates. *Nonlinear Dyn* 50(2):249–264
10. Tian Q, Zhang Y, Chen L et al (2009) Dynamics of spatial flexible multibody systems with clearance and lubricated spherical joints. *Comput Struct* 87(13):913–929
11. Reissner E (1972) On one-dimensional finite-strain beam theory: the plane problem. *Zeitschrift fur angewandte Mathematik und Physik ZAMP* 23(5):795–804
12. Simo JC (1985) A finite strain beam formulation. The three-dimensional dynamic problem, Part I. *Comput Methods Appl Mech Eng* 49(1):55–70
13. Simo JC, Vu-Quoc L (1988) On the dynamics in space of rods undergoing large motions - a geometrically exact approach. *Comput Methods Appl Mech Eng* 2(66):125–161
14. Romero I (2004) The interpolation of rotations and its application to finite element models of geometrically exact rods. *Comput Mech* 34(2):120–130
15. Bau OA, Han S (2014) Interpolation of rotation and motion. *Multibody Syst Dyn* 31(3):339–370

16. Fallah N, Hatami F (2006) A displacement formulation based on finite volume method for analysis of Timoshenko beam. In: Proceedings of the 7th international conference on civil engineering, Tehran
17. Masarati P, Mantegazza P (1997) On the C0 discretisation of beams by finite elements and finite volumes. *Aerotecnica Missilie Spazio* 75(1):77–86
18. Jing L, Zhang W, Ming P (2015) Static and dynamic analysis of Timo-shenko beam model based on the finite volume method. *Nonlinear Dyn* 36(9):1217–1222
19. Zhao Z, Ren G (2012) A quaternion-based formulation of Euler-Beroulli beam without singularity. *Nonlinear Dyn* 67(3):1825–1835
20. Zhang Z, Qi Z, Wu Z (2015) Rigid-flexible dynamics analysis of a large deformation beam based on interpolation of strains 28(3):993–998



An Image Mosaic Algorithm Based on Improved ORB

Xianjin Du^(✉) and Lin Li

School of Optical-Electrical and Computer Engineering,
University of Shanghai for Science and Technology, Shanghai 200093, China
15216779869@163.com, lulin0211@163.com

Abstract. Due to the lack of scale invariance in ORB (Oriented FAST and Rotated BRIEF) algorithm, image mosaic in complex scenes, high mismatch rate, and easy to produce mosaic seams in synthetic images, etc. In this paper, an improved ORB algorithm for image mosaic based on BRISK feature extraction is proposed. Firstly, the FAST9-16 of BRISK algorithm is introduced to detect corner points and solve the scale invariance problem. Then ORB algorithm is used to describe feature points, and Hamming distance is used to roughly match image features. Then KNN algorithm and PROSAC (Progressive Sample Consensus) algorithm are used to refine and purify the matching points to further improve the accuracy. Finally, image seamless fusion and mosaic are completed by gradual-in-gradual-out weighted average fusion. The experimental results show that the proposed algorithm has good performance in image scaling, rotation and illumination intensity. It is a real-time image mosaic method with high accuracy and good mosaic effect.

Keywords: Image mosaic · BRISK feature · ORB operator · Scale invariant

1 Introduction

Image stitching [12, 13] technology is composed of three processes: image preprocessing, image registration and image fusion. Image stitching technology is widely used in computer vision, medical image analysis, remote sensing UAV image stitching [16] and virtual reality. Image registration is the basis of image mosaic technology and becomes a key step in mosaic. There are two kinds of image registration algorithms: based on gray correlation and based on features. The feature-based registration method registers the overlapped parts of an image by extracting the local invariant features of the image. Because of its robust and fast algorithm, it has become a research hotspot.

Lowe [1] proposed a scale invariant feature transform (SIFT) matching algorithm, which is widely used in mosaic because of its good scale-based, rotation, illumination and other invariant characteristics. However, because of the large amount of computation, the matching speed is relatively slow. Calonder [2] and Rosten [3] proposed BRIEF and FAST operators successively. The descriptor was extracted by Hessian matrix and Haar wavelet, and the dimension was reduced to 64 dimensions. The speed of the algorithm was greatly improved. Rublee [4] proposed ORB algorithm, which has the advantage of improving matching speed and reducing noise interference, but the

matching accuracy is low, does not have scale invariance. BRISK [10, 11] (Binary robust invariant scalable keypoints) uses AGAST corner detection algorithm to extract corner features. It has scale and rotation invariance, fast calculation speed, but unstable. Domestic scholars such as Party Jianwu [18] proposed an image mosaic optimization algorithm based on SIFT feature detection, which is suitable for image mosaic in complex scenes, but the matching time is long. Zhang et al. [6] studied image region block matching, which shortened the matching time, but the accuracy was not high. Li et al. [5] proposed an image mosaic algorithm based on region partitioning and scale invariant feature transform (SIFT). This algorithm guarantees a certain matching accuracy and improves the matching speed. Image fusion is another key part of image mosaic, and it is an effective way to eliminate the phenomenon of mosaic seam and duplication. Aiming at the problem of stitching seam elimination, the paper [7–9] proposed some algorithms, such as combining stitching line with multi-resolution fusion and weighted average fusion, which eliminated the stitching seam problem to a certain extent, but the image transition is still not smooth.

Aiming at the problem that ORB algorithm does not have scale invariance and high mismatch rate, a new image mosaic algorithm based on BRISK feature and ORB descriptor is proposed to improve the matching speed and accuracy of image mosaic. It can realize seamless mosaic of image with rotation invariance and scale invariance accurately, quickly and steadily. This algorithm can reduce the geometric distortion of the image and improve the matching speed on the premise of ensuring the accuracy of image matching.

2 The Algorithm Application

ORB algorithm uses oFAST corner detection operator for feature detection and rBRIEF feature description operator for feature description. It solves the problem that FAST feature detection does not have direction and BRIEF feature description does not have rotation invariance. In order to solve the problems of scale invariance and poor robustness of ORB algorithm, BRISK algorithm is introduced to construct scale space to make feature point detection scale invariant.

2.1 Feature Point Detection

Feature points are the more prominent points in the image, called corner points. Feature detection is to detect and extract these prominent points and then describe them. BRISK algorithm expands FAST corner detection from image plane to scale space, and uses FAST score to search maximum value in image plane and scale space. Firstly, scale-space pyramids are constructed for images. The pyramids consist of ordinary and intermediate layers, each of which has four. The scaling scale of each layer is as follows (1):

$$t(c_\epsilon) = 2^\epsilon, \quad t(d_\epsilon) = 2^\epsilon \times 1.5, \quad \epsilon = 1, 2, 3, 4 \quad (1)$$

Among them: t represents scaling scale values. c_ϵ represents the common layer of the image and d_ϵ europium represents the middle layer of the image.

All common surface layers are obtained by continuous semi-sampling of the original image c_0 , and the middle layer d_0 is obtained by 1.5 times the original image c_0 , while the other middle layers are obtained by continuous semi-sampling of d_0 , so an image pyramid with eight layers structure can be obtained. The FAST score is calculated in the form of FAST9-16 for the eight-layer image. It basically requires that the gray value of a pixel in the image is larger than that of the nine consecutive circles of 16 pixels around it.

Initially, the FSAT9-16 detectors are used to identify potential areas of interest in the common and intermediate layers using the same threshold. Next, non-maximum suppression is performed on the points of these regions in position and scale space. The so-called non-maximum suppression is to compare the sum of absolute values of the difference between the two feature points and the 16 pixels around them. At this time, non-maximum suppression needs to consider the nearest 8 points in the key point layer and the $9 + 9 = 18$ points in the upper and lower layers, totally 26 points for calculation. The searched points need to satisfy the following conditions: (1) the FAST score of the detected points is greater than that of their neighborhood in the same layer; (2) the FAST score of the detected points is greater than that of all points in the upper and lower layers. Detection points satisfying two conditions are the key points.

Subpixel interpolation, considering that the image display changes along the scale dimension, we fit and interpolate each detected maximum point. Firstly, two-dimensional quadratic function fitting is used to fit the scoring values of the scale of the extreme points and the corresponding positions of the upper and lower levels, and three sub-pixel maximum coordinate points are obtained. Then, one-dimensional parabolic fitting of these maximum points on the scale axis is carried out, and the accurate FAST score and scale value of the maximum are obtained. Finally, the new image coordinates in the adjacent layer are determined by repeated coordinate interpolation.

2.2 Description of Feature Points

ORB algorithm uses rBRIEF algorithm for feature description. Because the descriptors generated by random points are sensitive to noise, the ORB algorithm randomly selects $5 * 5$ pairs of pixels in the field of $31 * 31$ pixels of FAST feature points, calculates the sum of pixels quickly by integral graph method, and obtains binary string. The pixel block is composed of the results of binary gray comparison. The calculation formulas of pixel block comparison and descriptor are as follows:

$$\tau(p; x, y) := \begin{pmatrix} 1 : p(x) < p(y) \\ 0 : p(x) \geq p(y) \end{pmatrix} \quad (2)$$

$$f_n(p) : \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i; y_i) \quad (3)$$

In the formula, $p(x)$ is the gray value of the image block p at point x . n can take 128, 256 and other values. Different values will affect the operation speed, recognition rate and storage rate. ORB itself does not have rotation invariance. The solution takes the centroid direction of feature points obtained by formula (4) as the main direction of BRIEF. Rotate the matching points x_i and y_i in our neighborhood. Let n pairs of test points (x_i, y_i) generate feature point descriptors, and define a $2 * n$ matrix.

$$\mathbf{M} = \begin{pmatrix} x_1, & \dots, & x_n \\ y_1, & \dots, & y_n \end{pmatrix} \quad (4)$$

For each test point of rotation matrix formed by the angle of the R_θ : $R_\theta = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$, The \mathbf{M} matrix becomes to form coordinates as shown in (5)

$$\mathbf{M}_\theta = R_\theta \mathbf{M} \quad (5)$$

So the descriptor with rotation invariance is obtained as follows:

$$g_n(p, \theta) := f_n(p) | (x_i, y_i) \in \mathbf{M}_\theta \quad (6)$$

ORB extracts BRIEF feature descriptors according to the direction of formula (4).

2.3 Feature Matching and Purification

After feature extraction, a binary coded descriptor is obtained for each feature point in the image. Hamming distance refers to the number of different characters in the corresponding position of two strings of the same length. To take the two strings XOR operation, statistical calculation result is 1 of the number, this number is the two string Hamming distance.

Let the descriptors of A and B be A 10101011 and B 10101010, then the Hamming distance is expressed as $d(A, B) = 1$. If a, b two string length is n , then the similarity calculation formula:

$$S = 1 - \frac{d(a, b)}{n} \%$$

There may be mismatches in Hamming distance rough matching. KNN algorithm and PROSAC [14] algorithm are selected to purify the rough matching. PROSAC (progressive sample consensus) algorithm firstly ranks the unordered matching pairs in descending order according to the quality of matching, and selects the data subset with better quality as the initial matching point set. When the model parameters are fitted, the initial matching set is iterated repeatedly to get the maximum internal points, that is, the correct matching points after screening. The sample extraction of PROSAC algorithm is computationally heavy and prone to deviation. At the same time, in order to

reduce the computational load and deviation of sample extraction in PROSAC algorithm, KNN algorithm is used to find matches whose distance is less than a certain threshold, which reduces the computational load and improves the matching speed and accuracy.

Image mosaic, usually used to describe the perspective transformation of reference image and the images of geometric transformation, change of perspective can be mapped into a rectangular quadrilateral. If the matching coordinates $p = (x, y)$ by projection mapping to $q = (x_0, y_0)$ transform formula as follows: $q = Hp$.

$$H = \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & 1 \end{bmatrix} \quad (7)$$

After single-strain transformation of registration image with formula (7), the matched image has certain overlapping area, that is, at the same pixel position, two registration images have different pixel expression, so it is necessary to fuse the overlapping pixels.

3 Fusion Image

Image fusion is to synthesize two or more images into a new image using a specific algorithm. By utilizing the spatial correlation and information complementarity of two (or more) images, the fused images can give a more comprehensive and clear description of the scene. In the process of image mosaic, due to the influence of illumination, rotation and translation, the effect of direct mosaic will appear such phenomena as stitching seam and duplication.

Therefore, we adopt the gradual-in-gradual-out weighted fusion method to solve the problems of stitching seam and duplication caused by large exposure difference and noise interference. The fusion methods are as follows:

$$f(x, y) = \begin{cases} f_1(x, y) & (x, y) \in f_1 \\ \alpha * f_1(x, y) + \beta * f_2(x, y) & (x, y) \in (f_1 \cap f_2) \\ f_2(x, y) & (x, y) \in f_2 \end{cases} \quad (8)$$

In the formula, f is the fused image, f_1 and f_2 are the stitching sutures in the fusion zone, and alpha and beta are the weights of the pixels in the reference image and the image to be stitched. There is a relationship between $\alpha + \beta = 1$, and $0 \ll \alpha \leq 1$, $0 \ll \beta \ll 1$. In the overlapping part, when alpha is gradually changed from 1 to 0, beta is gradually changed from 0 to 1, so the smooth fusion between the reference image and the mosaic image is completed. The image mosaic process based on the improved ORB is shown in Fig. 1.

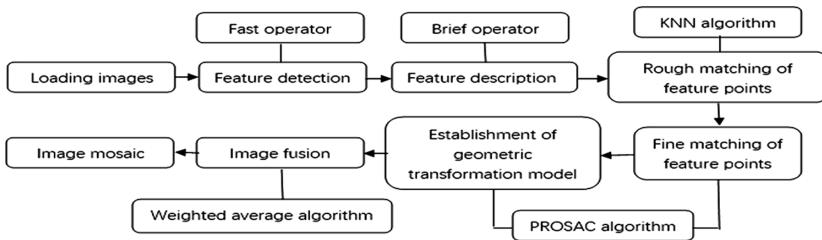


Fig. 1. Image mosaic process based on improved ORB algorithm

4 Analysis of Experimental Results

Three pairs of images under illumination, rotation and scale transformation are selected for image registration. Simultaneously, single object image and complex background image are selected for stitching. Hewlett-Packard Computer: Intel (R) Core (TM) i5-3230 M processor, CPU: 2.60 GHz, 4 GB memory, 64-bit Win10 operating system, Visual Studio 2013 and OpenCV2.4.13. The experimental data are Leuven data, Boat data and Ubc data in the data set provided by Mikolajczyk and Schmid. The data sets provided by Mikolajczyk and Schmid contain several image sets with different geometric and illumination intensity transformations. as shown in Fig. 2(a), (b) the image size was 850 * 680 pixel (c) image size is 800 * 640 pixel, The results of original matching and fine matching are shown below.

For the experiments, this paper uses the following parameters to evaluate the experiment: (1) The running time to evaluate the matching speed. (2) Using an objective evaluation index, the correct rate of CMR (Correct Matching Rate), the CMR value is greater, the better matching performance, defined as CMR:

$$\text{CMR} = N_c/N$$

N is the logarithm of all matching points and N_c is the logarithm of correct matching points.

Table 1 for the correct matching rate and schedule. Comparing Table 1, we can see that in the experimental pictures taken by scale reduction and left rotation (b) and (c), the ORB improved algorithm in this paper has a remarkable effect, with a slight increase in time consuming, but it enhances the scale scaling and rotation invariance, reduces the logarithm of matching points, and improves the average accuracy by about 20%. The experimental results are as follows:

4.1 Image Stitching Results

After purifying and eliminating the mismatched points, the image after registration is fused by the gradual-in-gradual-out weighted fusion algorithm. The final result of the mosaic is shown in Fig. 3. From the figure, we can see that the mosaic image has no ghost, no mosaic seam, and the brightness of the image changes uniformly, high fidelity, and good mosaic effect.

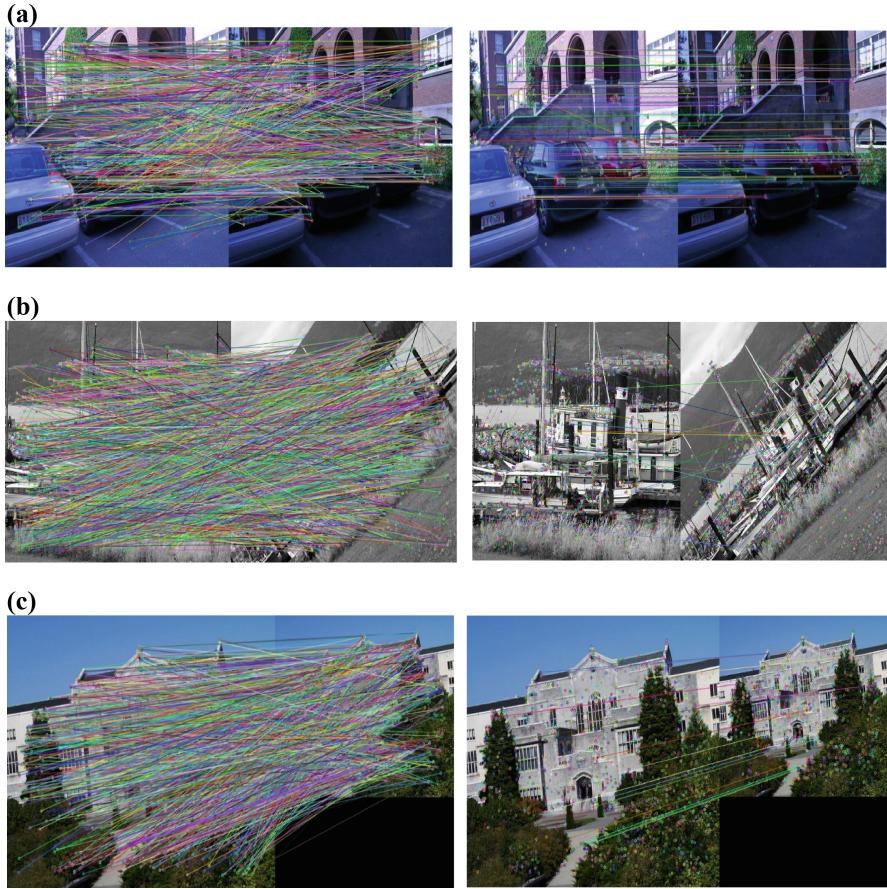


Fig. 2. (a) Illumination change feature point matching comparison diagram (b) Comparison of rotation feature points matching (c) Scale change feature point matching comparison diagram

Table 1. Correct matching rate comparison and time-consuming comparison unit: second

| No. | Algorithm | Matching point pair | Correct point pairs | CMR/% | Time/s |
|-------------|-----------------|---------------------|---------------------|--------|--------|
| Figure 2(a) | Brisk | 132 | 111 | 84.09% | 0.138 |
| | ORB | 173 | 162 | 93.64% | 0.0984 |
| | Literature [15] | – | – | 70.19% | 1.3876 |
| | This paper | 104 | 98 | 94.23% | 0.1343 |
| Figure 2(b) | Brisk | 6 | 5 | 83.33% | 0.3862 |
| | ORB | 214 | 167 | 78.04% | 0.1913 |
| | Literature [15] | 15 | 14 | 92.80% | 1.1816 |
| | This paper | 17 | 16 | 94.11% | 0.3305 |
| Figure 2(c) | Brisk | 48 | 44 | 91.67% | 0.1192 |
| | ORB | 476 | 333 | 69.65% | 0.0715 |
| | Literature [15] | – | – | – | – |
| | This paper | 13 | 12 | 92.30% | 0.1123 |



Fig. 3. Result of image stitching and fusion

5 Conclusion

This paper presents an improved ORB image mosaic algorithm based on BRISK. Combining BRISK feature point detection operator with Brief feature point description, the improved algorithm has the advantages of good rotation, scale invariance, fast and stable. KNN algorithm is used to match the feature points initially. Then PROSAC algorithm is used to match the matched feature points precisely and calculate the geometric transformation model. Finally, the image is fused by gradual gradual-out weighted fusion algorithm. It is compared with ORB algorithm, BRISK algorithm and literature [15]. The experimental results show that this algorithm makes up for the lack of scale invariance of the original ORB algorithm, and effectively improves the speed and accuracy of image matching when scale changes. When scale changes, the matching accuracy of this algorithm is improved by about 20% compared with that of ORB. In the case of image blurring, illumination change, rotation change and so on, although the time increases slightly, the matching accuracy has been improved. Generally speaking, the algorithm in this paper is an image mosaic method which can meet the real-time requirements and high registration accuracy.

References

1. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110
2. Calonder M et al (2010) BRIEF: binary robust independent elementary features. In: Proceedings of 11th European conference on computer vision, ECCV 2010, Part IV, Heraklion, Crete, Greece, 5–11 September 2010
3. Rosten E, Porter R, Drummond T (2008) Faster and better: a machine learning approach to corner detection. *IEEE Trans Pattern Anal Mach Intell* 32(1):105–119
4. Rublee E, Rabaud V, Konolige K et al (2011) ORB: an efficient alternative to SIFT or SURF. In: International conference on computer vision. IEEE, pp 2564–2571
5. Li Y, Li G, Gushaohu (2016) Image mosaic algorithm based on region partitioning and scale invariant feature transformation. *Opt Precis Eng* (05):1197–1205

6. Zhang H, Zhang X, Luo Y, Chen B (2017) A fast SIFT registration method based on image region partitioning. *Infrared Technol* 39(04):341–344
7. Luo Y, Wang Y, Zhang H (2018) Image mosaic algorithms combining optimal suture line and improved inward-out method. *Infrared Technol* 40(4):382–387
8. Guyu, Zhou Y, Ren G, Feng Q, Lu G (2017) Image mosaic combining optimal suture and multi-resolution fusion. *CSCD J Image Graph* 22(06):842–851
9. Huang L, Chen C (2014) Research on image fusion algorithm in panoramic mosaic. *J Electron Inf* 36(06):1292–1298
10. Dong Q, Liu J, Wang C et al (2017) Image mosaic algorithm based on improved BRISK. *J Electron Inf Technol* 39(02):444–450
11. Leutenegger S, Chli M, Siegwart RY (2012) BRISK: binary robust invariant scalable keypoints. In: IEEE International Conference on Computer Vision (ICCV), pp 1589–1596
12. Quzhong, Qiao G, Lin S (2015) A fast mosaic algorithms for eliminating image mosaic joints and ghost shadows. *Comput Sci* 42(03):280–283
13. Han M, Yan K, Qin C (2019) Image mosaic algorithm for UAV aerial photography based on improved KAZE. *Acta Automatica Sinica* 45(02):305–314
14. Wang J, Shijun, Wu X (2008) Overview of image mosaic technology. *Comput Appl Res* (07):1940–1943, 1947
15. Di Y, Chen Y, Chen Y, Chen Y (2011) Overview of image mosaic algorithms for UAV. *Comput Appl* 31(01):170–174
16. Chum O, Matas J (2005) Matching with PROSAC - progressive sample consensus. In: IEEE computer society conference on computer vision and pattern recognition, CVPR 2005. IEEE
17. Anweisheng, Yu R, Wu Y (2015) Image registration algorithm based on FAST and SURF. *Comput Eng* 41(10):232–235, 239
18. Dang J, Zongyan, Wang Y (2012) Research on image mosaic optimization algorithm based on SIFT feature detection. *Appl Res Comput* 29(01):329–332



Optimization Path Planning Algorithm Based on STL File Reconstruction for Automated Fiber Placement

Yanjuan Zhu^(✉) and Kanzhong Yao

School of Aerospace Engineering and Applied Mechanics, Tongji University,
Shanghai 200092, China
zhu_yj@tongji.edu.cn

Abstract. A path planning optimization algorithm for automated fiber placement is proposed. According to the STL model file triangle patch and the vertex index number of the triangle, the STL model file is converted into points and faces matrixes, the data reconstruction of the file is realized and a STL model file reading method is constructed. Based on the reconstructed data, a fiber placement path planning algorithm is proposed. The adjacent path to be planned are solved by points, normal vectors, tangent planes and coordinate axes of the laying surface and the corresponding mathematical model is established. The algorithm is simulated and the approximate iterative method is taken as a reference. The simulation verifies that the algorithm is accurate and reliable.

Keywords: Path planning · Automated fiber placement · Data reconstruction

1 Introduction

The automated fiber placement path planning is to establish a mathematical model on the surface of the laying member to form a composite prepeg (or tow) arrangement that meets the structural design requirements [1, 2]. Path planning is the prerequisite of the subsequent steps of automated placing CAD/CAM software. The proper path planning is the key to realize the automated fiber placement technology.

The study of placement path can be traced back to the early 1990s. In 1987, Lewis [3] first proposed the concept of “natural path” and applied it to path planning to realize the small deformation of the prepeg belt during the laying process. After that, the research on the method of laying out the path planning method is mainly based on the “natural path”. The research work in reference [4] was based on “natural path”, proposed a method for controlling the direction of the prepeg tape in the process of automated tape placement. Research [5] analyzed the geometric properties of the “natural path” on the developable surface, derived and proved that “natural path” is equivalent to the geodesic line on the developable surface, and introduced boundary control points on the general surface to adjust the center point of the prepeg to reduces the deformation of the prepeg belt, then a solution method for the general surface “natural path” was constructed. Wang [6] proposed a numerical calculation method to solve the geodesic line between two points. The boundary value problem of differential

equations was transformed into an initial value problem, which made the initial direction selection more intuitive. Luo [7] reconstructed the mesh information by approximating the free surface with a triangular mesh, and established the correlation information between points, edges and faces, completed the solution of the “natural path” on the surface, and calculated the boundary path of the surface. Researchers from Nanjing University of Aeronautics and Astronautics [8] studied the method of laying out the path of integrated fiber for composite materials. Through the finite element analysis of the component model, the internal stress distribution of the component is obtained, and then a number of reference fiber placement paths are planned based on the magnitude and direction of the principal stress. Reference [9] discussed the control strategy of the width of the laying and the accuracy of the placement. A calculation method for the control trajectory of the center of the machine press roll is proposed to meet the complex control requirements of the fiber placement manipulator.

Based on former research [10, 11], a path planning optimization algorithm for automated fiber placement of composite materials is proposed. First, we create the STL file triangle patch and the fixed point index number of the triangle, and convert the STL model file into a point and face matrix, realize the data reconstruction of files and construct STL model file reading method. Based on the reconstructed data, a path planning algorithm is proposed, the corresponding mathematical model is established and the algorithm is solved to obtain the optimized path planning method.

2 Modeling Data and Path Planning

2.1 STL Data Reconstruction

STL model file is the most commonly used system model file interface format for CAD/CAM in the world. Most model software support files in this format. The STL file stores the triangle information of the triangle shape of the model, including the vertex coordinates of the triangle, the normal vector of the triangle patch, and the structure of the file is simple and easy to read.

STL model is regular, spatially closed, uniquely expressed object and always bounded. The STL model file contains some key geometric information such as points, lines, and surfaces, and can usually express the solid model information. The standard STL file data model should satisfy three rules:

- Right-hand rule:** The preservation of the three vertices of the triangle and the direction of the normal vector needs to be in a specific order. According to the right-hand rule, with the right-hand gripping, the vertices of the triangle are saved in the order of the four-finger direction, and the direction in which the thumb points is also the direction of the normal vector of the triangle patch. The normal vector is expressed as a unit vector.
- Vertex rule:** All two adjacent triangle patches have only two shared vertices, and a triangle patch can have only one common edge with its adjacent patch. In addition, the vertex of the triangular patch cannot fall on the edge of another triangular patch.

3. **Edge rule:** Each edge of a triangle must and can only be shared by two triangles.

The data structure of the STL file triangle patch unit is shown in Table 1.

Table 1. Data structure

| | | |
|-----------------|---------------|-----------------|
| Triangular unit | Normal vector | x_n, y_n, z_n |
| | First vertex | x_1, y_1, z_1 |
| | Second vertex | x_2, y_2, z_2 |
| | Third vertex | x_3, y_3, z_3 |

Matlab can extract the normal vector of all triangle patches and the coordinates of the three vertices that make up the patch of a STL model. The data needed for path planning are: the coordinates of every point, the normal vector of each triangle patch, and the correspondence between points and triangular patches.

The following data structures (matrices) are obtained:

Faces: In the form of $3 \times N_f$ matrix, N_f is the total number of all triangle patches, the row number is the index of the patch, and the three columns of each row are the projections of the unit normal vector of the polygon on the x , y and z axes.

Points: In the form of $[3 + M] \times N_p$ matrix, N_p is the total number of all non-repeating points, the row number is the index of the point, the first three columns of each row is the x , y and z coordinates of the points, M is the number of faces shared by most points, and the following columns are the index numbers of the faces shared by the points from front to back. For example, if a point is shared M_i ($M_i < M$) times, then from the 4-th column to the $(4 + M_i)$ -th column are the index number of the shared face, and the following $(M - M_i)$ columns are 0.

2.2 Path Planning Algorithm

Automated fiber placement “natural path” path planning algorithm mainly includes fixed angle method, approximate iterative method, meshing method, or triangular patch and the numerical solution using the characteristics of geodesic line as the basis for deformation characterization. By using a suitable curve to characterize the lateral extension of the prepreg in the placement surface during placement, a corresponding iterative algorithm is constructed to control and minimize the difference in arc length on both sides of the prepreg due to curvature changes, achieving the laying process without wrinkles and tearing of the prepreg.

Fixed Angle Method: The traditional fixed angle method uses a certain reference axis as a reference line, and through stepwise iteration, planning path is obtained after laying a reference line cluster on the laying surface at a certain angle with the reference axis. The method can well meet the direction consistency requirements of fiber

direction and designed direction. The axis or specific designed curve is usually selected as the reference line. The methods for solving the reference curve include slice round cut method, grid method, point-to-point geodesic line solution, etc.

Approximate Iteration Method: Ingersoll company of the United States proposed a relatively simple “natural path” trajectory planning method: the shape of the composite member for the actual automated laying is mostly a curved surface with a small curvature and a large size, and the fixed step is used to gradually extend the surface line to the surface to replace the solution of a curve on a surface. However, in the specific path planning process, the initial starting point of the laying is required, and a curve is placed along the laying surface as a reference. Due to the technical blockade, the definition of the curve is not described in detail in the relevant literature, so this method is not suitable for most researchers.

However, in the path planning algorithm research, when an algorithm needs to be verified, a reference curve can be given in ideal state, and the approximate iterative method planning result can be used as the standard path to provide reference.

3 Optimization Path Planning Algorithm

3.1 Algorithm Principle

Based on the fixed angle method described in Sect. 2.2, a new path planning algorithm is proposed. The specific steps are as follows:

Step one: Find a point A_i on the laying surface, draw the normal vector \vec{n}_i of the surface through A_i and create a tangent plane S perpendicular to \vec{n}_i .

Step two: Take the intersection of the x-axis and the tangent plane S as point B, and take a point C randomly on the x-axis. (The x-axis here, can be replaced by the y-axis or the z-axis, depends on which axis is more consistent to the mandrel axis.)

Step three: Draw a straight line through point C along the direction of \vec{n}_i , and the point of intersection of the straight line and the tangent plane S is point D.

Step four: Connect points B and D to get vector \overrightarrow{BD} , move the starting point of the vector \overrightarrow{BD} to point A_i to get vector $\overrightarrow{BD_A}$, counterclockwise rotate $\overrightarrow{BD_A}$ around \vec{n}_i to get vector $(\overrightarrow{BD_A})'$, and the rotation angle is θ .

Step five: Based on point A_i , take step size λ along the direction of the vector $(\overrightarrow{BD_A})'$ to get point A_i' .

Step six: Through point A_i' , make a straight line along the direction of \vec{n}_i , find A_{i+1} , the point of intersection of the line and the surface, it is the next point of A_i along the laying path.

The fixed angle method path calculation process is shown in Fig. 1:

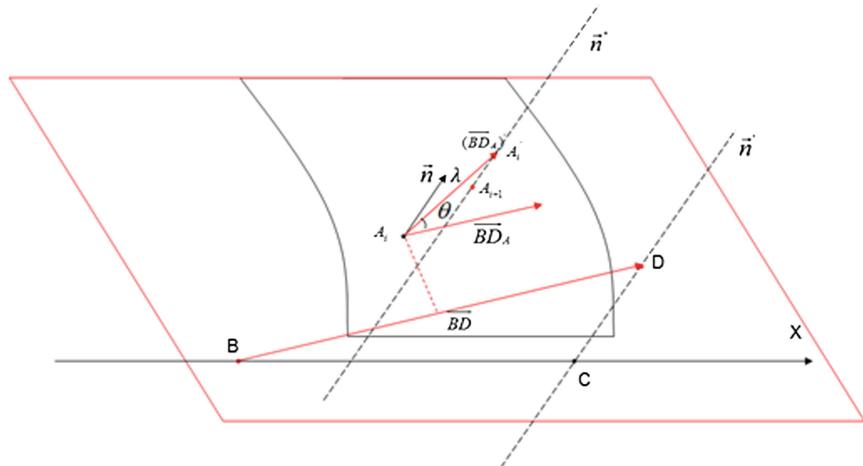


Fig. 1. Fixed angle method

3.2 Math Modeling

With the basis of data reconstruction in Sect. 2.1 and the algorithm principle in Sect. 3.1, the mathematical model of optimization algorithm is established. The solution process can be processed.

Assuming the index number of point A_i in data matrix is K, and its coordinate is (x_i, y_i, z_i) . From the information of the K-th row of the data matrix, it can be obtained which faces share the point numbered K and the average value of the normal vectors of these faces (n_{xi}, n_{yi}, n_{zi}) are taken as the normal vector \vec{n}_i of the point. Through this point, create a tangent plane S perpendicular to \vec{n}_i , the equation of the tangent plane can be obtained:

$$n_{xi}(x - x_i) + n_{yi}(y - y_i) + n_{zi}(z - z_i) = 0 \quad (1)$$

Take the point of intersection B of the x-axis and the tangent plane S, and solve the following equations:

$$\begin{cases} n_{xi}(x - x_i) + n_{yi}(y - y_i) + n_{zi}(z - z_i) = 0 \\ y = 0 \\ z = 0 \end{cases} \quad (2)$$

The coordinates of point B are (x_B, y_B, z_B) , and a point C is randomly taken on the x-axis. (The x-axis here, can be replaced by the y-axis or the z-axis, depends on which axis is more consistent with the mandrel axis.)

Draw a straight line through point C along the direction of \vec{n}_i :

$$\left\{ \begin{array}{l} n_{xi}(x - x_i) + n_{yi}(y - y_i) + n_{zi}(z - z_i) = 0 \\ \frac{x - x_i}{n_{xi}} = \frac{y}{n_{yi}} \\ \frac{y}{n_{yi}} = \frac{z}{n_{zi}} \end{array} \right. \quad (3)$$

The point of intersection of the straight line and the tangent plane S is D (x_D, y_D, z_D).

Connect points B and D to get vector \overrightarrow{BD} , move the starting point of the vector \overrightarrow{BD} to point A_i to get $\overrightarrow{BD_A}$, counterclockwise rotate $\overrightarrow{BD_A}$ around \vec{n}_i to get vector $(\overrightarrow{BD_A})' (n_{xi}^{(BD_A)'}, n_{zi}^{(BD_A)'}, n_{zi}^{(BD_A)'})$, and the rotation angle is θ .

Based on point A_i , take step size λ along the direction of the vector $(\overrightarrow{BD_A})'$ to get point $A'_i(x_i + \lambda n_{xi}^{(BD_A)'}, y_i + \lambda n_{yi}^{(BD_A)'}, z_i + \lambda n_{zi}^{(BD_A)'})$.

Through point A'_i , make a straight line along the direction of \vec{n}_i , the linear equation is:

$$\left\{ \begin{array}{l} \frac{x - (x_i + \lambda n_{xi}^{(BD_A)'})}{n_{xi}^{(BD_A)'}} = \frac{y - (y_i + \lambda n_{yi}^{(BD_A)'})}{n_{yi}^{(BD_A)'}} \\ \frac{y - (y_i + \lambda n_{yi}^{(BD_A)'})}{n_{yi}^{(BD_A)'}} = \frac{z - (z_i + \lambda n_{zi}^{(BD_A)'})}{n_{zi}^{(BD_A)'}} \end{array} \right. \quad (4)$$

Select 10 points on the surface where $(x_i + y_i + z_i)$ are relatively close to $(x_i + \lambda n_{xi}^{(BD_A)'}, y_i + \lambda n_{yi}^{(BD_A)'}, z_i + \lambda n_{zi}^{(BD_A)'})$, and calculate the distance between the points and the line obtained above. Select one point where the distance is the smallest as point A_{i+1} , and it is the next point of A_i on the laying path. Repeatedly, you can get a complete planning path.

4 Verification and Analysis

Figure 2(a) below shows the end of the fuselage (blue part) of the A380 aircraft model. Considering that this part is too large for the machine tool, select a small part of it, as shown in Fig. 2(b) below.

The model is meshed in the STL Rapid Prototyping module of CATIA with the parameter sag = 0.1 mm, step = 1 mm, and saved as STL file and imported into MATLAB. Use the path solving method described above, take 5 mm parameters step size and 45° reference angle in our calculation, the results are showing in Fig. 3:

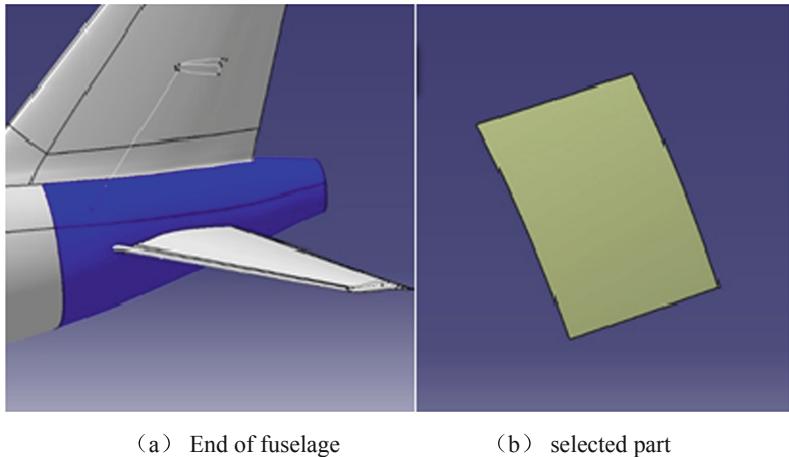


Fig. 2. Laying model

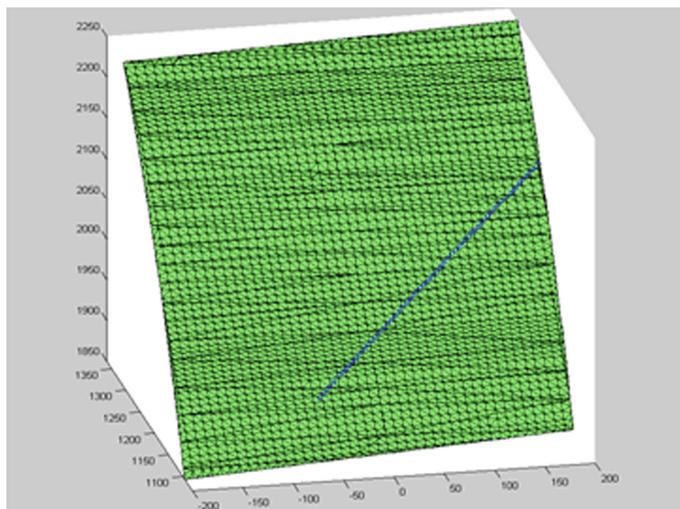


Fig. 3. Fixed angle method outputs

For better image observation, the parameters $sag = 0.1$ mm and $step = 10$ mm are set as the display grid and are not involved in the calculation. The blue * points in Fig. 3 are the points of the laying path calculated by fixed angle method.

In order to verify the accuracy of the solution results, the model file in Sect. 2.1 is still taken as an example. Given the ideal reference curve, the approximate iteration method is used to solve the natural path (see Fig. 4):

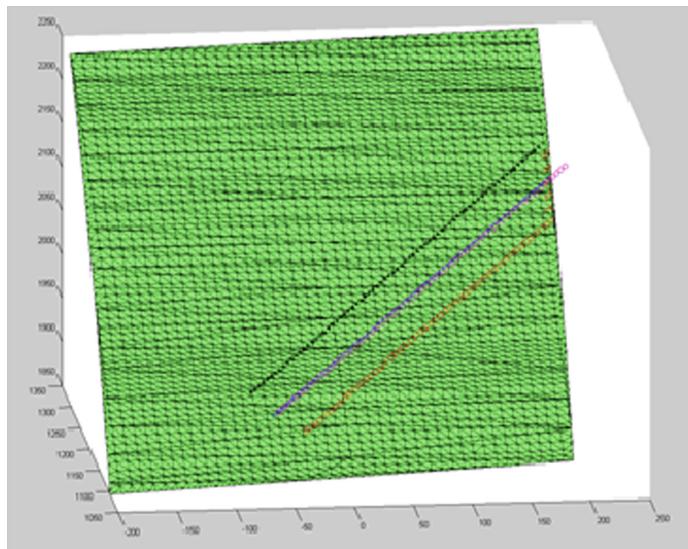


Fig. 4. Ideal natural path

In Fig. 4, black and red points are the borders, pink points are the ideal natural path, and blue points are the path solved by the fixed angle method.

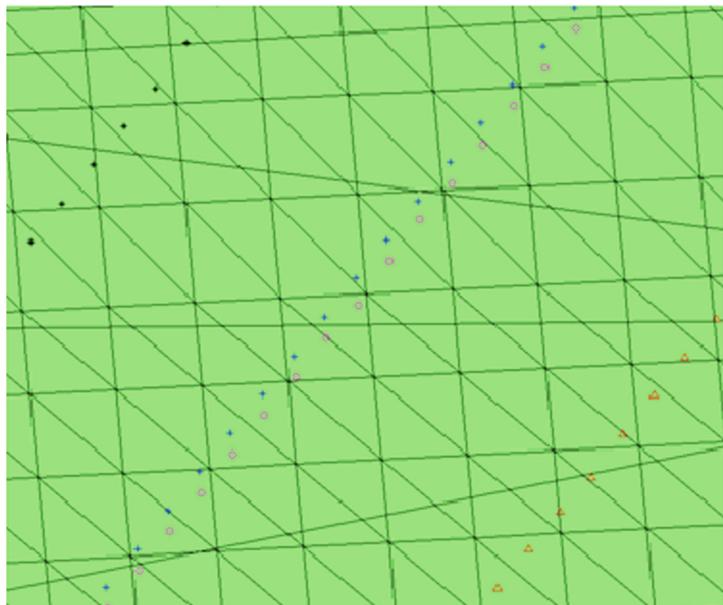


Fig. 5. Contrast display

It can be seen from Fig. 5 that in each 10 mm × 10 mm grid, the optimized fixed angle method and the ideal approximate iterative method are basically consistent under 3 mm precision. The subsequent trajectory can be obtained according to the trajectory parallel principle, and the feasibility of the algorithm is verified.

5 Conclusion

The data reconstruction of the STL model file transforms the boundary representation data structure into points and faces matrixes data which are easy to program, a general data reconstruction method is obtained. The existing fixed angle path planning algorithm is improved, the solution of the reference axis in the calculation is optimized, and the solution process is simplified. The mathematical model of the optimization algorithm is established and a simulation is proposed. The simulation results show that the algorithm is feasible and provides a basis for the subsequent systematic research of automated fiber placement.

References

- Yao J, Sun D, Yao Z et al (2011) Current situation and research progress of automated tape-laying technology for composites. *Mach Des Res* 27(4):60–65
- Schueler K, Miller J, Hale R (2004) Approximate geometric methods in application to the modeling of fiber placed composite structures. *J Comput Inf Sci Eng* 4:251–256
- Lewis HW, Romero JE (1987) Composite tape placement apparatus with natural path generation means. US Patent: 4696707
- Shinno N, Shigematsu T (1991) Method for controlling tape affixing direction of automatic tape affixing apparatus. US Patent: 5041179
- Hu C, Xiao J, Li Y et al (2007) Study on automated tape-laying technique for composites part I: natural path property analysis and calculation method. *Aerospace Technol* (1):40–43
- Wang S, Xiao J, Wu H (2007) Research of algorithm of geodesic lines involved in track programming of tape-laying. *Fiber Reinforced Plast Compos* (2):15–18
- Luo H, Li Y, Xiao J, Huan D (2009) Research on automatic tape-laying technique for composites—calculation method of tape-laying path on free-form surfaces. *Acta Aeronautica et Astronautica Sinica* 30(9):1782–1787
- Zhou Y (2006) Research on CAD of automatic composite fiber placement. NUAA, Nan Jing
- Wen L, Li J, Wang X, Xiao J (2013) Adjustment algorithm based on structural design for automated tape laying and automated fiber placement. *Acta Aeronautica et Astronautica Sinica* 34(7):1731–1739
- Koricho EG, Khomenko A, Fristedt T, Haq M (2015) Innovative tailored fiber placement technique for enhanced damage resistance in notched composite laminate. *Compos Struct* (120):378–385
- Hyer MW, Charette RF (1991) Use of curvilinear fiber format in composite structure design. *AIAA J* 29:1011–1015



Decoupling Design Based on Active Disturbance Rejection Control Approach for Near Space Vehicle Autopilot

Hui Li¹, Zhide Yin¹, Fengyi Li^{2(✉)}, and Hongyu Liu³

¹ Beijing Electro-Mechanical Institute, Beijing 100074, China

² Science and Technology on Complex System Control and Intelligent Agent Cooperation Laboratory, Beijing 100074, China
lh1999@sohu.com

³ Beijing Institute of Control and Electronic Technology, Beijing 100038, China

Abstract. A decoupled control approach based on linear active disturbance rejection control idea is proposed. The plant model, in which the coupling dynamics is as a part of disturbance, is established, and some virtual control variables are defined. Thus the original coupling plant is divided into some single-input single-output systems and the virtual control variables are willing to be determined by linear active disturbance rejection control method, and the real control variables are easily solved from virtual control variables according to their relationship. The approach is employed to a hypersonic near space vehicle lateral autopilot. The numerical simulations show that the proposed approach has good decoupling effectiveness and guarantees angle command signal tracking accuracy.

Keywords: Near Space Vehicle · Active Disturbance Rejection Control · Extended state observer · Decoupled control

1 Introduction

Near space is often, but not always, quantitatively defined as the range of Earth altitudes from 20 km to the “edge of space”, the Kármán line, at 100 km [1]. The Near Space Vehicle (NSV) has great strategic value in both military and civilian area and combines the advantages of traditional aircrafts and spacecrafts, the research on it attracts extensive attentions. In order to improve maneuverability and robustness of the NSV, the bank-to-turn (BTT) technique has been paid extensive investigations on the NSV autopilot design in recent decades.

With the use of BTT technique, the NSV needs a high roll rate and a small sideslip angle, furthermore, the NSV is generally designed to be the plane symmetrical type, thus the cross-coupling effect between the yaw channel and roll channel often appears in NSV. Maybe the coupling brings in some severe nonlinear characteristics and largely increases the difficulty of design of the NSV autopilot. Therefore, it is not appropriate any more to design the attitude controller individually for each independent channel ignoring the coupling effect. To address this issue, decoupled control has

received extensive researches during the last few decades. The pre-compensation method was to obtain block diagonal dominance in large-scale systems enabling the use of simpler control structure [2], but the structure of the pre-compensator has to be fixed in advance, which may lead to unnecessary dynamics. The modern frequency method, based on diagonal dominance and Nyquist stability criterion, was suitable for linear time-invariant multiple-input multiple-output (MIMO) systems [3]. Generally, the designed controller is strongly dependent on the transfer function of the plant, but it is hard to obtain the accurate model due to aerodynamic uncertainties of the NSV. Recently nonlinear control has been intensively studied on decoupled control issue. A nonlinear dynamic inversion method was used to design the decoupled controller [4], but robustness is not ensured in the case that the plant model cannot be accurately obtained. A sliding-mode variable structure controller, based on Lyapunov stabilization theory, was designed to achieve the BTT motion and guarantee the stability of the hypersonic vehicle [5], but the sliding mode controller is prone to give rise to high frequency vibration and large control quantity.

Active Disturbance Rejection Control (ADRC) is an adaptive nonlinear control firstly proposed by Han [6]. The ADRC method is similar to the conventional PID control method, which uses the error driven philosophy rather than model-based control law to eliminate errors. But ADRC can obtain better performance on disturbance rejection and dynamic performance than that of PID control, and realizes the idea of “small error using big gain and big error using small gain”. Although ADRC has been applied in different industrial control problems, the literatures on aircraft attitude control problem with this kind of approaches are really rare.

In this paper, a decoupled control approach based on the linear active disturbance control (LADRC) is studied. By establishing the coupling plant model taking the form of two-order derivative of system output, and defining some virtual control variables, the original coupling NSV plant is turn into some independent single-input single-out (SISO) systems. A linear extended state observer (LESO) is used to observe the total disturbance, which includes the coupling dynamics, furthermore, the conventional PD control is used in the LADRC for each SISO system, and the real control variables are accordingly obtained from virtual control variables.

2 NSV Lateral Motion Dynamics

The NSV lateral motion dynamics can be described using the short period approximation of the lateral motion equations. Based on the “small perturbation theory”, the mathematical model of lateral motion dynamics is written in state space equation notation as follows [7]

$$\begin{bmatrix} \dot{\omega}_x \\ \dot{\omega}_y \\ \dot{\beta} \\ \dot{\gamma} \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & b_{14} & 0 \\ 0 & b_{22} & b_{24} & 0 \\ \alpha & -(\alpha \tan \vartheta + b_{32}) & -(b_{34} - a_{33}) & -b_{36} \\ 1 & -\tan \vartheta & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \beta \\ \gamma \end{bmatrix} + \begin{bmatrix} b_{17} & b_{15} \\ b_{27} & b_{25} \\ 0 & -b_{35} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix} \quad (1)$$

Where ϑ is pitch angle, α is angle of attack, β is angle of sideslip, and γ is roll angle. ω_x and ω_y are NSV roll angular rate and yaw angular rate, respectively. δ_x and δ_y are fin deflections of roll channel and yaw channel, respectively. Parameters a_{33} , b_{11} , b_{14} , b_{15} , b_{17} , b_{22} , b_{24} , b_{25} , b_{27} , b_{32} , b_{34} , b_{35} and b_{36} are dynamics coefficients whose definitions and meanings are given in [7].

It is clear that the mathematical model described by Eq. (1) includes kinematic coupling, control coupling and aerodynamic cross-coupling. The design purpose is to stabilize the system and to achieve good performances in both angle of sideslip command responses and roll angle command responses. Since the NSV flies with great aerodynamic uncertainties and its lateral coupling dynamics is complex, the work on lateral autopilot design turns to be very difficult and boring, and the controller performance maybe deeply decade.

3 NSV Lateral Decoupled Control by ADRC

ADRC can be used to reduce the influence of the uncertainty and to make the system have good robustness. Here a decoupled control structure for the coupling dynamics of NSV lateral motion is proposed. A two-order derivative model on system output is established, and some virtual control variables are defined, furthermore, an LESO, proposed by Gao [8], is used to observe “total disturbance”, in which both the coupling and uncertainty are included. Then the conventional PD control is used to determine virtual control variables, and real fin deflections are obtained by the relationship between virtual control variables and real control variables.

Transform Eq. (1) into Eq. (2) with two-order derivative form

$$\begin{bmatrix} \ddot{\gamma} \\ \ddot{\beta} \end{bmatrix} = \begin{bmatrix} f_1(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y) \\ f_2(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y) \end{bmatrix} + B \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix} \\ = \begin{bmatrix} f_1(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y) \\ f_2(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y) \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix} \quad (2)$$

Where

$$\begin{aligned} f_1(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y) \\ = b_{11}\omega_x - \tan \vartheta b_{22}\omega_y + (b_{14} - \tan \vartheta b_{24})\beta \\ + [(b_{17} - \tan \vartheta b_{27}) - (b_{170} - \tan \vartheta b_{270})]\delta_x \\ + [(b_{15} - \tan \vartheta b_{25}) - (b_{150} - \tan \vartheta b_{250})]\delta_y \end{aligned}$$

$$\begin{aligned}
f_2(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y) &= [\alpha b_{11} - (b_{34} - a_{33})\alpha - b_{36}]\omega_x \\
&\quad + [-\alpha(\tan \vartheta + b_{32})b_{22} + (b_{34} - a_{33})\alpha(\tan \vartheta + b_{32}) + b_{36}\tan \vartheta]\omega_y \\
&\quad + [\alpha b_{14} - \alpha(\tan \vartheta + b_{32})b_{24} + (b_{34} - a_{33})^2]\beta + (b_{34} - a_{33})b_{36}\gamma - b_{35}b_{36}\dot{\vartheta}_y \\
&\quad + \{[\alpha b_{17} - \alpha(\tan \vartheta + b_{32})b_{27}] - [\alpha b_{170} - \alpha(\tan \vartheta + b_{320})b_{270}]\}\delta_x \\
&\quad + \{[\alpha b_{15} - \alpha(\tan \vartheta + b_{32})b_{25} + (b_{34} - a_{33})b_{35}] \\
&\quad \quad - [\alpha b_{150} - \alpha(\tan \vartheta + b_{320})b_{250} + (b_{340} - a_{330})b_{350}]\}\delta_y \\
B_{11} &= b_{170} - \tan \vartheta b_{270} \\
B_{12} &= b_{150} - \tan \vartheta b_{250} \\
B_{21} &= \alpha b_{170} - \alpha(\tan \vartheta + b_{320})b_{270} \\
B_{22} &= \alpha b_{150} - \alpha(\tan \vartheta + b_{320})b_{250} + (b_{340} - a_{330})b_{350}
\end{aligned}$$

Parameters a_{330} , b_{150} , b_{170} , b_{250} , b_{270} , b_{320} , b_{340} and b_{350} can be approximately fixed corresponding to parameters a_{33} , b_{15} , b_{17} , b_{25} , b_{27} , b_{32} , b_{34} and b_{35} , respectively.

Obviously, functions $f_1(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y)$ and $f_2(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y)$ include the coupling effect in NSV lateral dynamics.

Let

$$\mathbf{x} = \begin{bmatrix} \gamma \\ \beta \end{bmatrix}, \mathbf{f} = \begin{bmatrix} f_1(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y) \\ f_2(\omega_x, \omega_y, \beta, \gamma, \delta_x, \delta_y) \end{bmatrix}, \mathbf{u} = \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix}$$

Introduce virtual control vector

$$\mathbf{U} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \mathbf{B}\mathbf{u} \quad (3)$$

Where U_1 and U_2 are named as virtual control variables.

Usually, roll angle and angle of sideslip are anticipant outputs of the NSV lateral dynamics, then Eq. (2) turns to be

$$\begin{aligned}
\mathbf{x} &= \mathbf{f} + \mathbf{U} \\
\mathbf{y} &= \mathbf{x}
\end{aligned} \quad (4)$$

In system (4), the relationship between input and output of the i th channel can be expressed as

$$\begin{aligned}
\ddot{x}_i &= f_i + U_i \\
y_i &= x_i \\
i &= 1, 2
\end{aligned} \quad (5)$$

Therefore, each channel in Eq. (4) turns to be a corresponding SISO system.

Make $x_{i1} = x_i$ and make the extended state variable $x_{i3} = f_i$, and assume the derivative of the extended state variable as below

$$\dot{x}_{i3} = h_i \quad (6)$$

Where $h_i(t)$ is assumed to be a random signal.

Then the state space form of Eq. (5) is given by

$$\begin{aligned}\dot{x}_{i1} &= x_{i2} \\ \dot{x}_{i2} &= x_{i3} + b_{0i}U_i \\ \dot{x}_{i3} &= h_i \\ y_i &= x_{i1}\end{aligned}\quad (7)$$

where $b_{0i} = 1$.

A third-order LESO with corresponding to original third-order plant (7) can be designed as follow

$$\begin{cases} \dot{z}_{i1} = z_{i2} + l_{i1}(y_i - \hat{y}_i) \\ \dot{z}_{i2} = z_{i3} + b_{0i}U_i + l_{i2}(y_i - \hat{y}_i) \\ \dot{z}_{i3} = l_{i3}(y_i - \hat{y}_i) \\ \hat{y}_i = z_{i1} \end{cases} \quad (8)$$

Let $l_{i1} = 3\omega_{i0}$, $l_{i2} = 3\omega_{i0}^2$ and $l_{i3} = \omega_{i0}^3$, thus the characteristic roots of the LESO are uniformly valued ω_0 . For a well-tuned LESO, an estimate is obtained as $x_{i3} \approx z_{i3}$. If the virtual control variable takes the form

$$U_i = \frac{u_{0i} - z_{i3}}{b_{0i}} \quad (9)$$

Where u_{0i} is a virtual control output variable, too. Thus the original system may be approximately transformed into $\ddot{x}_{i1} = u_{0i}$.

u_{0i} may take the classical PD control form

$$u_{0i} = k_{pi}(y_i^* - y_i) - k_{di}z_{i2} \quad (10)$$

Then the virtual control signal is

$$U_i = \frac{k_{pi}(y_i^* - y_i) - k_{di}z_{i2} - z_{i3}}{b_{0i}} \quad (11)$$

Where y_i^* is the command signal of system (5), i.e., the roll angle command signal or angle of sideslip command signal.

From Eq. (7) and Eq. (11), we can approximately get

$$\ddot{y}_i + k_{di}\dot{y}_i + k_{pi}y_i = k_{pi}y_i^* \quad (12)$$

Then

$$\frac{y_i}{y_i^*} = \frac{k_{pi}}{s^2 + k_{di}s + k_{pi}} \quad (13)$$

The command signal will be reproduced only if the output y_i is measurable, thus the controlled output and the virtual control variable are completely decoupled. The control signal in Eq. (10) is named linear active disturbance rejection control (LADRC) for its purely linear form which is easy to conduct robustness analysis in frequency-domain.

If we design an ADRC controller for each channel in Eq. (5), the virtual control vector \mathbf{U} can be gotten. Hence the real control vector, i.e., the fin deflections, δ_x and δ_y , can be obtained by

$$\mathbf{u} = \mathbf{B}^{-1}\mathbf{U} \quad (14)$$

The block diagram of the decoupled LADRC is shown in Fig. 1.

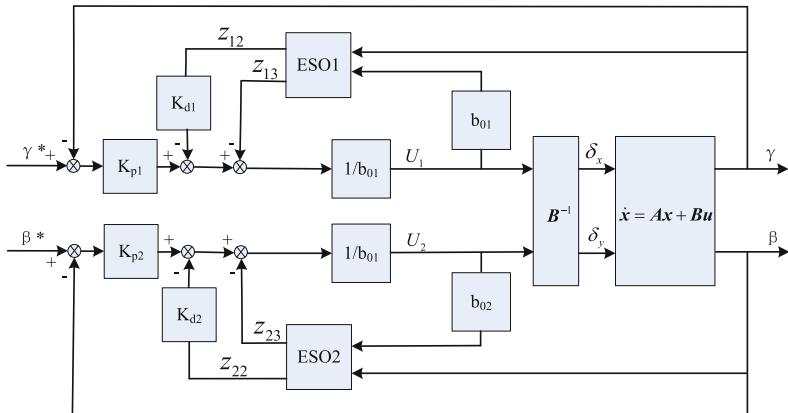


Fig. 1. The structure of decoupled LADRC algorithm for NSV lateral dynamics

Here f_i , including coupling effect and uncertainty, is treated as the “total disturbance”, which does not need to be expressively known in the process of LESO design. It can be concluded that LADRC is a PD control with observer compensation. The estimate of extended state counteracts the total disturbance of system, thus a majority of error is removed firstly, and then the residual, only a minority of error, is removed by PD control. Because f_i contains the perturbation in matrix \mathbf{B} and is able to be observed by LESO, matrix \mathbf{B} can bear some extent of uncertainty.

Therefore the job of decoupled LADRC controller design turns to be how to determine the parameters k_{pi} , k_{di} and ω_{0i} . Parameters k_{pi} and k_{di} reflect the response

velocity and the damping of the closed-loop system, respectively. The parameter ω_{0i} denotes the bandwidth of the LESO. The performance of the i th closed-loop channel can be easily guaranteed by specifying parameters k_{pi} , k_{di} and ω_{0i} .

From the job above, it can be concluded that decoupled LADRC artistically brings the capability of observation and compensation of ADRC into play. The decoupled LADRC is usually able to be realized by a set of fixed proportional parameters even though the total disturbance is very big.

4 Design and Simulation Results

To demonstrate the availability of the proposed approach, we use the same NSV model as the model in [9]. A hypersonic NSV glides from the height at 80 km to the height at 30 km. For the purpose of simplifying the study, we assumed that the initial roll angle, the initial angle of sideslip, and their initial error are zero degree. The trajectory is optimized based on the angle of attack and bank angle by an optimization method, which is not taken attention in this study.

Select six operating points along the geometric point trajectory, then six coupling dynamics systems of lateral dynamics are obtained. The vehicle dynamics are very different from each other due to its large envelope and wide speed range. A decoupled LADRC controller is designed for each dynamic system by the decoupled LADRC aforementioned. Then we make $k_{p1} = 2$, $k_{d1} = 4$, $\omega_{01} = 15$ and $k_{p2} = 2$, $k_{d2} = 4$, $\omega_{02} = 15$. In the case that a step signal is input into the roll channel, the angle of sideslip responses and the roll angle responses at specified operation points are illustrated in Figs. 2 and 3, respectively. In the case that a step signal is input into the yaw channel, the angle of sideslip responses and the roll responses at specified operation points are illustrated in Figs. 4 and 5, respectively. Obviously, the step responses, with corresponding to the channel inputting step signal, have well transient responses and are unanimously without the steady-state errors, furthermore, the step responses of the other channel are well suppressed.

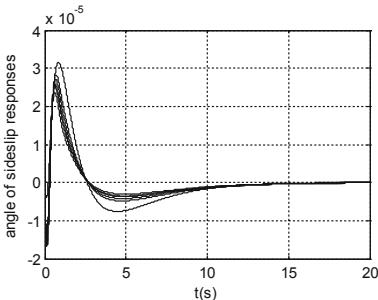


Fig. 2. Angle of sideslip responses at specified operation points by the decoupled LADRC approach when a step signal input into the roll channel

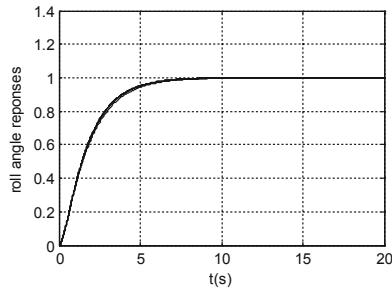


Fig. 3. Roll angle responses at specified operation points by the decoupled LADRC approach when a step signal input into the roll channel

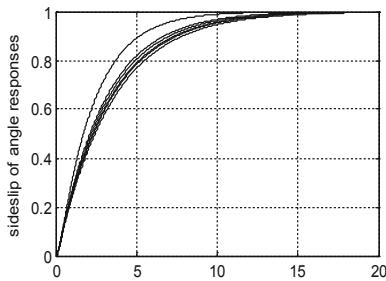


Fig. 4. Angle of sideslip responses at specified operation points by the decoupled LADRC approach when a step signal input into the yaw channel

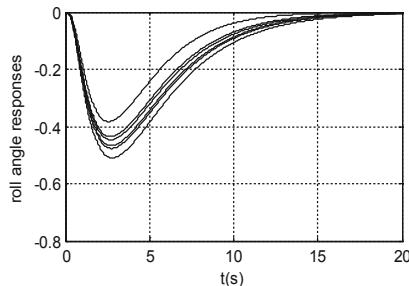


Fig. 5. Roll angle responses at specified operation points by the decoupled LADRC approach when a step signal input into the yaw channel

The decoupled LADRC controller is not difficult to be designed. The determination of LADRC parameters doesn't depend on the selected operating points, and good decoupling performance is realized only with a set of fixed parameters in the whole flight within a large envelope.

The results of tracking continuous command of both angle of sideslip and roll angle by the decoupled ADRC are illustrated in Figs. 6 and 7. They show that the decoupled LADRC law is very effective along the whole trajectory and the lateral dynamics is decoupled very well.

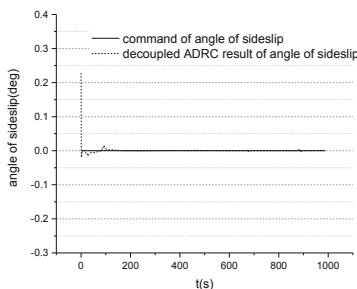


Fig. 6. History of angle of sideslip by the decoupled ADRC approach

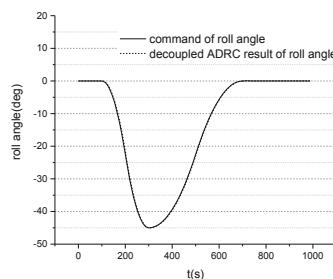


Fig. 7. History of roll angle by the decoupled ADRC approach

5 Conclusion

An LADRC approach to decouple the lateral dynamics is obtained for the hypersonic NSV lateral autopilot. The decoupling mechanism is very clear, and the ADRC parameters can be easily determined. The simulation for a hypersonic NSV flying

within a large envelope shows that the proposed approach can guarantee the decoupling performances only with a set of fixed parameters, and the continuous command of both angle of sideslip and roll angle along the trajectory are well tracked.

References

1. Young M, Keuth S (2009) An overview of advanced concepts for near-space systems. In: AIAA 2009, vol 4885, pp 1–18
2. Chen C, Munro N (1989) Procedure to achieve diagonal dominance using a PI/PID controller structure. *Int J Control* 50:1771–1792
3. Rosenbrock H (1969) Design of multivariable control systems using the inverse Nyquist array. *Proc Inst Electr Eng* 116:1929–1936
4. Snell S, Enns D, Garrard W et al (1992) Nonlinear inversion flight control for a supermaneuverable aircraft. *J Guid Control Dyn* 15:976–984
5. Xu H, Mirmirani M, Ioannou P (2004) Adaptive sliding mode control design for a hypersonic flight vehicle. *J Guid Control Dyn* 27:829–838
6. Han J (2009) From PID to active disturbance rejection control. *IEEE Trans Ind Electron* 56:1–7
7. Qian X, Lin R, Zhao Y (2000) Missile flight mechanics. Beijing Institute of Technology Press, Beijing (in Chinese)
8. Gao Z (2003) Scaling and bandwidth-parameterization based controller tuning. In: Proceedings of the American control conference, Denver, Colorado, pp 4989–4996
9. Li H, Yin Z, Fan Y et al (2015) A linear active disturbance rejection control approach for near space vehicle autopilot. In: 2015 IEEE the fourth international conference on control automation and information sciences, Suzhou, China, pp 330–335



Design of Speed Regulation of Ping-Pong Ball Ejector Based on RBF Neural Network PID Control

Min Xu^(✉) and Chaoli Wang

University of Shanghai for Science and Technology, Shanghai 200000, China
1881753723@163. com

Abstract. In order to keep the height of the table tennis ball machine from being stable, it is difficult to achieve fast and accurate speed regulation in the brushless DC motor control system for the traditional PID control. A method of controlling the RBF neural network PID is proposed. Based on the mathematical model of the brushless DC motor, the self-learning ability of the RBF neural network is used to adjust the parameters of the PID controller in real time to adjust the motor speed so that the table tennis ball-out mechanism is highly stable. The experimental results show that the motor control system based on the RBF neural network PID control strategy has stable start-up, good static and dynamic performance, and strong robustness, which can meet the requirements of high stability of the ball-out machine.

Keywords: Brushless DC Motor · Mathematical model · RBF neural network · Table tennis ball machine

1 Introduction

The invention of the table tennis ball machine has brought great convenience to the society, the school, the family and the professional training ground. It can be used for entertainment and fitness as well as for professional training. The ping-pong ball-off machine designed this time is different from the traditional ball-feeding machine. It needs to be combined with the mechanical arm. The ball is taken out by the ball-out machine. The robot arm swings and imitates people to issue various types of balls. Therefore, ping-pong is required. The ball is highly stable.

Brushless DC Motor (BLDCM) has become more and more widely used in defense, industrial control and household appliances because of its simple structure and high efficiency [1], so it has become the first choice for this table tennis ball machine motor. Conventional proportional-integral-derivative (PID) controllers are widely used in many control applications due to their simplicity and effectiveness, but the parameters of PID controllers are fixed and are not suitable for controlling complex nonlinear control systems, while brushless DC motor control systems. It is a multivariable, nonlinear, and strongly coupled system [2], which makes the table tennis ball machine susceptible to environmental influence under the traditional PID control. It often appears that the ball is highly unstable and difficult to adjust. Therefore, in order to

solve this problem, the motor control strategy must have strong adaptability and robustness.

Radial basis function (RBF) networks, as a kind of neural network, are often used in the control domain to form radial basis function neural network control [3]. The fuzzy PID control method proposed in [4–6] relies too much on the setting of the PID initial parameter value. Literature [7] proposed using BP network fuzzy control method to improve system stability, but the dynamic performance is not ideal. In this paper, the RBF network and the traditional PID control are combined to control the PID controller's control parameters in real time through the RBF network. It does not depend on the PID initial parameter setting, thus improving the performance of the control system and making the table tennis out. The ball machine is highly stable. And verify the effect through Matlab/Simulink simulation environment.

2 Mathematical Model of Brushless DC Motor

The brushless DC motor is composed of a motor body and a driver, and is a typical mechatronic product. In order to simplify the analysis, a reasonable assumption is made for the brushless DC motor [8]:

- (1) Ignore motor core saturation, ignoring eddy current and hysteresis loss;
- (2) Excluding the armature reaction, the air gap magnetic field is approximately considered to be a trapezoidal wave with a flat top width of 120° electrical angle;
- (3) Ignoring the cogging effect, the armature conductor is continuously and uniformly distributed on the armature surface;
- (4) The power tube and freewheeling diode of the drive system inverter circuit have ideal switching characteristics.

The equivalent circuit diagram of the inverter circuit of the brushless DC motor is shown in Fig. 1.

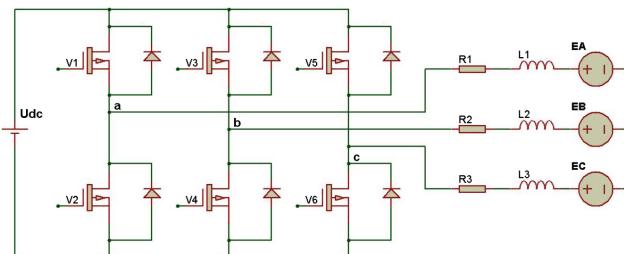


Fig. 1. Equivalent circuit diagram of BLDCM

Therefore, the matrix form of the phase voltage equation of the brushless DC motor can be expressed as:

$$\begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} = \begin{bmatrix} R & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & R \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \begin{bmatrix} L - M & 0 & 0 \\ 0 & L - M & 0 \\ 0 & 0 & L - M \end{bmatrix} \frac{d}{dt} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \begin{bmatrix} e_a \\ e_b \\ e_c \end{bmatrix} \quad (1)$$

$$i_a + i_b + i_c = 0 \quad (2)$$

Where: u_a, u_b, u_c are the stator winding phase voltage; R is the resistance of stator winding; i_a, i_b, i_c are the stator winding phase currents; L is the self-inductance of each phase winding of stator; M is mutual inductance between stator two-phase windings; e_a, e_b, e_c are the counter electromotive force of stator phase winding.

It can be seen from the expression of the electromagnetic torque of the motor that the electromagnetic torque of the motor is proportional to the flux and the current amplitude, so controlling the amplitude of the square wave current output by the inverter can control the electromagnetic torque [9]. The equation of motion of the motor is:

$$J \frac{d\omega}{dt} = T_e - T_L - B\omega \quad (3)$$

Where: ω is the mechanical angular velocity of the rotor; T_L is the load torque; B is viscous damping coefficient; J is the moment of inertia of the rotor and the load.

3 The Traditional PID Control BLDCM Speed Control Principle

The block diagram of the PID control system is shown in Fig. 2 [2].

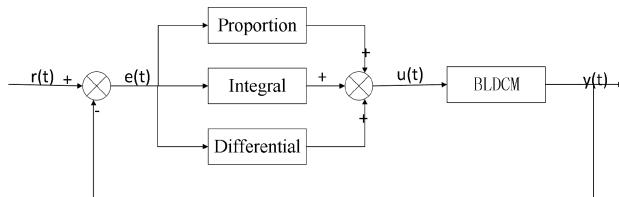


Fig. 2. PID control block diagram

The traditional PID control formula is:

$$u(k) = K_P e(k) + K_I \sum_{j=0}^k e(j) + K_D [e(k) - e(k-1)] \quad (4)$$

Can be obtained by:

$$\begin{aligned}\Delta u(k) &= u(k) - u(k-1) \\ &= K_p[e(k) - e(k-1)] + K_I e(k) + K_D[e(k) - 2e(k-1) + e(k-2)]\end{aligned}\quad (5)$$

Where: $e(k) = r(k) - y(k)$ is error; $r(k)$ is the given value; $y(k)$ is the actual output value.

Traditional PID control has better control effects in linear systems and simple non-linear systems [8]. In the ideal state, the motor in the table tennis ball-out machine can use this control strategy to respond well to the input signal and run smoothly. In fact, the motor has over-speed regulation during the starting process, and there are often various uncertain factors in the operation process. The trial-and-error method of the traditional PID control system design has the accuracy problem in the actual motor control process. Therefore, the motor speed cannot be maintained in a smooth running state.

It can be seen from Fig. 3 that a speed error is obtained by subtracting the feedback speed from a given speed. The speed error passes through the speed controller (PID regulator), outputs a value to the current loop for a given current, and then subtracts the feedback from the given current. The current gets a current error that passes through the current controller (PID regulator) and outputs a value, the duty cycle.

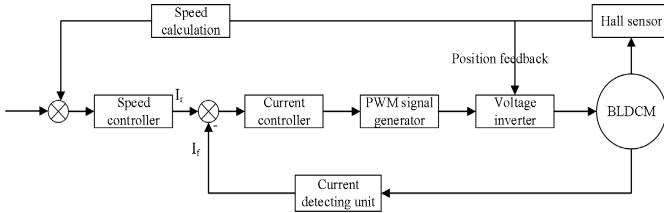


Fig. 3. Basic block diagram of double closed loop control of BLDCM

The PWM controller generates a pulse width modulation signal of equal amplitude, equal pulse width and constant period to control the opening and closing of the switching tube in the inverter circuit. When the duty ratio is large, the armature voltage is high and the winding current is large; otherwise, it is small. Since the magnitude of the electromagnetic torque is proportional to the current, closed-loop control of the torque and thus the speed can be achieved.

4 Parameter Self-tuning PID Control Based on RBF Neural Network

4.1 Neural Network Learning Algorithm

The RBF neural network is consistent with the function approximation theory. It is suitable for multivariate function approximation. As long as the central point set is properly selected, only a few neurons can obtain good approximation effects, and it

also has the advantage of the only optimal approximation point [8]. The neural network is a three-layer forward network, including an input layer, an implicit layer, and an output layer, as shown in Fig. 4 (the 3-6-1 structure is selected).

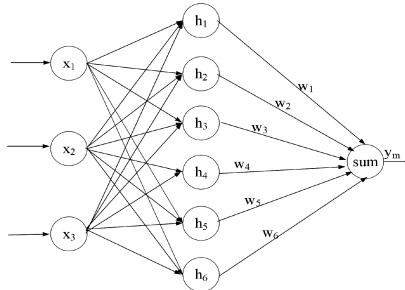


Fig. 4. RBF network structure diagram

Since the structure of the Gaussian function is simple, better analysis results can be obtained when performing analysis processing. Handling such an RBF network does not increase the computational complexity in the case of a large number of variable inputs. The Gaussian function is as follows:

$$h_j = \exp\left(-\frac{\|X - C_j\|^2}{2b_j^2}\right) \quad (j = 1, 2, \dots, 6) \quad (6)$$

Where: $C_j = [c_{j1}, c_{j2}, c_{j3}]^T$ is the center vector of the forth node of the network, $X = [x_1, x_2, x_3]^T$ is the input vector, the basis width vector is $B = [b_j]^T$, b_j is the basis width parameter of the node, and $b_j > 0$.

The weight vector of the network is:

$$W = [\omega_j]^T, j = 1, 2, \dots, 6 \quad (7)$$

Identify network output is:

$$y_m(k) = \sum_{j=1}^6 \omega_j(k) h_j(k) \quad (8)$$

Identifier performance indicator function is:

$$J(k) = \frac{1}{2} (y(k) - y_m(k))^2 \quad (9)$$

According to the network model of the data, the iterative method is used to calculate the parameter values as follows:

$$\omega_j(k) = \omega_j(k-1) + \eta(y(k) - y_m(k))h_j(k) + \alpha(\omega_j(k-1) - \omega_j(k-2)) \quad (10)$$

$$\Delta b_j(k) = (y(k) - y_m(k))\omega_j(k)h_j(k) \frac{\|X - C_j\|^2}{b_j(k)^3} \quad (11)$$

$$b_j(k) = b_j(k-1) + \eta\Delta b_j(k) + \alpha(b_j(k-1) - b_j(k-2)) \quad (12)$$

$$\Delta c_{ji}(k) = (y(k) - y_m(k))\omega_j(k) \frac{x_j(k) - c_{ji}(k)}{b_j(k)^2} \quad (13)$$

$$c_{ji}(k) = c_{ji}(k-1) + \eta\Delta c_{ji}(k) + \alpha(c_{ji}(k-1) - c_{ji}(k-2)) \quad (14)$$

Where η is the learning rate and α is the momentum factor.

Jacobian array is used to detect the output and input of the controlled system in real time, so as to compare the two values and determine the control sensitivity of the system. The Jacobian array algorithm is as follows:

$$\frac{\partial y(k)}{\partial \Delta u(k)} \approx \frac{\partial y_m(k)}{\partial \Delta u(k)} = \sum_{j=1}^m \omega_j(k)h_j(k) \frac{c_{ji}(k)}{b_j(k)^2} \quad (15)$$

4.2 RBF Network PID Parameter Tuning

Three inputs of PID as follow:

$$xc(1) = e(k) - e(k-1) \quad (16)$$

$$xc(2) = e(k) \quad (17)$$

$$xc(3) = e(k) - 2e(k-1) + e(k-2) \quad (18)$$

The control algorithm is:

$$\Delta u(k) = k_p xc(1) + k_i xc(2) + k_d xc(3) \quad (19)$$

The setting index of the neural network is:

$$E(k) = \frac{1}{2}e(k)^2 \quad (20)$$

Three parameters of PID controller used to adjust the gradient method:

$$\Delta k_p = -\eta \frac{\partial E}{\partial k_p} = -\frac{\partial E}{\partial y} \frac{\partial y}{\partial \Delta u} \frac{\partial \Delta u}{\partial k_p} = \eta e(k) \frac{\partial y}{\partial \Delta u} xc(1) \quad (21)$$

$$\Delta k_i = -\eta \frac{\partial E}{\partial k_i} = -\frac{\partial E}{\partial y} \frac{\partial y}{\partial \Delta u} \frac{\partial \Delta u}{\partial k_i} = \eta e(k) \frac{\partial y}{\partial \Delta u} x(2) \quad (22)$$

$$\Delta k_d = -\eta \frac{\partial E}{\partial k_d} = -\frac{\partial E}{\partial y} \frac{\partial y}{\partial \Delta u} \frac{\partial \Delta u}{\partial k_d} = \eta e(k) \frac{\partial y}{\partial \Delta u} x(3) \quad (23)$$

The structure of RBF network tuning PID control system is shown in Fig. 5.

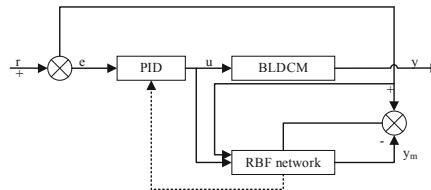


Fig. 5. RBF network tuning PID control block diagram

In Fig. 5, the current loop uses a PI regulator to improve the dynamic response capability, the current loop is corrected to a typical I-type system, and the speed loop uses the RBF neural network PID controller, and the PID parameters are determined by the RBF network. Improve the robustness and anti-interference ability of the motor, so that the motor in the table tennis ball machine can run stably.

5 Analysis of Experimental Results

In order to deeply study the motor drive system of table tennis server, this paper establishes a brushless DC motor simulation model with RBF neural network PID control in this environment, and compares it with the traditional PID control model to test the rationality and reliability of the whole control system.

The parameters of the brushless DC motor are as follows: the rated voltage of the motor is $U = 48$ V, the power is $P = 80$ W, the number of poles is $n_p = 3$, the simulation time is 0.2 s, the starting load of the motor is $T = 0.136$ Nm, and the rated speed is 1500 rpm.

Figures 6 and 7 show the motor speed simulation curve, which adopts PID control and RBF neural network PID control. By comparing the two curves, it can be found that in the RBF neural network PID control mode, the system response speed is faster and the speed is super. The adjustment is small, after the speed is stable, the simulation curve is smoother, and the better speed regulation effect can be obtained, which basically meets the requirements of the ball-out speed and high stability of the table tennis ball-out mechanism.

When the motor speed runs at a constant speed of 1500 rpm and 10% transient load disturbance is suddenly added at $t = 0.1$ s, the motor speed characteristic curve is shown in Figs. 8 and 9.

From the simulation curves of Figs. 8 and 9, when the load disturbance of the motor occurs, the RBF neural network PID control can adaptable learn according to the

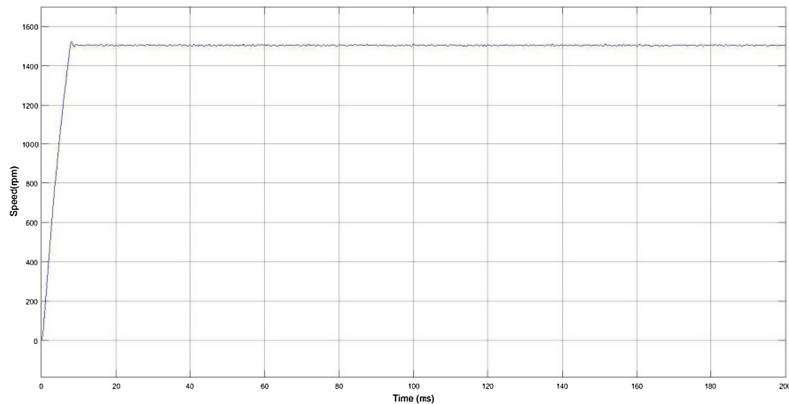


Fig. 6. Traditional PID control simulation curve of BLDCM

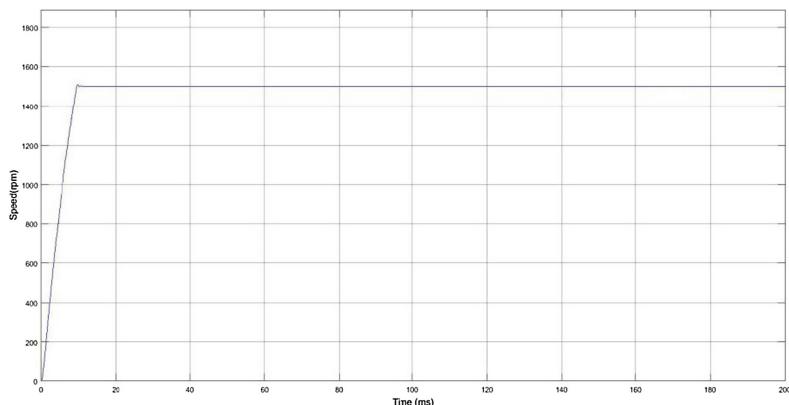


Fig. 7. RBF neural network PID control speed simulation curve of BLDCM

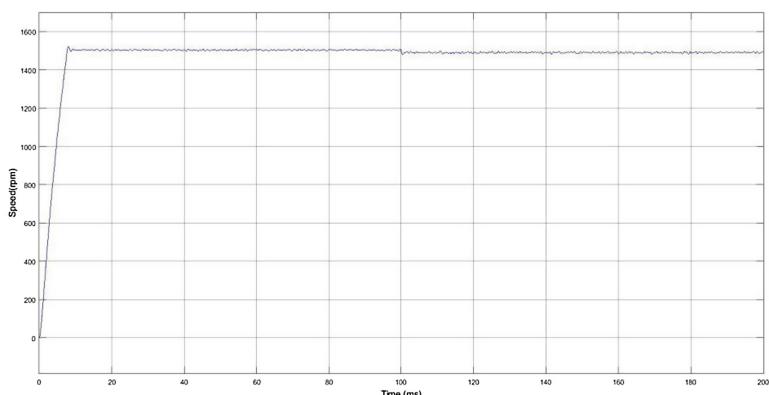


Fig. 8. PID control sudden load speed recovery characteristic curve

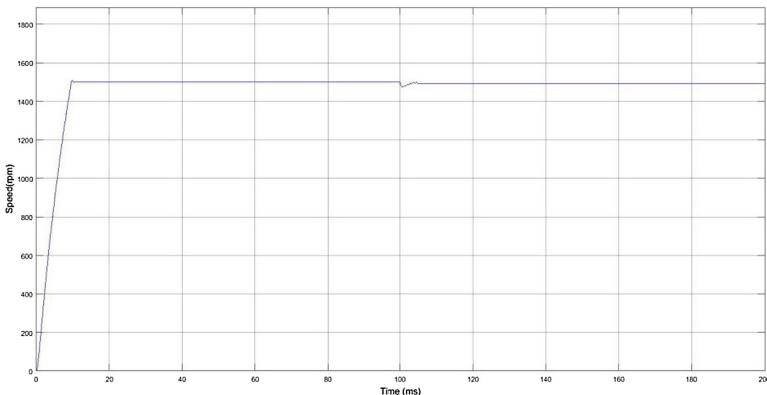


Fig. 9. RBF neural network PID control sudden load speed recovery characteristic curve

environmental change, adjust the PID parameters in real time through the environmental change, adjust the PID parameters in real time through the change of the environment, and improve the system adaptive ability so that it can overcome the external disturbance causes the motor to quickly return to a stable operating state.

6 Conclusion

For the conventional PID control method, it is unable to adapt to the nonlinearity, time-varying, uncertainty of the double-closed-loop DC speed regulation system of the brushless DC motor in the table tennis ball-out mechanism, and the over-speed of the motor during the starting process, it is difficult to suppress the parameter change. The application of the RBF neural network PID control to the system, using the self-learning ability of the RBF neural network, continuously adjust the proportional, integral and differential control parameters of the PID controller in real time, realizing the rapid start of the motor and effectively suppressing external interference. Make the table tennis ball machine highly stable. The simulation results show that the control strategy of using the neural network to control the motor speed of the table tennis ball machine has strong self-learning, self-adaptive ability and strong robustness, which has obvious advantages over the traditional PID double closed-loop control system.

References

- 夏长亮 (2009) 无刷直流电机控制系统. 科学出版社
- Fan X, Meng F, Fu C et al (2009) Research of brushless DC motor simulation system based on RBF-PID algorithm. In: 2009 second international symposium on knowledge acquisition and modeling. IEEE Computer Society
- 裴玉兵, 朱学来 (2017) 基于 RBF-PID 控制的直流调速系统仿真. 盐城工学院学报(自然科学版) (4):21–25
- 王晓近, 傅涛 (2015) 基于模糊 RBF 神经网络的无刷直流电机控制. 微电机 48(11)

5. Cheng J, Zhang G, Lu C, et al (2017) Research of brushless DC motor control system based on RBF neural network. In: Automation. IEEE
6. Cao SY, Tang WJ (2018) 2018 37th Chinese control conference (CCC) - speed control system based on fuzzy neural network of BLDCM, Wuhan, China, 25–27 July 2018, pp 3295–3297
7. Cao SY, Tang WJ. Speed control system based on fuzzy neural network of BLDCM. In: 第 37 届中国控制会议
8. 马晓爽, 石征锦 (2016) 基于 Simulink 的无刷直流电机双闭环调速系统仿真研究. 制造业自动化 38(7):82–88
9. Rank E (2003) Application of Bayesian trained RBF networks to nonlinear time-series modeling. Signal Process 83(7):1393–1410



Structural Analyses of Chinese Passenger Airline Network from the Perspective of Multi-layer Networks

Hao Wang¹, Shiwen Sun¹, Li Wang¹, and Chengyi Xia^{1,2(✉)}

¹ Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Tianjin University of Technology, Tianjin 300384, People's Republic of China
cyxia@email.tjut.edu.cn, xialooking@163.com

² Key Laboratory of Computer Vision and System (Ministry of Education), Tianjin University of Technology, Tianjin 300384, People's Republic of China

Abstract. By utilizing the theory of complex network, we investigate the structural features of China's passenger ATNs (Air Transport Networks) from the perspective of multi-layer networks. The results indicate that with the superposition of the ATNs of four major airline groups, the aggregated ATNs will gradually present the features of high clustering coefficient, short average path length and high average degree. We also find that when the number of node classes in the aggregation ATNs changes, the variation range of the structural properties of the aggregate ATNs exhibit the different behaviors in the process of superposition. Finally, by adding other small airlines into the above-mentioned aggregated ATNs, it is found that the high clustering coefficient of China's ATNs and the uneven load are mainly determined by the ATN of four major airline groups, but the average degree of the network is greatly affected by the small airlines ATN. The current findings are beneficial for us to understand the structural properties of Chinese ATNs and further improve the design of domestic flight routs in the future.

Keywords: Multi-layer networks · Structural properties · Air transportation networks

1 Introduction

As the traffic system becomes more and more important in the modern society, traffic problems have aroused the wide attention in theoretical and experimental researches. With the acceleration of globalization, compared with other transportation systems, air transportation system plays an increasingly key role in local, national and international economy due to its features of high speed, high safety and low transportation cost. Over the past two decades, complex network theory has been widely used to analyze different modes of transportation. Among them, several studies [1] have utilized nodes to represent airports and edges to denote flights between airports to reveal the topology

and dynamic features of ATNs. Since the 20th century, with the development of air transport industry, the number of airports and flights has increased rapidly, which created many complex ATNs of different size. Therefore, the analysis of the features of ATN in different spatial and temporal scales has attracted much attention. For instance, Refs. [2–4] studied the global ATNs, while ATNs in Europe [5], North America [5–8], China [9–16] and other regions [17, 18] have also been studied in depth, to name but a few. It is uncovered that most ATNs hold the small-world properties, degree heterogeneity and negative correlation characteristics and so on.

However, previous researches have often ignored the multi-layer nature of real air transport systems. Based on the multi-layer network theory, Cardillo et al. [19] established the European ATNs with multiple layers, which divide 15 European largest airlines into 15 layers; Verma et al. classified the global ATNs into three layers based on k-core decomposition method [20]; These works provide a new angle to analyze the structural properties of ATNs. At present, China's air transport industry experienced a very rapid development in the past decades, and ATNs consist of airport nodes with different flight zone classes and flight routes of many companies. Obviously, the multi-layer network model is more suitable for the analysis of China's ATNs since it can more accurately define the different dynamic development of each layer of the complex system.

To this end, we analyze the air transport industry related data and attempt to characterize the Chinese ATNs through the multi-layer networks, found that although Chinese ATNs are composed of flight routes of nearly 40 airlines, most of them are controlled by China Eastern airlines, Air China, HNA airlines and China southern airlines. Meanwhile, these four major airline groups account for nearly 80% of the total number of flights in China. In addition, it is also found that there exist some differences in the air transport structure between airline groups and between these airline groups and other small airlines. In what follows, we firstly describe some frequently-used quantities related with the ATNs; Then, we present some analyses of structural properties of Chinese ATNs from the perspective of multi-layer networks. Finally, we end this paper with some concluding remarks. Under this framework, we hope that the structural parameter analysis of this network model can help to thoroughly understand the route distribution law of China's ATN, so as to manage the network more effectively, improve the network performance, and even enhance the air transport efficiency and flight schedule optimization. Meanwhile, we hope that by more accurately grasping the structural differences of ATNs of various airline groups in China so that the further researches can be inspired on some subtle practical issues.

2 Data Collection and Related Measures

2.1 Data Collection on the Chinese ATNs

We crawled the data of China's ATNs from January 22, 2018 to January 28, 2018 from ctrip.com (<http://www.ctrip.com>). We apply these data to build an unweighted and

undirected Chinese ATNs, in which the airport acts as a node (cities with multiple airports are represented by a node), and the edge means the direct flight between two airports (the ATNs are set to be undirected). In addition, we only consider the flights from mainland China (departure and destination airports outside mainland China will not be considered).

As mentioned above, the number of flights of four major airline groups accounted for nearly 80% of the total number of flights during this period. Thus, we mainly took four major airline groups in China (i.e., China Eastern airlines, Air China, HNA airlines and China southern airlines) as four ATNs graded on the node dimension for analysis. According to the airport flight zone rating index, the airport nodes of each network are divided into five classes (i.e., 4F, 4E, 4D, 4C and 3C from high to low in order). In the following section, we mark the flight zone class of each airport as the corresponding node level of the airport in the network, based on which all nodes in the network are classified. In addition, if a city has multiple airports with different flight zone classes, the highest class will be taken as the node class of the city airport. The class of an airport's flight zone is determined by two factors: one is the length of the aircraft reference field length (i.e., runway length); The other one includes two main parameters of the aircraft, that is, the wingspan and the aircraft outer main gear wheel span. The airport with the higher class can support the takeoff and landing of the larger aircraft. It also means higher rated airports, which are often the backbone of the entire air transport network, will accommodate more passengers and have more influence.

To explore the structural properties of China's major airlines, we first constructed an ATN with only 4F-class nodes, and then added other low-grade airport nodes step by step to form the ATNs of four major airline groups with node hierarchical structure. Next, we aggregate the single-layer networks of the airline groups so that ATNs are formed by the aggregation of multiple single-layer networks with hierarchical structure. Finally, we will classify the nodes in the ATN composed of routes of other small airlines and aggregate them into the above multi-layer network as the last layer to form the whole Chinese ATNs.

2.2 Network Structure Properties

We select several commonly used properties in the study of complex network to accurately describe the structural features of the ATNs, which include as follows.

- (1) **The cumulative distribution $p(k)$:** $p(k)$ denotes the probability of finding a node with a degree value not less than k , which reflects the structural heterogeneity of ATN [9, 21–23].
- (2) **The average degree $\langle k \rangle$:** The average degree of all nodes $\langle k \rangle$ is used to reflect the average number of routes among all airports in the network.
- (3) **The clustering coefficient C :** The clustering coefficient of a single node is defined as $C_i = \frac{2e_i}{k_i(k_i-1)}$, where e_i is the actual number of edges between neighbors of the

node i , and k_i is the degree of the node i . The clustering coefficient $C = \sum_{i \in N} C_i / N$ represents the average clustering coefficient in the network. Since the clustering coefficient reflects the density of triangles formed between nodes in the network, we use it to estimate the ease of flight transfer and the possibility of round-trip travel with dual destinations [23–26].

- (4) **The average shortest path length $\langle L \rangle$:** The average shortest path length describes the average length of the shortest paths between all node pairs in the network [17], which means the average number of hops between two airports.
- (5) **The network assortativity coefficient r :** r is usually used to denote the degree correlation of all nodes in a complex network, and r is defined as

$$r = \frac{M^{-1} \sum_i j_i k_i - [M^{-1} \sum_i \frac{1}{2} (j_i + k_i)]^2}{M^{-1} \sum_i \frac{1}{2} (j_i^2 + k_i^2) - [M^{-1} \sum_i \frac{1}{2} (j_i + k_i)]^2}, \quad (1)$$

where j_i and k_i are the degree of two nodes at the end of the i -th edge ($i = 1, \dots, M$, M is the total number of edges). The value of r lies in the interval $[-1, 1]$. $r > 0$ implies the assortative networks, but $r < 0$ denotes disassortative networks. Here, r is used to measure the assortative mixing level of each ATN.

- (6) **The network polarization π :** The network polarization is a measure related with the betweenness of nodes and can represent the load homogeneity of the network, and can be defined as follows [9]

$$\pi = \frac{B_{\max} - \langle B \rangle}{\langle B \rangle} \quad (2)$$

where B_{\max} is the maximum betweenness of the network, and $\langle B \rangle$ is the average betweenness [9]. As the value of π increases, the network becomes more and more inhomogeneous.

3 Simulation Results and Discussion

3.1 Chinese ATNs with Node Classification for Major Airline Groups

At first, in Fig. 1, we show the gradual expansion of the ATNs of China Eastern Airlines, Air China, HNA (Hainan) Airlines and China Southern Airlines after they join other nodes from the high-class to low-class on the basis of the most core ATN with only 4F-class nodes. The airports with the highest class of flight zones are observed to be concentrated on the densely populated southeastern part of China. They are both economically developed cities or the popular tourist destinations in China. Moreover, the core ATNs of the four major airline groups are all strongly connected.

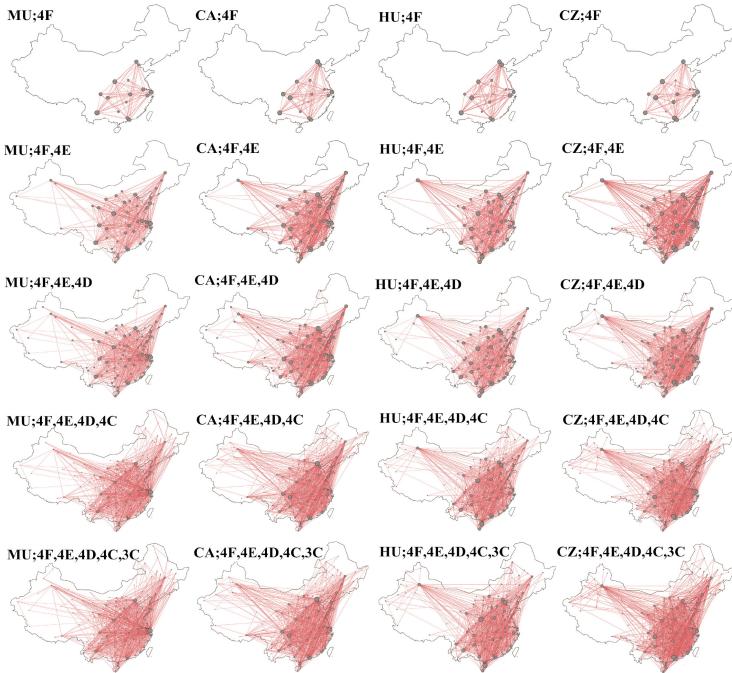


Fig. 1. Visualization of ATNs of each airline groups from high class to low class by adding nodes of other classes based on the most core ATN with only 4F-class nodes. The figure indicates the airline group to which the network belongs and all classes of nodes added in the current network (MU represents China Eastern Airlines group; CA is Air China group; HU for HNA group; CZ stands for China Southern Airlines group).

Secondly, we present the analysis of above-mentioned 6 significant quantities for four major airline groups in Fig. 2. Among them, Fig. 2(a) shows the cumulative distribution $p(k)$ of the ATNs containing all categories of nodes. It can be clearly seen that the ATNs of all groups exhibit the highly heterogeneous distribution. Figure 2(b) presents the average degree of the ATNs of each group as the number of node classes X added into the network increases. As an example, $X = 1$ means that the network just include the 4F-class airports, while $X = 2$ includes the 4F-class and 4E-class airports, and so on. Overall, the average degree of all groups reach a peak after 4E-class nodes were added, and then begin to decline since the core ATN of each group is mainly composed of 4F-class and 4E-class airports located in China's first and second-tier cities. In contrast, nodes of other classes are mostly low-degree nodes, and these nodes are only regional hubs or peripheral airport nodes that do not undertake hub tasks.

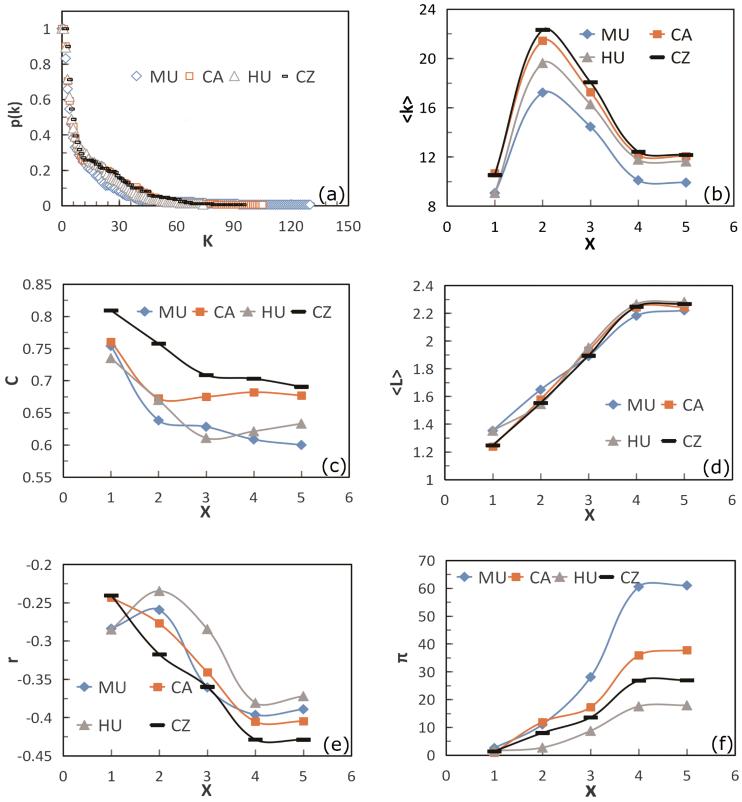


Fig. 2. The structural features of the ATNs of four major Chinese airline groups. In panel (a), K denotes the node degree; From panel (b)–(f), X represents the number of airport node classes contained in the ATN at this time, and vertical axis characterize the related quantities.

In addition, Fig. 2(c) pictures the clustering coefficient C as a function of X . In terms of the absolute value, the clustering coefficient of China Southern Airlines is significantly higher than other groups. This indicates that, in the event of an obstruction to the direct flight of China Southern Airlines, the transfer scheme could be used to link up other airports more flexibly. In addition, with the addition of nodes of different classes in the network, the variation trend of clustering coefficient of various airline groups is very different. It is speculated that there is a big difference in the strategy of each airline group to allocate flight routes between airports at all classes. In Fig. 2(d), with the addition of nodes of different classes in the network, the average shortest path length $\langle L \rangle$ of the four airline groups also increases gradually. Of course, the average shortest path length $\langle L \rangle$ of each airline group is relatively close under different values of X .

On the one hand, it can be clearly seen from Fig. 2(e) that the network assortativity coefficient r is less than 0 regardless of the value of X , which means the disassortative property of all ATNs. On the other hand, the correlation between r and X is not monotonic. It is worth noting that the network assortativity coefficient of China Eastern Airlines and HNA airlines reached the peak when the 4E-class nodes joined the network, while the assortativity coefficient of Air China and China Southern Airlines both decreased at this stage. Meanwhile, it can be observed in Fig. 2(f) that the value of network polarization π increases with the increase of the X , indicating that with the addition of nodes at all classes in the network, the load of the ATNs become more and more inhomogeneous. In particular, the changes in the polarization of the ATNs of China Eastern Airlines are the most obvious.

3.2 Multi-layered Aggregate ATN with Node Classification of China's Major Airline Groups

In order to further explore the structural nature of the ATNs, we aggregated the single-layer network of each group with node classification. The ATNs are superimposed in the order of HNA airlines, China Eastern Airlines, China Southern Airlines and Air China group (according to the order of the total number of network nodes in each group) to construct the multi-layered aggregated ATNs.

Figure 3 illustrates the structural properties of the aggregate multi-layer ATNs. Among them, Fig. 3(a) shows the trend of the cumulative distribution $p(k)$ of the multi-layered aggregate ATN as the number X of node classes in the network increases. This indicates that with the addition of nodes of different classes in the aggregation network, the aggregate ATNs gradually change from the situation that the number of nodes with high degree are more than the number of nodes with low degree to the heterogeneous network in the opposite case.

Figure 3(b) presents the average degree $\langle k \rangle$ as a function of number of layers of networks. Each curve describes the change of parameter values when the four airline groups' ATNs with same value of X are superimposed. It can be clearly seen that, under the same Y value, the relative magnitudes of $\langle k \rangle$ with different X values basically remains unchanged, which is similar for the clustering coefficient in Fig. 3(c) and the average shortest path length $\langle L \rangle$ in Fig. 3(d). Moreover, when $X = 2$, the value of average degree $\langle k \rangle$ is highest, which may be due to the strong radiation force of the hub airport with high classes. In Fig. 3(d), at different values of X , the shortest path length $\langle L \rangle$ of all aggregation networks decreases steadily with the increase of Y . In addition, the order of the function curve from top to bottom is the same as the descending order of X values.

In Fig. 3(e), it can be shown that the aggregation network still has a negative correlation between nodes since the value of r is always less than 0. Furthermore, the correlation between r and Y may be distinct for different X values, that is, as Y increases, the assortativity coefficient r of networks of display the converse tendency within different airport classes. At last, Fig. 3(f) shows that when $X = 4$ or 5, the

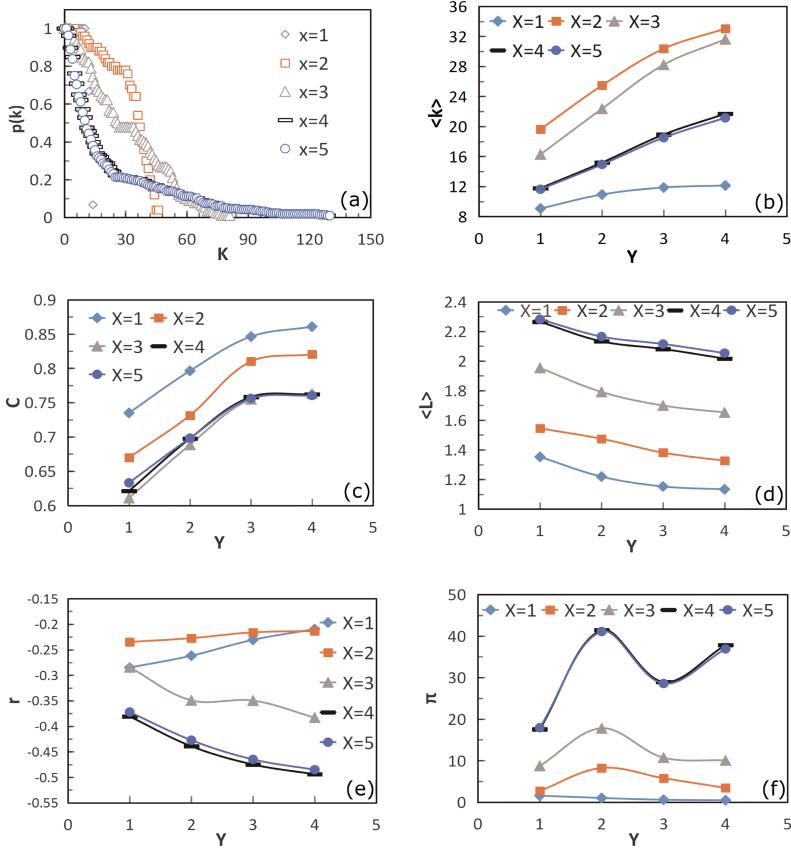


Fig. 3. The structural features of four major airline groups' aggregated ATN. X represents the number of airport node classes contained in the ATN at this time. Y represents the total number of superposition layers of the aggregation network.

fluctuation of the network polarization π is large, and the value of the network polarization is also significantly higher than that of other X -valued aggregation networks. As the value of X increases, joining of different classes of airport nodes makes the network gradually become a hub-and-spoke network composed of hub nodes, bridge nodes and periphery nodes, which renders the load of network become inhomogeneous.

3.3 Multi-layered Aggregate Chinese ATN with Node Classification

After exploring the structural properties of ATNs composed of China's four major airline groups, we will use the ATNs with node classification formed by more than ten other small airline routes as the fifth layer ($Y = 5$) superimposed on the above

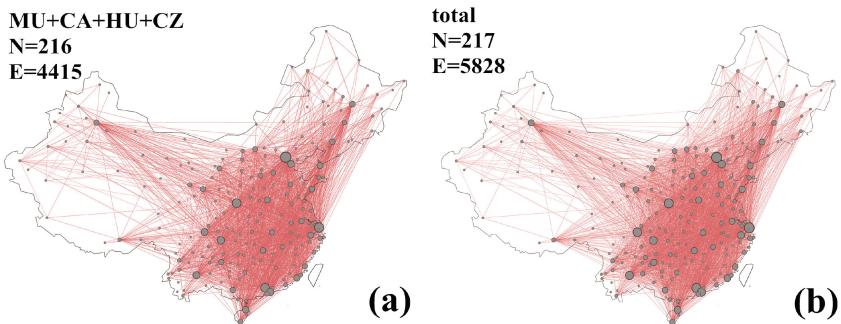


Fig. 4. (a) The structure of China's major airline groups' multi-layer aggregated ATN, which contains $N = 216$ nodes and $E = 4415$ edges. (b) The structure of China's aggregated ATN, which contains $N = 217$ nodes and $E = 5828$ edges.

aggregation ATNs. After the superposition, the ATN changed from Fig. 4(a) to (b), and the scale difference of the two types of aggregation networks is not obvious. In order to investigate the differences in route allocation between small airlines and major Chinese airline groups, we will mainly analyze the significant changes in the structural features of the aggregated ATN after the final layer is added into the networks.

From the cumulative distribution of China's major airline groups aggregation ATN and China's aggregate ATN with $X = 5$ and $Y = 5$ shown in Fig. 5(a), it can be observed that both networks exhibit the heterogeneous distribution, and the attenuation speeds of degree distribution of two networks are relatively close, which is also related to the little difference in the scale of the two aggregation ATNs.

It is also found that after superimposing ATN composed of all other small airlines, the aggregated ATNs under certain X values has a larger change than the network structure feature parameters of the previous major airline groups in China. In Fig. 5(b), in the case of $X = 4$ and $X = 5$, when $Y = 5$, the relative value of the average degree $\langle k \rangle$ increases significantly with respect to the superposition with the previous layer network; when $X = 3$, the superposition of the fifth layer also makes $\langle k \rangle$ increase significantly, even exceeding the average degree of the aggregated network at $X = 2$. This indicates that the endpoints of the edges of the small airline ATN are mostly 4D-class, 4C-class, 3C-class nodes. In Fig. 5(c), in the case of $X = 4$ and $X = 5$, with the addition of the fifth-tier small airline network, the clustering coefficient C is lower than that of the previous aggregation network, which implies that the small airline ATN has a low C , and the high clustering property of Chinese ATNs is mainly determined by ATNs of four major airline groups. In Fig. 5(d), with the addition of an ATN consisting of small airlines, the network polarization curves of different X values are reduced to a relatively low level at $Y = 5$, which indicates that the uneven load of China's ATNs is also mainly affected by ATNs of four major airline groups.

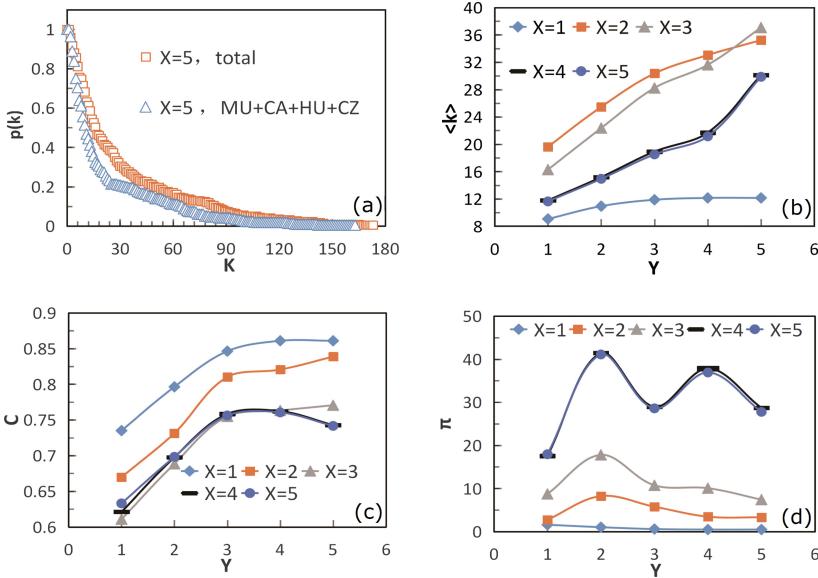


Fig. 5. Structural features of China's ATN formed by superimposing a small airline ATN. X represents the number of airport node classes contained in the ATN at this time. Y represents the number of superposition layers of the aggregation network.

4 Conclusion

In summary, we firstly explore the structural features of China's four major airline group ATNs. The results show that ATNs of each airline groups exhibit the feature of heterogeneous distribution, high clustering coefficient and short average path length. Moreover, in the process of gradually joining the nodes at various classes, most of the structural parameters of each network show the same trend as the network gradually becomes a hub-and-spoke network composed of hub nodes, bridge nodes and periphery nodes. Yet, the clustering coefficient and the network assortativity show the different trends due to the differences in the route allocation between the airline groups at different classes of airports.

When we consider the multi-layer aggregate ATN of four major airline groups in China, we find that with the superposition of multiple layers, the aggregate ATNs will gradually show higher clustering coefficient, shorter average path length, higher average degree. From the perspective of the change of node class number X in the aggregation network, with the increase of X , the aggregation ATNs gradually change from the situation that the number of high degree nodes is more than the number of low degree nodes to the opposite case. In addition, the variation ranges of five curves with the same feature parameters under different X values are obviously different, but the relative position of each curve is almost unchanged, which reveals that the classification of the airport is also reflected in the network properties of the node classification. Meanwhile, China's ATNs are finally formed by superimposing ATNs comprising all

airlines, it is evidenced that high clustering coefficient and load unevenness of China's ATNs are mainly determined by the ATN of the four major airline groups, but the high average degree of the network is greatly affected by the ATN of small airlines. The current results are helpful to understand the route distribution law of China's ATNs, manage the network more effectively, improve the network performance and enhance the air transport efficiency *etc.*

Acknowledgments. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant Nos. 61773286, and Tianjin Municipal Natural Science Foundation under grant no. 18JCYBJC87800.

References

1. Colizza V, Barrat A, Barthélemy M et al (2006) The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc Natl Acad Sci USA* 103(7):2015–2020
2. Guimera R, Amaral LAN (2004) Modeling the world-wide airport network. *Eur Phys J B* 38(2):381–385
3. Guimera R, Mossa S, Turtschi A et al (2005) The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. *Proc Natl Acad Sci USA* 102(22):7794–7799
4. Hsu C, Shih H (2008) Small-world network theory in the study of network connectivity and efficiency of complementary international airline alliances. *J Air Transp Manag* 14(3):123–129
5. Paleari S, Redondi R, Malighetti P (2010) A comparative study of airport connectivity in China, Europe and US: which network provides the best service to passengers? *Transp Res Part E Logist Transp Rev* 46(2):198–210
6. Barrat A, Barthélemy M, Vespignani A (2005) The effects of spatial constraints on the evolution of weighted complex networks. *J Stat Mech Theory Exp* 2005(05):P05003
7. Xu Z, Harriss R (2008) Exploring the structure of the US intercity passenger air transportation network: a weighted complex network approach. *GeoJournal* 73(2):87
8. Liu H, Tao Z (2007) Empirical study of Chinese city airline network. *Acta Phys Sinica* 56(1):106–112
9. Hong C, Zhang J, Cao X et al (2016) Structural properties of the Chinese air transportation multilayer network. *Chaos, Solitons Fractals* 86:28–34
10. Wei L, Cai X (2004) Statistical analysis of airport network of China. *Phys Rev E* 69(4):046106
11. Lin J (2012) Network analysis of China's aviation system, statistical and spatial structure. *J Transp Geogr* 22:109–117
12. Liu H, Zhang X, Lang C et al (2009) Analysis on the connecting mechanism of Chinese city airline network. *Sci China Ser G Phys Mech Astron* 39(7):935–942
13. Peng Y, Zhou Y (2009) Network structure of air transport in China. *Econ Geogr* 29:1850–1854
14. Wang J, Mo H (2014) Complex evolution process of China's air transport network. *J Transp Syst Eng Inf Technol* 14(1):71–80
15. Zeng X, Tang X, Jiang K (2011) Empirical study of Chinese airline network structure based on complex network theory. *J Transp Syst Eng Inf Technol* 11(6):175–181

16. Zhang Y, Round DK (2009) Policy implications of the effects of concentration and multimarket contact in China's airline market. *Rev Ind Organ* 34(4):307–326
17. Bagler G (2008) Analysis of the airport network of India as a complex weighted network. *Phys A* 387(12):2972–2980
18. Da Rocha LEC (2009) Structural evolution of the Brazilian airport network. *J Stat Mech Theory Exp* 2009(04):P04020
19. Cardillo A, Zanin M, Gómez-Gardenes J et al (2013) Modeling the multi-layer nature of the European air transport network: resilience and passengers re-scheduling under random failures. *Eur Phys J Spec Top* 215(1):23–33
20. Verma T, Araújo NAM, Herrmann HJ (2014) Revealing the structure of the world airline network. *Sci Rep* 4:5638
21. Wang P, Hunter T, Bayen AM et al (2012) Understanding road usage patterns in urban areas. *Sci Rep* 2(12):1001
22. Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512
23. Newman ME (2003) The structure and function of complex networks. *SIAM Rev* 45(2):167–256
24. Battiston F, Nicosia V, Latora V (2014) Structural measures for multiplex networks. *Phys Rev E* 89(3):032804
25. Costa LF, Rodrigues FA, Travieso G et al (2007) Characterization of complex networks: a survey of measurements. *Adv Phys* 56(1):167–242
26. Boccaletti S, Bianconi G, Herrero RC et al (2014) The structure and dynamics of multilayer networks. *Phys Rep* 544(1):1–122



Observer-Based Leader Following Consensus of Nonlinear Multi-agent Systems with a Leader of Bounded Input

Ziyi Wang and Lin Li^(✉)

Department of Control Science and Engineering,
University of Shanghai for Science and Technology, Shanghai 200093, China
lilin0211@163.com

Abstract. The leader-following consensus of multi-agent systems with Lipschitz nonlinear dynamics and a leader of bounded input is investigated in this paper. Firstly, for dealing with the difficult measurement of the real-time states, an state observer is designed for each agent. Considered the bounded input of leader, a consensus control protocol is proposed for the multi-agent systems. Then, the validity of the control protocol is proved by Lyapunov stability theory and graph theory. Finally, a numerical example is given to demonstrate the effectiveness of the control protocol.

Keywords: Leader-following consensus · Observer · Multi-agent systems · Bounded input

1 Introduction

Recently, with the rapid development of computer science and embedded technology, multi-agent systems have been gradually applied to various scientific fields e.g. flocking control [11], agreement or consensus laws [5] and formation control [7]. In these areas, consensus problem of multiagent systems is one of the core issues in cooperative control of multiagent systems. The model of consensus was firstly raised by Vicsek, then it was further developed under the researches of Olfati-Saber [1, 3] and Murray [2]. Subsequently, domestic scholars Ren [4, 5] and Li [6, 8, 10, 12] have done a lot of works on the basis of existing researches. Ren initially enumerated the basic knowledge needed to study multi-agent consensus, and then applied consensus theories on aircraft and vehicles to solve practical control problems. Li focused on the connectivity and directionality of the topological graph and the influence of different control strategies on system consensus.

Currently, the research object of multi-agent consensus has changed from first-order systems to second-order [13, 14] or high-order systems. The topological structure of the relationships among agents has also expanded from strong connectivity and undirection [17, 18] to low connectivity and direction, and even to more stringent directed spanning trees or directed forests [9]. In addition, nonlinear systems have

become the main research object instead of linear systems [19, 20]. In recent years, while studying the consensus of non-linear multi-agent systems, the leader-following consensus has also been widely discussed [15, 16]. Choosing a leader for a multi-agent system can help other agents follow the leaders real-time states, which may improve to reach the consensus efficiency and accuracy.

There are lots of literatures studying the consensus control of nonlinear multi-agent systems with leader-following scheme [20, 22, 23]. Inspired by above analysis, this paper devotes to design a consensus protocol for achieving the consensus tracking of nonlinear multi-agent systems with a leader of bounded input. Our contributions can be summarized as follows: (1) We design an state observer for each agent considering the impossibility to track all the states of agents, The states of the observer can be used as the real-time states of the agent during the analysis. (2) The leader included should be contained. In the process of reaching the consensus, the leader can solve the issue of resources and his states are independent of followers. Additionally, as time goes by all followers can follow the states of leader. (3) In many cases, the bounded input should not be ignored especially when it is applied to the leader to achieve certain objective, e.g. reach a desirable consensus value or to avoid hazardous obstacles.

Notations: I_N denotes the identity matrix of N order, 1_N represents a column vector with all entries equal to one, $R^{n \times n}$ be the set of $n \times n$ real matrices. If the real matrix $P > 0$, which means P is positive definite; The superscript T of a vector or a matrix represents the transpose it. The symbol \otimes means Kronecker product; $\|x\|$ and $\|A\|$ denote its Euclidean norm.

2 Problem Formulation and Preliminaries

The topology of this paper is modeled by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, where $\mathcal{V} = \{1, 2, \dots, N\}$ is the set of nodes, $\mathcal{E} = \mathcal{V} \times \mathcal{V}$ is the set of edges, and a weighted adjacency matrix $\mathcal{A} = (a_{ij}) \in R^{N \times N}$, in which, the entry $a_{ii} = 0, a_{ij} = 1$ if the edge $(v_i, v_j) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. The node v_i is called a neighbor of node v_j , then the neighbor index set of agent v_i can be denoted by $N_i = \{j \mid (i, j) \in \mathcal{E}\}$. The degree matrix of \mathcal{G} is defined as $\Delta = \{\Delta_1, \Delta_2, \dots, \Delta_N\}$, where $\Delta_i = \sum_{j \in N_i} a_{ij}$, The Laplacian matrix of the graph is defined as $L = \Delta - \mathcal{A}$.

In this undirected topology, let $\mathcal{D} = \{d_1, d_2, \dots, d_N\}$ be the leader adjacency matrix, and the union graph $\mathcal{H} = \mathcal{L} + \mathcal{D}$.

We consider a group of $N+1$ agents with nonlinear dynamics, consisting of N followers and one leader. The dynamics of each follower is given as

$$\begin{cases} \dot{x}_i = Ax_i + Bu_i + f(x_i), \\ y_i = Cx_i \end{cases}, \quad i = 1, 2, \dots, N \quad (1)$$

where $x_i \in R^n$ is the state, $u_i \in R^m$ is the control input, and $f(x_i) \in R^n$ is the intrinsic nonlinear dynamics.

The leader is described as

$$\begin{cases} \dot{x}_0 = Ax_0 + Bu_0 + f(x_0) \\ y = Cx_0 \end{cases} \quad (2)$$

where $x_0 \in R^n$ is the state of leader, which is independent of followers' states. $f(x_0) \in R^n$ is the intrinsic nonlinear dynamics. And A, B, C are known constant matrices with appropriate dimensions.

Considering the difficulty of reaching the states of agents, we consider the following observers

$$\begin{cases} \dot{\tilde{x}}_i = A\tilde{x}_i + Bu_i + G(\tilde{y}_i - y_i) + f(\tilde{x}_i), \\ \tilde{y}_i = C\tilde{x}_i \end{cases}, \quad i = 1, 2, \dots, N \quad (3)$$

where $\tilde{x}_i \in R^n$ is the observer state, and $f(\tilde{x}_i) \in R^n$ is the intrinsic nonlinear dynamics, whose internal structure is known, G is the observer gain matrix.

The leader-following consensus of system (1) and (2) is said to be achieved, if with the proposed control input $u_i(t)$, the close-loop system satisfies $\lim_{t \rightarrow \infty} x_i(t) = \lim_{t \rightarrow \infty} x_0(t)$ for any initial condition $x_i(0)$.

Before giving the main result, we should list some useful lemmas and assumptions.

Assumption 1: The interaction graph among followers is undirected and the leader has a directed path to each other.

Assumption 2: The leaders control input u_0 is continuous and bounded, i.e., that means $\|u_0\|_\infty = \gamma$, where γ is a positive constant.

Assumption 3: It is assumed that the pair (A, B) and (C, A) are controllable and detectable, respectively.

Lemma 1 [5]: Suppose the graph \mathcal{G} is connected, then we can find the matrix \mathcal{H} has a zero eigenvalue and the corresponding eigenvector is 1_N .

Lemma 2 [7]: The eigenvalues of the matrix \mathcal{H} are nonnegative. The matrix \mathcal{H} is positive definite if and only if the graph \mathcal{G} is connected.

Lemma 3 [21]: $f(x) \in R^n$ is nonlinear dynamics which satisfies the Lipschitz condition with a Lipschitz constant $\beta > 0$, i.e., for any $x, y \in R^n$, we have

$$\|f(x) - f(y)\| \leq \beta \|x - y\|$$

3 Main Result

In this section, with the observer (3), systems (1) and (2) can reach the leader following consensus, design the consensus protocol as follow:

$$u_i = K \left[\sum_{j \in N_i} a_{ij} (\tilde{x}_j - \tilde{x}_i) + d_i (x_0 - \tilde{x}_i) \right] - c \operatorname{sgn} K \left[\sum_{j \in N_i} a_{ij} (\tilde{x}_j - \tilde{x}_i) + d_i (x_0 - \tilde{x}_i) \right] \quad (4)$$

where K is the control gain matrix to be determined and $K = B^T Q$. c is the parameter related to the bounded input u_0 .

Theorem 1: Given positive scalars β , if there exist a positive definite matrix P such that the following Riccati inequality

$$P(A + GC) + (A + GC)^T P + \beta^2 P^2 + I_n < 0 \quad (5)$$

holds. Then N follower agents can track the leader agent from any initial conditions.

Proof: Let $\xi_{1i} = \tilde{x}_i - x_i$, $\xi_{2i} = \tilde{x}_i - x_0$. Then, we can get

$$\begin{cases} \dot{\xi}_{1i} = (A + GC)\xi_{1i} + (f(\tilde{x}_i) - f(x_i)) \\ \dot{\xi}_{2i} = A\xi_{2i} + BK \sum_{j \in N_i} h + \lambda - Bu_0 + (f(\tilde{x}_i) - f(x_0)) \end{cases} \quad (6)$$

In which,

$$\begin{aligned} h &= a_{ij} (\xi_{2j} - \xi_{2i}) - BKd_i \xi_{2i} + GC\xi_{1i}, \\ \lambda &= cB \operatorname{sgn} K \left(\sum_{j \in N_i} a_{ij} (\xi_{2j} - \xi_{2i}) - d_i \xi_{2i} \right) \end{aligned}$$

Set $\xi_1^T = [\dot{\xi}_{11}^T, \dot{\xi}_{12}^T, \dots, \dot{\xi}_{1N}^T]$ and $\xi_2^T = [\dot{\xi}_{21}^T, \dot{\xi}_{22}^T, \dots, \dot{\xi}_{2N}^T]$, then we get

$$\begin{cases} \dot{\xi}_1 = (I_N \otimes (A + GC))\xi_1 + (I_N \otimes I_n)(f(\tilde{x}_i) - f(x_i)) \quad (7a) \\ \dot{\xi}_2 = (I_N \otimes A)\xi_2 - (H \otimes BK)\xi_2 + (I_N \otimes GC)\xi_1 + \varphi \quad (7b) \\ \quad -(I_N \otimes B)u_0 + (I_N \otimes I_n)(f(\tilde{x}_i) - f(x_0)) \end{cases}$$

Where $\varphi = c(I_N \otimes B)\operatorname{sgn}((H \otimes K)\xi_2)$.

In order to simplify the derivation, the error system in (7a) is analyzed first. Consider the following Lyapunov function candidate $V_1(t)$:

$$V_1(t) = \xi_1^T (H \otimes P) \xi_1$$

Where P is positive definite matrix. The following inequality is used in,

$$\begin{aligned}
& 2\xi_1 P(f(\tilde{x}_i) - f(x_i)) \\
& \leq 2\|P\xi_1\| \|f(\tilde{x}_i) - f(x_i)\| \\
& \leq 2\beta\|P\xi_1\| \|\xi_1\| \\
& \leq \beta^2 \xi_1^T P^2 \xi_1 + \xi_1^T \xi_1
\end{aligned}$$

The time derivative of V_1 along can be obtained as

$$\begin{aligned}
V_1(t) &= 2\xi_1^T (H \otimes P) \left[\begin{array}{l} I_N \otimes (A + GC) \xi_1 \\ + (I_N \otimes I_n)(f(\tilde{x}_i) - f(x_i)) \end{array} \right] \\
&\leq \xi_1^T [I_N \otimes P(A + GC) + (A + GC)^T P] \xi_1 \\
&+ \xi_1^T [I_N (\beta^2 P^2 + I_n)] \xi_1 \\
&= \xi_1^T \left[\begin{array}{l} I_N \otimes P(A + GC) \\ + (A + GC)^T P + \beta^2 P^2 + I_n \end{array} \right] \xi_1
\end{aligned}$$

Denote

$$\Phi_1 = P(A + GC) + (A + GC)^T P + \beta^2 P^2 + I_n$$

According to Theorem 1, $\Phi_1 < 0$, it is easy to know that for the estimation error ξ_1 , $\lim_{t \rightarrow \infty} x_i(t) = \lim_{t \rightarrow \infty} \tilde{x}_i(t)$, which means the error system of (7a) is stable. So, in the following derivation, the effect of (7a) system errors can be neglected. The part contains ξ_1 may not be discussed.

Remark 1: An observer is designed for each agent, the state of the observer can be approximately regarded as the real-time states of the agent. In the existing literatures, the consensus error is always designed as $e_i = x_i - x_0$. In this article, with the definition $\lim_{t \rightarrow \infty} (x_i(t) - x_0(t)) = \lim_{t \rightarrow \infty} (\tilde{x}_i(t) - x_0(t))$, therefore, the consensus problem of systems (1) and (2) can be transformed to (7b).

Follow the reasons above, the derivation of $\dot{\xi}_2$ can be written as

$$\begin{aligned}
\dot{\xi}_2 &= (I_N \otimes A) \xi_2 \\
&- (H \otimes BK) \xi_2 \\
&+ c(I_N \otimes B) \text{sgn}((H \otimes K) \xi_2) \\
&- (I_N \otimes B) u_0 \\
&+ (I_N \otimes I_n)(f(\tilde{x}_i) - f(x_0))
\end{aligned} \tag{8}$$

Theorem 2: Given positive scalars β, μ, η , if there exist a positive definite matrix Q such that the following Riccati inequality

$$QA + A^T Q - 2\mu QBB^T + \beta^2 Q^2 + I_n < -\eta I_n \tag{9}$$

holds. Then N follower agents can track the leader agent from any initial conditions.

Choose Lyapunov function $V_2(t)$ as

$$V_2(t) = \xi_2^T (H \otimes Q) \xi_2$$

where Q is a positive definite matrix. The time derivative of $V_2(t)$ along the trajectory can be given as

$$\begin{aligned} V_2(t) &= 2\xi_2^T (H \otimes Q) \Omega \\ &\leq \xi_2^T [I_N \otimes (QA + A^T Q + \beta^2 Q^2 + I_n)] \xi_2 \\ &\quad - 2\xi_2^T (H \otimes QBK) \xi_2 \\ &\quad + 2c(I_N \otimes B) \operatorname{sgn}(H \otimes K) \xi_2 \\ &\quad - 2\xi_2^T (H \otimes QB) u_0 \\ &= T_1 + T_2 \end{aligned} \tag{10}$$

Where

$$\left\{ \begin{array}{l} T_1 = \xi_2^T [I_N \otimes (QA + A^T Q + \beta^2 Q^2 + I_n)] \xi_2 - 2\xi_2^T (H \otimes QBK) \xi_2 \\ T_2 = 2c(I_N \otimes B) \operatorname{sgn}(H \otimes K) \xi_2 - 2\xi_2^T (H \otimes QB) u_0 \end{array} \right. \tag{11}$$

$$\left\{ \begin{array}{l} T_2 = 2c(I_N \otimes B) \operatorname{sgn}(H \otimes K) \xi_2 - 2\xi_2^T (H \otimes QB) u_0 \end{array} \right. \tag{12}$$

$$\begin{aligned} \Omega &= (I_N \otimes A) \xi_2 - (H \otimes BK) \xi_2 + c(I_N \otimes B) \operatorname{sgn}(H \otimes K) \xi_2 \\ &\quad - (I_N \otimes B) u_0 + (I_N \otimes I_n)(f(\tilde{x}_i) - f(x_0)) \end{aligned}$$

For T_1 , define $\Phi_2 = QA + A^T Q - 2\mu Q BK + \beta^2 Q^2 + I_n$, choose a positive scalar η , we can find that

$$\begin{aligned} T_1 &\leq \sum_{i=1}^N \xi_2^T (QA + A^T Q - 2\mu Q BK + \beta^2 Q^2 + I_n) \xi_2 \\ &< \xi_2^T (-\eta I_n) \xi_2 \\ &< 0 \end{aligned} \tag{13}$$

Constructing the matrix K in consensus protocol as

$$K = B^T Q \tag{14}$$

We also have the fact that $x^T \operatorname{sgn}(x) = \|x\|_1$, by using these elements, it follows from that

$$\begin{aligned}
T_2 &= 2c(I_N \otimes B)\text{sgn}(H \otimes K)\xi_2 - 2\xi_2^T(H \otimes QB)u_0 \\
&= 2c(I_N \otimes B)\text{sgn}(H \otimes B^T Q)\xi_2 - 2\xi_2^T(H \otimes QB)u_0 \\
&= 2(c - \gamma)\|H \otimes BQ\|\xi_2^T
\end{aligned} \tag{15}$$

Hence \dot{V}_2 is

$$\dot{V}_2 \leq T_1 + T_2 \tag{16}$$

As long as u_0 is bounded, combine with the parameter c , which yield, $\dot{V}_2 < 0$. From Lyapunov theorem $V_2 > 0$ and $\dot{V}_2 < 0$. Therefore, $\lim_{t \rightarrow \infty} x_0(t) = \lim_{t \rightarrow \infty} \tilde{x}_i(t)$, it follows that all followers can track the leader agent for any initial condition, this completes the proof of the convergence.

Remark 2: Due to the nonzero control input u_0 , It is necessary to design an independent adjustable parameter c . Compared with [20, 21], the parameter c in this paper can be updated with input u_0 . When $u_0 = 0$, setting the $c = 0$, the system becomes a general nonlinear system, which is not subject to input constraints. It ensures the consistency of the systems and makes the control protocol more flexible and effective.

4 Simulation Example

Assume that the network of this system is made up of four follower and one leader. The communication topological graph $\bar{\mathcal{H}}$ given in Fig. 1.

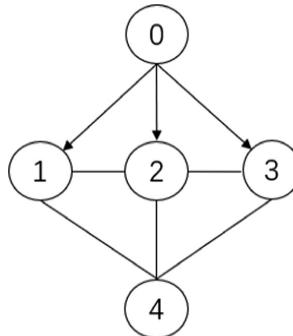


Fig. 1. Communication topology

From the topology above, its easily to get the Laplacian matrix and the adjacency matrix

$$L = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}, \quad D = \text{diag}(1, 0, 1, 0)$$

The dynamics of the leader and followers satisfied system (1) and (2) are given

$$A = \begin{bmatrix} 0 & -1 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C = [1 \quad 0]$$

By solve the Riccati inequality (9), we have

$$Q = \begin{bmatrix} 2.4341 & 0.5614 \\ 0.5614 & 1.3772 \end{bmatrix}$$

Given the observer feedback matrix G and control gain matrix K as follows

$$G = \begin{bmatrix} -5.1578 \\ 4.4173 \end{bmatrix}, K = [0.5614 \quad 1.3772]$$

Considering the random initial state, the parameter is chosen to be $c = 0.5$ and the bounded input of leader selected

$$u_0 = [0.5\sin t \quad 0]^T$$

The state trajectories of each agents shown in Fig. 2 and the states error between observers and leader shown in Fig. 3, respectively.

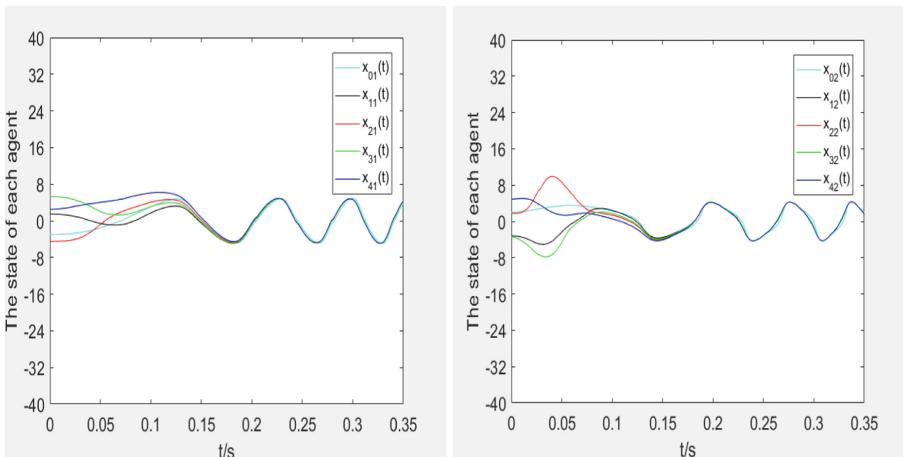


Fig. 2. Trajectories of agents' states

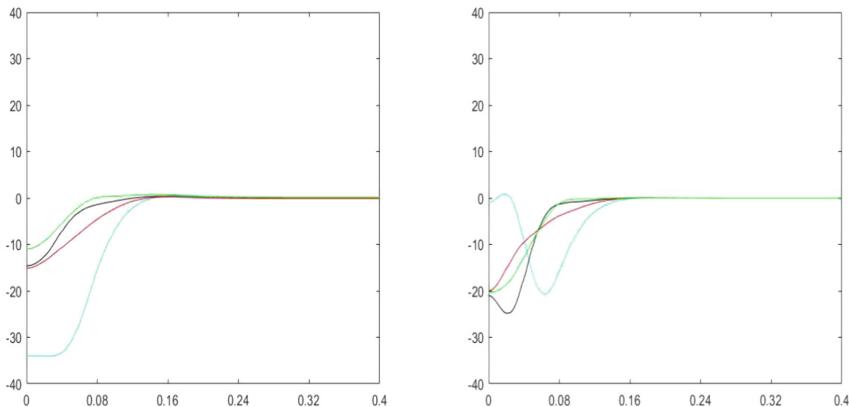


Fig. 3. The measurement errors between observers and leader

The above results show that under the control protocol (4) with state observer, the multi-agent system with five agents can achieve the leader-following tracking control consensus.

5 Conclusion

Leader-follower consensus of nonlinear multi-agent systems is studied in this paper. Firstly, considering the immeasurability of the agents states tracking, a state observer is designed for each agent, which used to track the real-time states of agents. Secondly, considering the system contained a leader with bounded input, the corresponding consensus control protocol is designed. The theoretical deduction proves that leader following consensus of the system can achieved. Finally, a simulation example verifies the effectiveness of the proposed method.

References

1. Olfati-Saber R, Murray R (2004) Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans Autom Control* 49(9):1520–1533
2. Fax JA, Murray RM (2004) Information flow and cooperative control of vehicle formations. *IEEE Trans Autom Control* 49(9):1465–1476
3. Olfati-Saber R, Murray R (2003) Consensus protocols for networks of dynamic agents. In: American control conference, June, vol 2, pp 951–956
4. Ren W, Beard RW (2007) Information consensus in multivehicle cooperative control. *IEEE Control Syst Mag* 1066-033x
5. Ren W, Atkins EM (2007) Distributed multi-vehicle coordinated control via local information exchange. *Int J Robust Nonlinear Control* 17(1011):1002–1033
6. Li ZK, Liu XD (2012) Distributed tracking control for linear multi-agent systems with a leader of bounded input using output information. In: Proceedings of the 10th world congress on intelligent control and automation, Beijing, China, pp 1756–1761

7. Yang YB, Wang CL (2013) Obstacle avoidance method for mobile robots based on improved artificial potential field method and its implementation on MATLAB. *J Univ Shanghai Sci Technol* 35(5):496–500
8. Li ZK, Ren W, Liu XD (2013) Distributed consensus of linear multi-agent systems with adaptive dynamics protocol. *Automatica* 49:1986–1995
9. Cheng TH, Kan Z, Shea JM, Dixon WE (2009) Decentralized event-triggered control for leader-follower consensus. In: The 53rd IEEE conference on decision and control, Los Angeles, USA, pp 1244–1249
10. Li ZK, Wen GH, Duan ZS (2014) Designing fully distributed consensus protocols for linear multi-agent systems with directed graphs. *IEEE Trans Autom Control* 60:1152–1157
11. Tanner H, Jadbabaie A, Pappas G (2007) Flocking in fixed and switching networks. *IEEE Trans Autom Control* 52(5):863–868
12. Li ZK, Ren W, Liu XD (2013) Consensus of multi-agent systems with general linear and Lipschitz nonlinear dynamics using distributed adaptive protocols. *IEEE Trans Autom Control* 58(7):1786–1791
13. Meng XY, Chen TW (2013) Event based agreement protocols for multi-agent networks. *Automatica* 49(7):2125–2132
14. Zhang H, Feng G, Yan HC (2014) Observer-based output feedback event-triggered control for consensus of multi-agent systems. *IEEE Trans Ind Electron* 61(9):4885–4894
15. Zhou B, Liao XF, Huang TW (2015) Leader-following exponential consensus of general linear multi-agent systems via event-triggered control with combinational measurements. *Appl Math Lett* 40(2):35–41
16. Deng C, Yang GH (2019) Leaderless and leader-following consensus of linear multi-agent systems with distributed event-triggered estimators. *J Franklin Inst* 356(1):309–311
17. Dimarogonas DV, Frazzoli E, Johansson KH (2012) Distributed event-triggered control for multi-agent systems. *IEEE Trans Autom Control* 57(5):1291–1297
18. Miao W, Zhang H, Yan H, Ren H (2016) Self-triggered output feedback control for consensus of multi-agent systems. *Neurocomputing* 190:179–187
19. Zhang H, Feng G, Yan H, Chen Q (2014) Distributed self-triggered control for consensus of multi-agent systems. *IEEE/CAA J Autom Sinica* 1(1):40–45
20. Shi Y, Yin YY (2017) Distributed leader-following consensus of a class of nonlinear multi-agent systems. In: 11th Asian control conference. Gold Coast Convention Centre, Australia
21. Zhang XX, Liu XP (2017) Consensus of linear multi-agent systems with disturbance generated from heterogeneous nonlinear exosystems. In: Proceedings of the 36th Chinese control conference, Dalian, China, pp 8063–8067
22. Zou WC, Guo J, Xiang ZR (2018) Leader-following consensus protocol design of second-order nonlinear multi-agent systems with unmodeled dynamics. In: Proceedings of the 37th Chinese control conference, Wuhan, China, pp 6884–6889
23. Lu J, Li L (2018) Leader-following consensus of nonlinear multi-agent systems using observer-based event-triggered control. In: 5th international conference on information, cybernetics and computational social systems, Hangzhou, China, pp 247–252



Target Recognition Method of Street Lamp Based on Point Cloud Data

Pan Cheng, Xiaobin Li^(✉), and Tianyang Yu

School of Shanghai Institute of Technology, Shanghai 201418, China
lixiaobinauto@163.com

Abstract. There are several problems need to be solved in the process of the recognition of the point cloud target in the urban traffic scene: Low target recognition accuracy due to noise and low real-time performance when dealing with a large amount of data. Therefore, in order to improve the situation, a new method, SVM target recognition algorithm based on fast density clustering is proposed in this paper. The workflow of the algorithm includes: (1) Point cloud denoising. In this step, most noise points are removed, the adverse effect of noise on subsequent processing is avoided. (2) A fast density clustering algorithm, GLDBSCAN, is used for clustering and then the eigenvectors are extracted. (3) A SVM classifier is used to identify the point cloud target. Experimental results show that the algorithm is effective. The algorithm presented in this paper is more accurate and efficient than the previous point cloud recognition algorithms.

Keywords: Point cloud · Fast density clustering · SVM · Target recognition

1 Introduction

Target recognition has a wide range of applications in many aspects, such as face recognition, intelligent transportation, environmental perception. At present, target recognition can be carried out through image or point cloud data. Since the accurate distance information cannot be provided, the image-based target recognition method is difficult to provide reliable basis for the anti-collision warning and path planning of vehicles. Point cloud data contains a large amount of information and is less disturbed by external factors. Therefore, point cloud data can be used for target recognition, and the results will be more in line with people's expectations. Due to the high frequency and high precision of 3d point cloud data, this method has become a research hotspot.

For the recognition of point cloud targets, Lu et al. [1] extracted features based on the sparse theory. This method reduces the amount of data, thus reduces the amount of calculation and improves the efficiency. This is a good method in addition to its poor anti-noise performance. Kong et al. [2] improved the k-means algorithm to extract the feature points inside clustering targets for target recognition.

However, the application scenarios of this method are very limited, and the effects vary greatly under different conditions. Li et al. [3] presented the idea of combining particle swarm optimization with genetic algorithm to optimize SVM parameters, the method can extract feature vectors more efficiently. The drawback of this method is that

when a large amount of data cannot be provided for training, the effect is not satisfactory. Although Zhang et al. [4] proposed the method of sample matching to realize the fast identification of street lamps. It has the obvious deficiency of low recognition accuracy. The author in [5] developed the idea of separating objects above the ground from the ground by the slope between adjacent points. Unsupervised clustering was used to cluster the point clouds on the ground, and the least square method was used to fit the circle to identify the rod-shaped structure. This method can effectively identify the target, but it takes a long time.

In order to get the accuracy and real-time of the target recognition algorithm at the same time, this paper improves the method proposed by Shaker et al. [5]. In order to improve the time performance of clustering, we use GLDBSCAN algorithm to cluster point cloud data and then the radial kernel SVM classifier is used to complete the final recognition task.

2 Clustering and Classification of Street Lamp Targets

Target recognition based on point cloud data includes several steps, such as point cloud filtering, point cloud clustering and feature extraction, the following is the structural framework of the algorithm (Fig. 1).

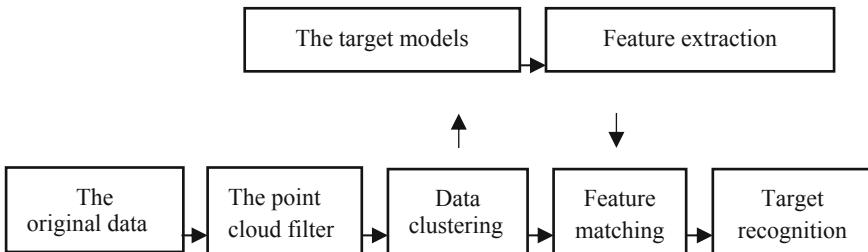


Fig. 1. Algorithm structure framework

2.1 Clustering of Point Cloud

Density-based clustering method divides different regions according to different data densities. Noise points in data have little impact on clustering results, so DBSCAN clustering can be used to cluster point clouds. DBSCAN density clustering starts from a point of high density, connecting the points with similar density one by one from this point, and finally forming clusters. For any data point, denoted as P , the Eps neighborhood of P is represented by $N_{\text{Eps}}(P)$, defined as [6]

$$N_{\text{Eps}}(P) = \{q \in D | \text{dist}(p, q) \leq \text{Eps}\} \quad (1)$$

The number of data points in the Eps neighborhood of P is calculated, and the number is the density of data points. And then select a density threshold, compare the number of points in the circle with the size of the threshold, if the number of points is less than the density threshold, then the central point is judged to be a low density point, on the contrary, the central point is judged to be a high density point, high density point is also called the core point. If the high-density point A is in the range of another high-density point B, connect the two points and connect the different high-density points according to this rule. The low-density points in the range of the high-density circle are defined as the boundary, and the remaining low-density points are identified as noise points. The four pictures below are the density clustering results when different Esp parameters are selected. The results show that the clustering results are closely related to the parameters (Fig. 2).

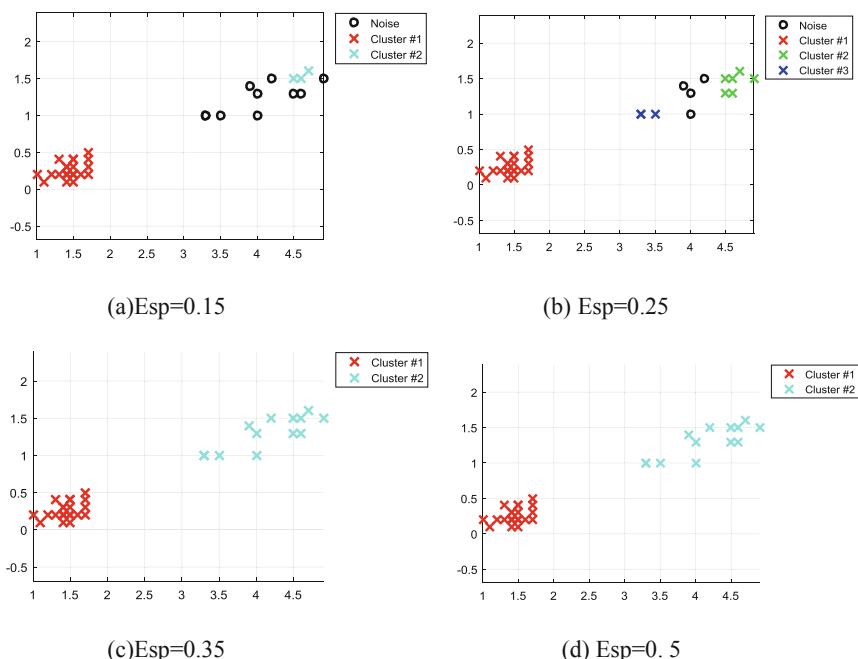


Fig. 2. Effect of parameters on DBSCAN algorithm

In addition to the above disadvantages, since the implementation of DBSCAN algorithm requires scanning the Esp neighborhood of each data point, therefore, it takes a long time to complete the clustering, so the algorithm is inefficient. Moreover, when the density is not uniform, the quality of clustering cannot be guaranteed.

In order to achieve better clustering effect, we adopted a fast density clustering algorithm called GLDBSCAN, the algorithm has high efficiency and good clustering

effect. Firstly, a grid structure is constructed based on spatial partition to divide the data space, and then the value of the central object of each non-empty grid is determined. Finally, the grid is clustered and classified [6].

The specific process is as follows:

Step 1. Grid generation: the interval length of the i th dimension of the data space is $I_i = [\text{high}_i - \text{low}_i]$, and Num is the number of objects in the data space.

$$l = \left[\sqrt[d]{\frac{I_1 \times I_2 \times \dots \times I_k}{\text{Num}/k}} \right] \quad (2)$$

Step 2. Build the SP – Tree: According to the formula given in Step 1, the grid G to which the object p belongs is determined, and then the node of the corresponding layer searched from the first according to the interval ordinal number. If there is a corresponding cell number on the sp-tree, start the search from the next dimension, modify the leaf node, and store the information contained in the object. Otherwise, a new node is created for the existence of the cell.

Step 3. Identify the grid center object: The non-empty grid center object is used instead of the grid to complete the clustering, the method will accelerate the speed of clustering. We define the absolute error to determine the central object:

$$E = \sum_{p \in G} |p - o| \quad (3)$$

Step 4. Data clustering: Select a core network and create a new class C that takes o as the core object and contains all the objects in G. When two adjacent grids are both core grids, the locally accessible density ratio between them can be calculated by the following formula:

$$\text{ratio} = \text{LRD}_k(o_1) - \text{LRD}_k(o_2) \quad (4)$$

Given a δ , if $|\text{ratio} - 1| < \delta$ is satisfied, we consider that the adjacent grid and the grid belong to the same class. On this basis, we judge other grids until a grid of other classes is selected. If the above conditions are not met, the two grids are marked as different classes and each of them will be extended. If it is a non-core network, then find out whether the adjacent grid contains the core network. If it does, set the corresponding identity to 1 and divide it into boundary grid. The minimum distance between the core grid and the boundary grid is calculated to determine whether an object is noise or not. Follow these steps until all the grids are processed.

The GLDBSCAN algorithm is not sensitive to parameters. The following pictures are the results of GLDBSCAN algorithm under different parameters (Fig. 3):

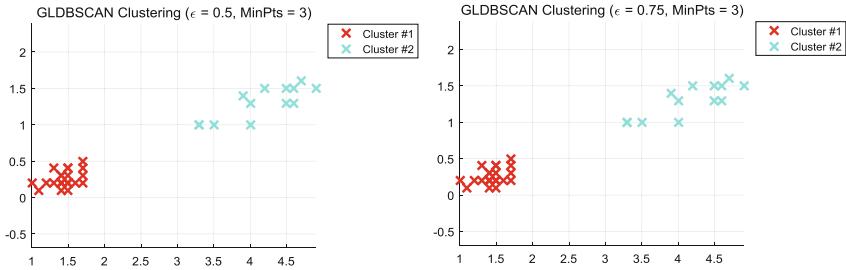


Fig. 3. Effect of parameters on GLDBSCAN algorithm

2.2 The SVM Classifier

SVM is a fast and reliable algorithm for target recognition. Its purpose is to establish a hyperplane, separating the two classes on both sides of the hyperplane, and the distance between them and the nearest vector reaches the maximum. In short, the hyperplane is a nonlinear decision boundary and the optimal boundary [7].

When we construct a hyperplane, we only need to consider the inner product between the eigenvector and the support vector, and we do not need to consider it in the way of display. The general expression of inner product is:

$$z_i \cdot z = K(x, x_i) \quad (5)$$

In the formula, z represents the mapping of vectors in the input space to the eigenspace.

When SVM is used as a binary classifier, the output of SVM is:

$$f(x, W) = \text{sgn} \left(\sum_{i=1}^N w_i K(x, x_i) + b \right) \quad (6)$$

The optimization objective function is

$$J = W^T W = \|W\|^2 \quad (7)$$

The constraint condition is:

$$y_j = \left[\sum_{i=1}^N w_i K(x, x_i) + b \right] \geq 1, \quad j = 1, \dots, N \quad (8)$$

In the formula, N represents the total number of samples, W is a parameter of the support vector machine, and (x_i, y_i) is the sample. In order to get the best classification effect, an objective function J and its constraints are defined, and a relaxation variable is introduced. This processing method ensures the accuracy of the results and increases the anti-noise performance of the algorithm.

$$J = \frac{1}{2} W^T W + C \sum_{i=1}^N \xi_i, \quad \xi_i \geq 0 \quad (9)$$

$$y_j = \left[\sum_{i=1}^N w_i K(x, x_i) + b \right] \geq 1 - \xi_j, \quad j = 1, \dots, N \quad (10)$$

The following is the result of using the SVM classifier (Fig. 4):

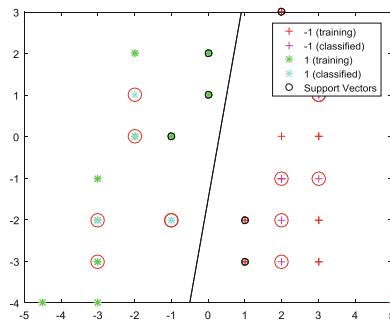


Fig. 4. Results of SVM classifier

3 Results

3.1 Data Visualization

As shown in Fig. 5, data visualization is conducive to the analysis of data and the presentation of information carried in data.

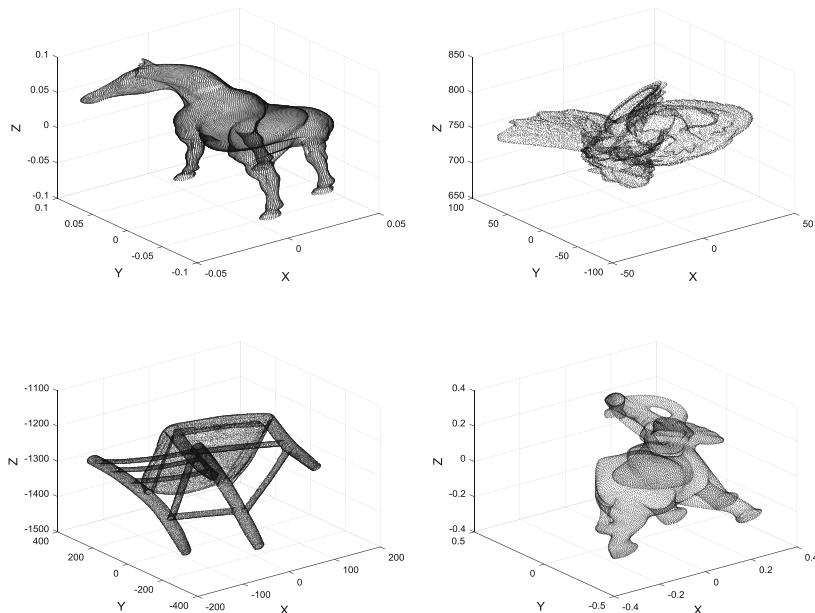


Fig. 5. Visualization of point cloud data

The visualization results of experimental scene are shown in Fig. 6:

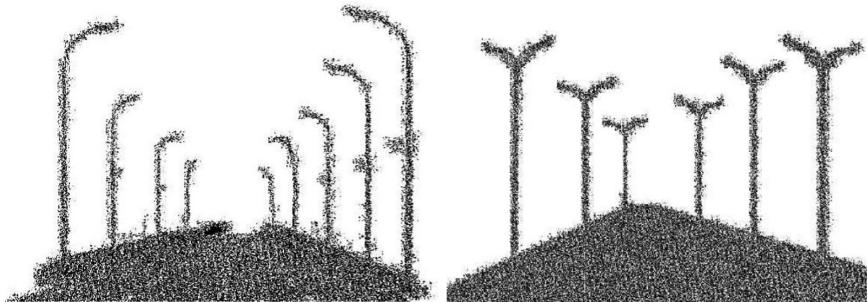


Fig. 6. Visualization of experimental scene

The visualization results of street lamp point cloud are shown in Fig. 7:

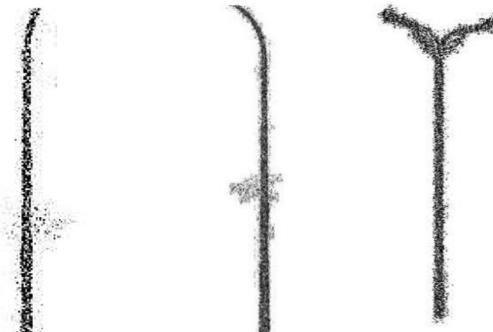


Fig. 7. Visualization of single street lamp point clouds

3.2 Identification of Street Lamp

We carried out experiments in MATLAB to realize target recognition of street lamps on the road, and the results are shown in Fig. 8:

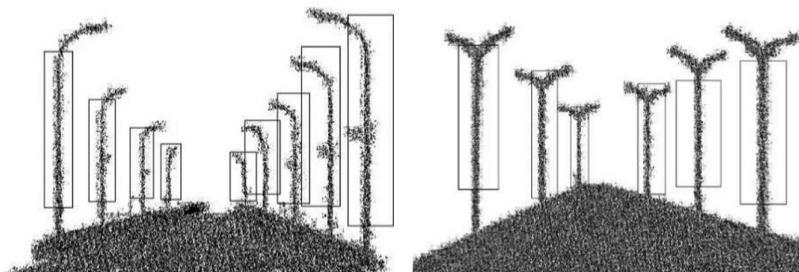


Fig. 8. Results of street lamp target recognition

4 Conclusions

- (1) This paper proposed a SVM algorithm based on fast density clustering to solve the problems of poor anti-noise performance, low recognition accuracy and low algorithm efficiency in point cloud object recognition.
- (2) GLDBSCAN algorithm combined with SVM classifier was used for clustering and recognition, which improved the accuracy and efficiency of the algorithm. Experimental results verify the feasibility of street lamp target recognition by SVM classifier based on GLDBSCAN algorithm.
- (3) Experiments results show that the proposed method can identify specific targets effectively. Compared with previous methods, the algorithm is more accurate and efficient.

Acknowledgments. Firstly, I'd like to express my sincere gratitude to my supervisor, professor Li Xiaobin. I really appreciate the guidance and advice given by him during my thesis writing.

Secondly, I would like to thank my seniors and classmates for their selfless help and encouragement during this period.

Finally, I would like to thank all those who helped me, directly or indirectly, during my thesis writing.

References

1. Lu J, Wang Z, Sun W et al (2018) LiDAR point cloud target recognition based on sparse representation. *Comput Era* (11):1–4
2. Kong D, Wang X, Liu Y et al (2008) Vehicle target recognition algorithm based on vehicle-mounted 32-line lidar point cloud. *Sci Technol Eng* 18(05):81–85
3. Li H, Wang H, Li Y et al (2016) Vehicle-borne radar point cloud target recognition of support vector machine. *Sci Surv Mapp* 41(05):45–49
4. Zhang X, Liu H, Li Y et al (2016) Extraction and recognition of street lamps in vehicle-mounted LiDAR scene. *Surv Mapp Eng* 25(09):50–54
5. Yan WY, Morsy S, Shaker A et al (2016) Automatic extraction of highway lamp poles and towers from mobile LiDAR data. *Opt Laser Technol* 77:162–168
6. Qi X (2009) An improved fast clustering algorithm and its parallelization. Lanzhou University
7. Zhang H, Han Z, Li C (2002) Support vector machine. *Comput Sci* 29(12):135–137



Research on Combination and Optimization of Path Planning Algorithm Based on Bridge Crane

Shuai Shao^(✉) and Wei Lin

Shanghai Institute of Technology, Shanghai 201499, China
1261028688@qq.com

Abstract. The path planning of bridge crane is affected by multiple factors (crane spreader size, cargo size, obstacles, frequent braking of cart and trolley), so it is not simple to carry out traditional path planning, regardless of safety and practical operability. In the paper, a visual sensor equipped with a panoramic camera is used to collect environmental information of the work area to create an environmental map, which provides guarantee for subsequent path planning. The unit decomposition method is used to divide the whole environment map into several sub-regions to generate fuzzy paths, and then the grid method is used to generate a plane grid map based on the unit decomposition method to decompose several sub-areas. The above steps lay the foundation for the subsequent improved A* algorithm to search for a specific path. By combining the three methods of unit decomposition method, grid method and improved A* algorithm to make best use of the advantages and bypass the disadvantages. Thereby better achieving the effect of path planning.

Keywords: Bridge crane · Path planning · Unit decomposition method · Grid method · Improved A* algorithm

1 Introduction

The bridge crane is mainly composed of a cart running mechanism, a trolley running mechanism, a hoisting mechanism (spreader) and a bridge. This type of crane is the most widely used and largest number of lifting appliances [1, 2]. Its appearance has greatly reduced the pressure of hand-handling workers and greatly improved work efficiency. However, with the rapid development of science and technology, people have become increasingly concerned about the safety and reliability of bridge crane handling cargo. The handling of bridge crane is also gradually transformed from the traditional type (operators rely on operational experience, operating bridge crane to transport cargo) to the intelligent type (the system directly generates the path and controls the bridge crane to carry the cargo along the established path). Therefore, as the core problem in the intelligent type: the path planning of the bridge crane has gradually become a hot spot of research.

The path planning problem of bridge cranes is to give the starting point and the end point in the working environment with obstacles, so as to find a safe and collision-free

optimal path (the optimality here is optimal under certain constraints). Therefore, the constraints of the path planning optimization algorithm in this paper are based on security and practical operability, so as to solve the practical optimal path.

Literature [3–5], although the shortest path is solved by genetic algorithm or intelligent water droplet algorithm, the turning point in the path (the node between the oblique line and the straight line) is too much, and it takes more time in actual operation. This is because the cart and the trolley run to the turning point for forward, and the speed needs to be decelerated to zero beforehand. When the path planning has obtained too many turning points. Although the path is optimal, it takes more time and energy consumption, which is not conducive to improving the efficiency of the bridge crane.

Literature [6, 7], extends 8 adjacent nodes in the parent node extension domain to 24 neighbor nodes in the traditional A* algorithm. The searched path is relatively shorter and smoother. However, the spreader and the suspended cargo are directly regarded as the mass point without considering the size of the spreader and the cargo to be hoisted. In this way, the spreader and the suspended cargo are bound to collide with obstacles, causing a safety accident.

Literature [8], an important cause of frequent crane safety accidents is that the distance between the crane and the obstacle is too small.

Therefore, this paper solves the above problems by combining the path planning algorithm with the combination of the unit decomposition method, the grid method and the improved A* algorithm, so as to obtain the optimal path under the constraint conditions.

2 Acquisition of Environmental Map

The bridge crane is divided into two working modes, a two-dimensional working mode and a three-dimensional working mode in actual work. As shown in Fig. 1, the two-dimensional working mode: refers to the bridge crane stopping the work after lifting the cargo to the specified safe height using the spreader at the starting point. The cart and the trolley begin to perform the task of moving the cargo. When the cart and trolley are running above the end point, the spreader works again and the goods are placed at the end. This working mode is mainly driven by the cart and the trolley in their respective horizontal planes, so it is called a two-dimensional working mode. As shown in Fig. 2, the three-dimensional working mode: refers to the bridge crane in the process of moving cargo, the cart, the trolley and the spreader work simultaneously. In this working mode, both the longitudinal displacement action of the spreader and the displacement action of the cart and the trolley on the horizontal plane. Therefore, it is called a three-dimensional working mode. This paper mainly focuses on the path planning of the cart and the trolley from the starting point to the end point in the two-dimensional working mode, and does not involve the three-dimensional working mode.

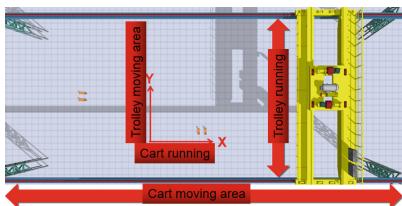


Fig. 1. Two-dimensional working mode

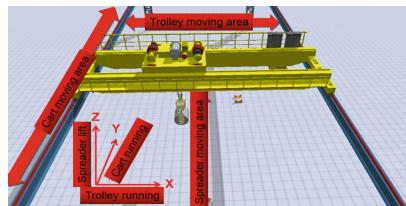


Fig. 2. Three-dimensional working mode

A vision sensor equipped with a panoramic camera is installed in the working area of the bridge crane to collect and analyze the environmental information of the working area of the bridge crane. Thereby obtaining real-time environmental information of the entire work area. On the one hand, it is beneficial for the visual sensor to obtain the latest environmental information in time when the obstacles in the work area change. On the other hand, it is beneficial to avoid the disadvantages of being susceptible to interference from external environmental signals such as photoelectric sensors and sonar sensors. Finally, the image processing of the acquired environmental information and the creation of an environmental map provide the possibility for the bridge crane to follow the fast and accurate path planning.

3 Environmental Map Modeling

After the environmental map is acquired, a $10 * 10$ two-dimensional model is established on the environment map based on the two-dimensional working mode. Among them, the direction in which the cart running mechanism runs on the bridge is the X-axis. The direction in which the trolley running mechanism runs on the cart running mechanism is the Y-axis. The origin of the coordinates is the lower left corner of the bridge.

4 Traditional A* Algorithm

The A* algorithm is a typical heuristic search algorithm. The essence of the search algorithm is to evaluate each search location node in the environment map. Search for the best location node for evaluation, then search from the best location node until the end of the search, and end the search. The evaluation method used in heuristic search is expressed using a cost function (1),

$$F(N) = G(N) + H(N) \quad (1)$$

Among them, $F(N)$ is the cost function of node N. $G(N)$ is the cost from the starting point to the N node. $H(N)$ is the cost from the N node to the end point. A total of two lists are used in the search process, namely the OpenList list and the CloseList list. The point to be tested is placed in the OpenList list, and the detected point is placed in the

CloseList list. During the run, the starting point is first placed in the OpenList list as the current parent. After searching for the adjacent 8 nodes in the extended domain around the parent node, the starting point is placed in the CloseList list, and the 8 nodes that are not in the CloseList list and not the obstacle point are placed in the OpenList list. Calculate the G and H values in these nodes, and solve the nodes with the minimum F value among these nodes. Set the node with the minimum F value as the parent node, and then put the other nodes into the CloseList list. Loop through the above steps. When the end point is placed in the OpenList list, it means that the path has been searched and the current path is saved. When the end point is never placed in the OpenList list, or the OpenList list is always empty, there is no path and the path search fails.

5 Combination and Optimization of Path Planning Algorithm

5.1 Combination of Path Planning Algorithm

In this paper, the path planning algorithm combines the three methods of unit decomposition method, grid method and improved A* algorithm to make best use of the advantages and bypass the disadvantages. Thereby better achieving the effect of path planning. The combination of the unit decomposition method and the grid method is beneficial to increase the flexibility of the path search. Avoid the erratic search of the bridge crane on the fuzzy path generated by the unit decomposition method. At the same time, the improved A* algorithm uses the heuristic function to speed up the specific path search speed based on the above two methods, and also avoids the need for the global decomposition of the unfamiliar environment map by the unit decomposition method and the grid method. Through the above methods, the search efficiency is greatly improved [9–11].

5.2 Optimization of Path Planning Algorithm

1. Based on the two-dimensional model of the environment map, when the two-dimensional grid is established by the grid method, the selection of the side length of each two-dimensional grid is very important. Because, on the one hand, if the grid side length is larger. Although the path planning time is shorter, the positioning of the bridge crane in the actual operation is less accurate. On the other hand, if the grid side length is too small. Although the actual environment is divided accurately, the turning point of the bridge crane will be correspondingly increased, and the braking of the cart and the trolley will be correspondingly increased, so that the bridge crane will consume more time in actual operation. At the same time, the anti-sway performance of the bridge crane will be deteriorated to some extent. Therefore, the larger of the spreader radius and the cargo radius is selected as the size of each grid, which can improve the safety and reliability of the path planning.
2. In order to improve work efficiency and save running time, the cart and the trolley should take as many straight lines as possible and less oblique lines. Therefore, the

8 adjacent nodes in the extended area around the parent node in the traditional A* algorithm are reduced to 4 adjacent nodes. The angle of the coordinated motion of the cart and the trolley is limited to an integral multiple of $\pi/2$, which can minimize the occurrence of turning points. Not only can it improve work efficiency, but it can also reduce the risk of bridge crane. Because the bridge crane is in the process of running, the frequent acceleration and deceleration of the cart and the trolley will increase the swing frequency and amplitude of the cargo. At the same time, it is not conducive to the positioning of the bridge crane during the automatic operation, which affects the work efficiency.

3. In the traditional A* algorithm, adjacent nodes of the extended area around the parent node are placed in the cost function F for descending order, and the node with the minimum F value is obtained. Sometimes the F values of multiple nodes are equal. In this case, the priority of the adjacent nodes needs to be set. Starting from the node to the right of the parent node, it has the highest priority and rotates counterclockwise. The priorities of adjacent nodes are arranged in order, which facilitates the orderly selection of node selection.
4. The traditional A* algorithm performs a path search along the starting point to the end point. In this paper, in order to speed up the path search, the efficiency of the bridge crane is improved. It is proposed to improve one-way search to two-way search. That is, the search results from the starting point to the end point and the end point to the starting point, but sometimes the two paths in the two-way search result are inconsistent, and the short path is taken as the standard.

6 Specific Use of Path Planning Algorithm

Regarding the use of path planning algorithms, this paper considers the security from the actual operation. First, the unit decomposition method is used to divide the entire two-dimensional model into several sub-regions (see Fig. 3). Each sub-region acts as a unit, with the unit as the vertex and the adjacent relationship between the units as the side, thus forming a connected graph (see Fig. 4). Find the unit containing the start point and the end point in the connected graph, and search for the fuzzy path connecting the start unit and the end unit. It lays the foundation for the specific path search of the improved A* algorithm. Because the subsequent improved A* algorithm will search based on the fuzzy path searched by the unit decomposition method. Moreover, the ultimate goal of the path search method is for practicality. In the actual operation, the operator only needs to inform the host computer of the starting point and the end point, and the system will automatically search the path. When the path is searched out. The bridge crane will carry the cargo according to the searched path, and the specific path search of the improved A* algorithm is based on the unit decomposition method to search for the fuzzy path. This guarantees the accuracy of the searched path and the security in actual operation (double verification). At the same time, it can also save the time of path search (in this case, it is no longer necessary to consider obstacles and path search into an infinite loop). As shown in Fig. 4, we can easily see that there are three fuzzy paths from the start point to the end point. The starting point from unit 1 to

unit 2 or unit 19 will enter an infinite loop with no way to go. Need to turn around and search again, which will bring uncertainty to the path search. At the same time, from the three fuzzy paths, we are not difficult to select, from 1 unit, after 4, 6, 7, 9, 11, 12, 14, 16 unit, and finally to 18 unit for the optimal path among the three fuzzy paths (here the fuzzy path is named the optimal fuzzy path).

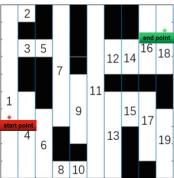


Fig. 3. Several sub-regions

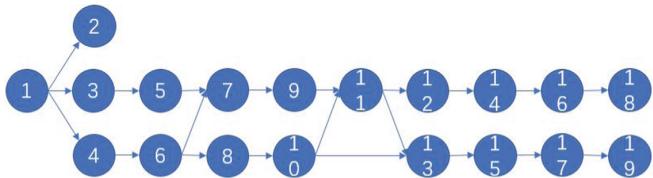


Fig. 4. Connected graph

Secondly, a grid map is created on the two-dimensional model of the environment map using the grid of the same size (100 grids). The size of each grid in the grid map is the larger of the spreader radius and the cargo radius of the lift. There are two types of grid in the grid map. One is that the free grid represents no obstacles; the other is that the barrier grid represents obstacles. Each grid is identified using cartesian coordinates. The center coordinate of each grid is the coordinates of the grid and represents the grid. The origin of the coordinates is the lower left corner of the grid map. The free grid is shown in white and the barrier grid is shown in black.

Finally, based on the unit decomposition method and the grid method, the improved A* algorithm is used to perform specific path search. That is, the two-way search is performed on the plane grid map of the fuzzy path solved by the unit decomposition method. With the starting point as the current parent node (place the OpenList list), 4 adjacent grids are searched for the surrounding extended area (the integer multiple angle of $\pi/2$ is searched). The starting point is placed in the CloseList list and the 4 adjacent grids just searched are placed into the OpenList list. The grid with the lowest cost function F value in the adjacent 4 grids is used as the next parent node (When multiple grid cost functions have the same F value, the adjacent nodes on the right have the highest priority. Counterclockwise rotation, the priorities of adjacent nodes are arranged in order). Search backwards until the end point is placed in the OpenList. In this process, there is no need to consider the obstacle grid, and the fuzzy path searched by the unit decomposition method has eliminated the obstacle grid. After completing the search of the specific path, compare the specific paths two-way searched on all the different fuzzy paths (there may be multiple fuzzy paths solved by the unit decomposition method), and find out the optimal path (evaluation indicators: minimum turning point, shortest path, shortest path search time), which is the path obtained by the final path planning.

7 Simulation Result Analysis

This paper uses MATLAB 2017b software to simulate and based on the optimal fuzzy path. The classical classic A* algorithm in the path search algorithm is compared with the optimization algorithm proposed in this paper. Solve the optimal path solution based on the constraint security and actual operability. The path search time T from Figs. 5, 6, 7, 8, 9 and 10 is: 0.2829 s; 0.2151 s; 1.094 s; 0.5435 s; 0.3427 s; 0.3309 s. It can be seen that the search from the starting point to the end point in the path search takes longer than the end point to the starting point. Explain that the two-way search in the optimization algorithm is practical and better. From the point of view of the turning point of the path, the path of the traditional A* algorithm is 4 turning points (see Fig. 5), and the optimization algorithm is 3 turning points (see Fig. 10). It shows that the optimization algorithm is optimized by 8 adjacent nodes to 4 adjacent nodes, which is practical and better. If viewed from the path length alone, the path length of the traditional A* algorithm is 11 (see Fig. 5), even shortened to 10.8 through the reverse search path (see Fig. 6), which is shorter than the path 14 after the optimization algorithm (see Fig. 10). However, from the perspective of constraints, the paths shown in Figs. 5 and 6 are not desirable in the actual operation of the bridge crane. Because the figure is to show the searched path more clearly, the spreader and the hanging cargo grid are represented by the grid center coordinate points. The size of the actual spreader and the cargo to be lifted is as large as the size of each grid in the figure, so the path has coincided with the obstacle (the coordinates are (3, 4), (5, 8), (8, 8)). The path of the optimization algorithm does not coincide with the obstacle and collides. It shows that the optimization algorithm is practical under the constraint conditions.

As shown in Figs. 8 and 10, in the path search, the F values of multiple nodes appearing in adjacent nodes around the parent node are the same. As shown in Fig. 8, the node with the highest priority is selected when the multiple F values are equal in the optimization algorithm in the node selection process with the same F value. As a result, the number of turning points increases, and the path search time becomes longer. When the multiple F values in the optimization algorithm are equal, the priority of the node selection is set, which is practical and better. In summary, under the constraint condition, the path searched by the optimization algorithm is the optimal path, the path search time is 0.3309 s, the number of path turning points is 3, and the path length is 14.

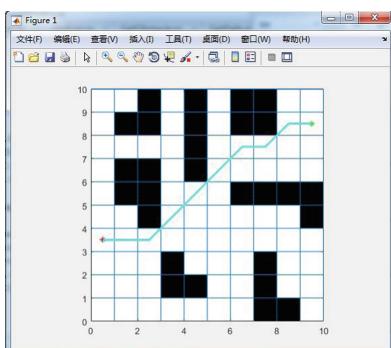


Fig. 5. Forward search (8 adjacent grids)

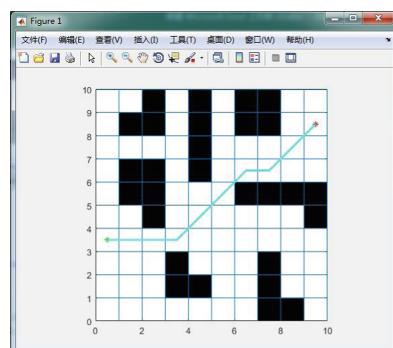
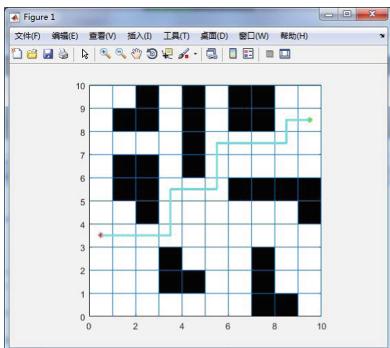
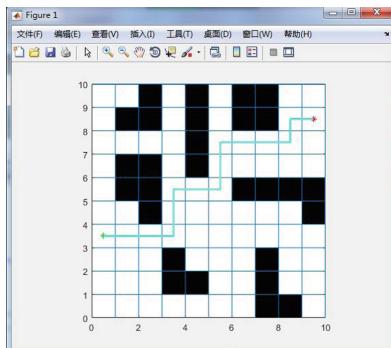
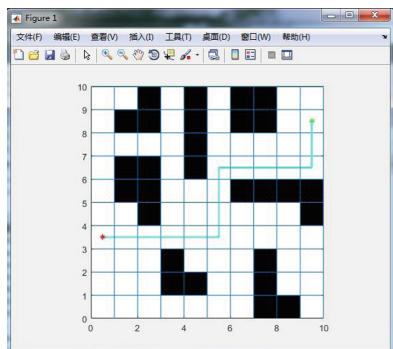
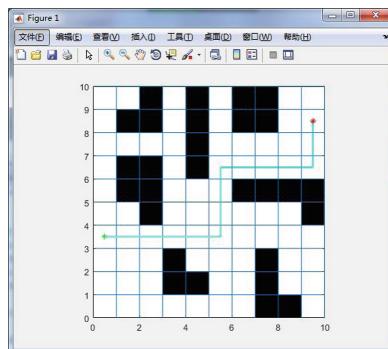


Fig. 6. Reverse search (8 adjacent grids)

**Fig. 7.** Forward search (4 adjacent grids)**Fig. 8.** Reverse search (4 adjacent grids)**Fig. 9.** Optimized forward search (4 adjacent grids)**Fig. 10.** Reverse search after optimization (4 adjacent grids)

8 Conclusion

The ultimate goal of the study of path planning algorithms is to produce economic value for better practical use. Therefore, the constraints of the actual application scenario must be taken into account in this process. According to the actual operating conditions of the bridge crane, the optimal path under the constraint condition of the bridge crane in the actual running process is solved by combining and optimizing the three algorithms. The feasibility of combining and optimizing the three algorithms in this paper is verified. Thereby improving the safety and reliability of the bridge crane during operation.

References

1. Yu J, Li Y, Liu J (2018) Overview of patent technology of bridge crane. *Technol Innov Appl* 30:25–26
2. Wang J (2014) Status and development trend of cranes. *Sci Technol Inf* (13):161

3. Dan Z (2013) Intelligent operation system and path planning research for overhead crane. Dalian University of Technology
4. Li X, Chen Z (2017) Research on path planning method of bridge crane hoisting. *J Taiyuan Univ Sci Technol* (6):445–449
5. Wang X, Ye T (2018) Research on robot path planning based on improved A* algorithm. *Comput Meas Control* 26(07):282–286
6. Li D, Yang X (2019) Path planning of mobile robot based on improved intelligent water drop algorithms. *Comput Technol Dev* (11)
7. Cui B, Wang M, Duan Y (2018) Path planning for A* algorithm based on searching 24 neighborhoods. *J Shenyang Univ Technol* 40(2):180–184
8. Gao Q (2011) Risk analysis and accident prevention on crane. Northeastern University
9. Chen Y, Zhu B (2014) Complete coverage path planning of mobile robot based on partition of unity method. *Equip Manuf Technol* 4:148–149
10. Wang W, Wu Y, Zhang X et al (2013) Research of the unit decomposing traversal method based on grid method of the mobile robot. *Tech Autom Appl* 32(11):34–38
11. Liu K, Zhang C, Yi H (2005) Complete coverage navigation of cleaning robots using square-cell-based map. *Syst Simul Technol* (3): 136–140



Extrinsic Parameter Co-calibration of a Monocular Camera and a LiDAR Using Only a Chessboard

Shuguang Tu^{1(✉)}, Yafei Wang¹, Chenyang Hu², and Yuneng Xu²

¹ Shanghai Jiao Tong University, Shanghai 200240, China

tushugaung@sjtu.edu.cn

² SAIC Volkswagen, Shanghai 201800, China

Abstract. Cameras and LiDARs (Light Detection and Ranging) are often equipped for autonomous driving applications, and need to be calibrated together to unify the coordinates. Conventional calibration approaches need complex calibration boards which are inconvenient to use. This paper proposes a novel and accurate calibration method using a simple rectangle planar chessboard with known size, and the key idea is to use 2D-3D point correspondences in LiDAR and camera frames. The accuracy of the calibration is improved by fusing the information of a lot of corresponding point clouds and images. In addition, the whole calibration process requires less manual intervention. Experiments show that this approach is more accurate compared to a conventional widely used method.

Keywords: Extrinsic calibration · Monocular camera · LiDAR · EPnP

1 Introduction

Nowadays, robots and autonomous vehicles are equipped with a variety of sensors such as cameras, LiDARs and IMUs (Inertial Measurement Unit). Each sensor can provide a lot of valuable information. But different sensors have different advantages. For instance, the LiDAR can provide accurate 3D locations of objects and the camera can provide rich color and feather information which can be used to detect objects of interest [16]. Therefore, it is necessary to fuse complementary information from different sensors to obtain higher quality information. The LiDAR and the camera are often used together to build maps [1], detect objects [19], solve navigation tasks and reconstruct scenes. The LiDAR and the camera are usually mounted at different locations. Therefore, it is necessary to make an extrinsic calibration to know their relative position and orientation before fusing their information [20]. The main task of extrinsic calibration between a LiDAR and a camera is to find the 3D coordinates of feather points (feature points are the corners of the chessboard in this paper) in the LiDAR frame.

Some papers have proposed many methods of the extrinsic calibration between the monocular camera and the LiDAR. Dhall et al. [2] used 3D-3D point correspondences and ICP algorithm to do the extrinsic calibration. But this method was required to use stereo cameras and ArUco markers to get 3D coordinates of feather points. Park et al.

[3] proposed a new method which using a polygon checkerboard. It got 3D coordinates of checkerboard edges by calculating intersection points of straight lines fitting the side of the checkboard [10]. This method ignored the errors resulted from the lateral distance resolution of the LiDAR. The method proposed by Ha [4] has the same shortage.

This paper proposes a new method to make an accurate calibration between the monocular camera and the LiDAR. This method is more convenient and has low requirements for the condition of the calibration compared to the method mentioned above. The whole process of the calibration requires only an ordinary chessboard with known size and requires less manual intervention. this method can keep high accuracy of the calibration at the same time. Most existing methods only try to minimize the errors resulted from the vertical distance resolution of the LiDAR and ignore the influence of the lateral distance resolution of the LiDAR. The method in this paper minimizes the errors resulted from both the vertical distance resolution and the lateral distance resolution of the LiDAR, which is achieved by fusing the information of a lot of corresponding images and point clouds. The experiments demonstrated the accuracy of the method proposed in this paper.

2 Calibration Model Setup

The calibration system consists of a monocular camera, a LiDAR and a chessboard with known size (see Fig. 1).

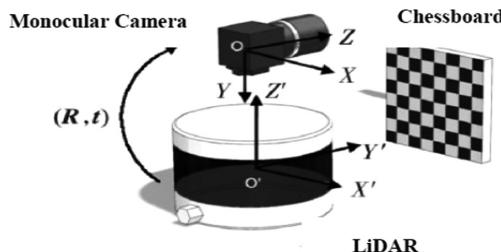


Fig. 1. Calibration system

The image data captured by the camera are formed by two-dimensional coordinate system (u, v) and the range data of the 3D point clouds are represented by three-dimensional coordinate system (x, y, z). Then the transformation from a 3D point (x, y, z) to a 2D point (u, v) can be represented by $P_{uv} = KTP_L = MP_L$:

$$\begin{aligned} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \end{aligned} \quad (1)$$

The \mathbf{K} is the intrinsics of the camera, which is known by the intrinsic calibration of the camera [13]. Our goal is to get transformation matrix \mathbf{T} by estimating the projection transformation matrix \mathbf{M} consisting of the intrinsic and extrinsic parameters [14]. The Eq. (1) can be written as the following equations:

$$u = \frac{m_{11}x + m_{12}y + m_{13}z + m_{14}}{m_{31}x + m_{32}y + m_{33}z + m_{34}} \quad (2)$$

$$v = \frac{m_{21}x + m_{22}y + m_{23}z + m_{24}}{m_{31}x + m_{32}y + m_{33}z + m_{34}} \quad (3)$$

The Eq. (2) and the Eq. (3) can be generated using every pair of 2D-3D corresponding point. Therefore, a sufficient number of the 2D-3D corresponding points are needed to estimate the projection transformation matrix. The EPnP (Estimation of Perspective-n-Point) algorithm [6, 17] used in this paper at least needs four pairs of the 2D-3D corresponding points [18]. In other words, the accuracy of the extrinsic calibration is determined by the accuracy of the 2D-3D corresponding points. The corresponding 2D coordinates in the image frame [11] can be found very accurately by the OpenCV (see Fig. 2). How to get the accurate 3D coordinates in the LiDAR frame is the main problem which has a big impact on the accuracy of the extrinsic calibration. Therefore, this paper proposed a new method to improve the accuracy of the corresponding 3D coordinates in the LiDAR frame by fusing the information of a lot of corresponding images and point clouds, which is the main innovation of this paper.



Fig. 2. Extraction of 2D coordinates

After the estimation of the projection transformation matrix through the EPnP algorithm, it is necessary to minimize the reprojection errors (see Eq. 4) between 3D coordinates and 2D coordinates, which is achieved by using G2O (Graph Optimization)

algorithm [8] to make bundle adjustment (BA) [7] for the projection transformation matrix. The cost function is:

$$\arg \min \|KTP_L - P_{uv}\|^2 \quad (4)$$

This problem is the non-linear least square problem [12] which can be solved by the Levenberg -Marquardt method (LM) [9].

3 3D Coordinates Extraction

How to extract the corresponding 3D coordinates accurately is the innovation of this paper. This section would introduce it in detail. The 3D coordinates of the feature points in the LiDAR frame can't be obtained from the point cloud directly, which is resulted from the sparsity of the point cloud. Four steps are taken in this paper to find the accurate 3D coordinates from the point cloud:

- (1) Data Preparation and Preprocessing
- (2) Point Cloud Clustering: Plane Segmentation
- (3) Find the Rough 3D Coordinates of the Feature Points
- (4) Optimize the 3D Coordinates of the Feature Points

3.1 Data Preparation and Preprocessing

The data used for extrinsic calibration between the monocular camera and the LiDAR is required to contain a lot of corresponding images and point clouds. When collecting the data, the chessboard is required to move very slowly in vertical and horizontal direction respectively to keep the data containing enough information which would be used to optimize the estimated result in step 3.4. In addition, if the data is stored in.bag format, it is necessary to convert it into images (.jpg) from \mathbf{IM}_I to \mathbf{IM}_k and point clouds (.pcd) from \mathbf{OPC}_I to \mathbf{OPC}_k which are associated with the timestamp.

3.2 Point Cloud Clustering: Plane Segmentation

The chessboard used in this paper is a rectangular planar board. The plane model (Eq. 5) can be calculated through three points which are non-collinear theoretically. In fact, the points in point cloud contain the noise which makes the points hitting on the chessboard are not on the plane of the chessboard finally.

$$ax + by + cz + d = 0 \quad (5)$$

Therefore, the Random Sample Consensus (RANSAC) [15] based method is applied to segmenting the plane of the chessboard. It tries to make the most of only the point information corresponding to the plane by using the error value of the sensor itself as the threshold value of RANSAC. The all valid points are projected to the corresponding segmented plane to make up new point clouds from \mathbf{NPC}_I to \mathbf{NPC}_k . It's

obvious that every new point cloud is on the corresponding plane fitting the chessboard (see (b) in Fig. 3).

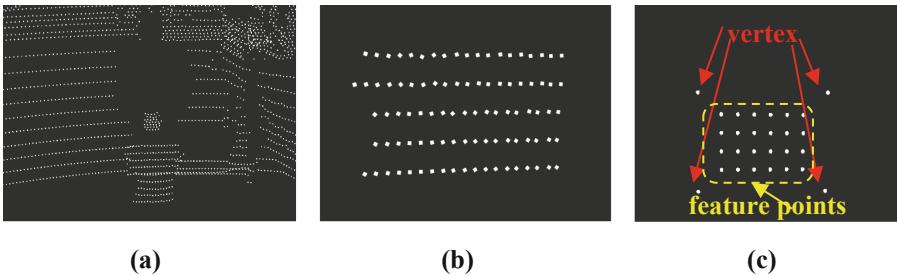


Fig. 3. The process of extracting the 3D coordinates. (a) is the original point cloud OPC_I . (b) is the new point cloud NPC_I on the segmented plane. (c) is the point cloud which only contains four rough vertexes and twenty-four rough feature points.

3.3 Find the Rough 3D Coordinates of the Feature Points

The goal of this step is to find rough 3D coordinates of feature points from the point cloud NPC_I . Because the 3D coordinates of the feature points are calculated according to the four vertexes of the chessboard. Actually, the goal is to find the four vertexes of the chessboard. The first thing is to find a rough point on the plane segmented in the last step as a vertex of the chessboard. This paper makes the point located at the top left of the NPC_I as the top left vertex of the chessboard. The other three vertexes of the chessboard can be calculated according to the rectangle model of the chessboard that is the known parameter. After that, the rough 3D coordinates of the feature points can be estimated successfully using the positions of the four vertexes (see (c) in Fig. 3). Certainly, the 3D coordinates of the feature points are inaccurate in this step and they would be optimized in next step using the information of all images (from IM_I to IM_k) and point clouds (from NPC_I to NPC_k).

3.4 Optimize the 3D Coordinates of the Feature Points

The optimization of 3D coordinates is the most important step in this paper. It makes the most contribution to the accuracy of the extrinsic calibration. Through the steps above, the rough 3D coordinates and corresponding 2D coordinates of feature points have been obtained from the NPC_I and IM_I respectively. Therefore, the rough projection transformation matrix can be solved by EPnP algorithm and optimized by G2O algorithm. Until now, only one point cloud NPC_I and one image IM_I are used to estimate the projection transformation matrix and no optimization of the rough 3D coordinates of the feature points is carried out, which would result in the inaccuracy of the projection transformation matrix estimated. In order to improve the accuracy of the projection transformation matrix, this paper would make full use of the information of all images (from IM_I to IM_k) and point clouds (from NPC_I to NPC_k) collected in step 3.1.

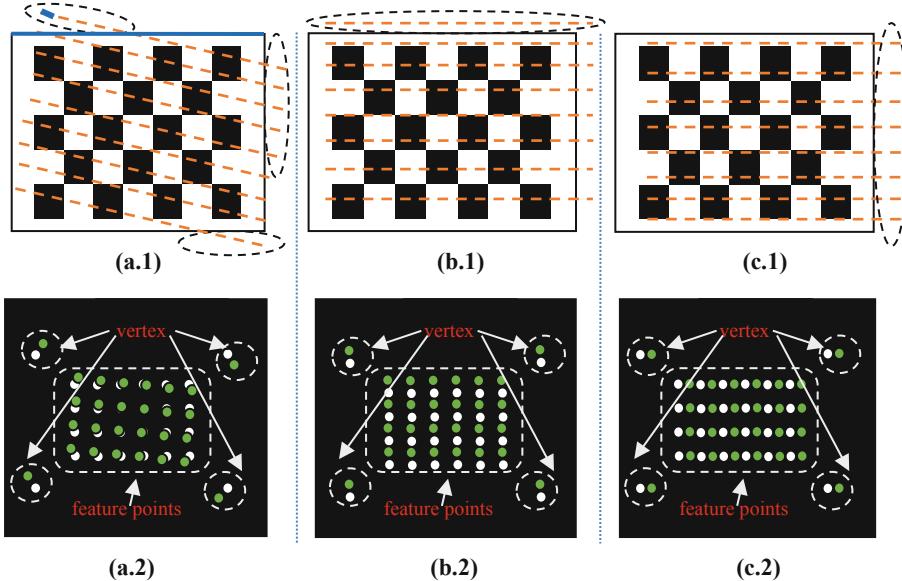


Fig. 4. The green points in **a.2(b.2, c.2)** are the new 3D feature points which have been carried out the rotation and translation optimizations respectively according to the location of the outliers in **a.1(b.1, c.1)**.

Firstly, the other images and point clouds are used to illustrate the inaccuracy of the estimated projection transformation matrix. The point clouds from NPC_1 to NPC_k are projected into the corresponding images from IM_1 to IM_k in turn. If there are some points (called outliers) out of the chessboard in the images (see the circled points in a.1, b.1 and c.1 of Fig. 4) and the distance between the outlier and corresponding line of rectangular chessboard exceeds the specific threshold determined by the resolution of the camera (for instance, as a.1 in Fig. 4 shows, if the blue point is an outlier, the corresponding line is the blue line.), the projection transformation matrix is considered to be inaccuracy. Secondly, the other images and point clouds are also used to optimize the 3D coordinates of the four vertexes in NPC_1 (white vertexes of a.2, b.2 and c.2 in Fig. 4), which would optimize the 3D coordinates of the feature points in NPC_1 (white feature points of a.2, b.2 and c.2 in Fig. 4). In other words, the position of the rectangle made up of four vertexes of the chessboard in NPC_1 is required to be changed. Because the rectangle is on the plane segmented in step 3.2, the freedom of the rectangle is three: a rotation and two translations (in two different directions) on the plane. It's obvious to take three steps to optimize the position of the rectangle in NPC_1 according to the position of the outliers in other images. The specific steps are in the following:

- (1) If there are some outliers distributed besides both parallel sides of the chessboard (see a.1 in Fig. 4) and the distance between the outlier and corresponding line is beyond the threshold, the rotation is necessary in NPC_1 . After a little rotation for the rectangle made up of four vertexes of the chessboard is carried out in NPC_1 , four new vertexes in NPC_1 would be obtained which makes the new 3D

coordinates of the feature points in NPC_1 be estimated again. Finally, the new projection transformation matrix would be solved again using the new 3D coordinates. Every time the new projection transformation matrix is estimated, the accuracy of the projection transformation matrix would be illustrated by all the point clouds and images. If the result doesn't satisfy the requirement, a little rotation would be carried out again in NPC_1 until the projection transformation matrix estimated satisfies the requirement.

- (2) This step would be carried out after the step (1) is finished. If there are some outliers distributed besides only one side of the long parallel sides (see b.1 in Fig. 4) and the distance between the outlier and corresponding line is beyond the threshold, the translation along the short parallel sides is necessary. The process of the optimization is same with the step (1) which makes a little translation many times until the projection transformation matrix satisfies the requirement.
- (3) This step would be carried out after the step (2) is finished. If there are some outliers distributed besides only one side of the short parallel sides (see c.1 in Fig. 4) and the distance between the outlier and corresponding line is beyond the threshold, the translation along the long parallel sides is necessary. The process of the optimization is same with the step (2) which makes a little translation many times until the projection transformation matrix satisfies the requirement.

Because the data used in the process of the calibration contains the images and point clouds that the chessboard is at different locations (see 3.1 section), the result would become more and more accurate through many optimizations. When every optimization step is finished, the accuracy of the projection transformation matrix would reach the requirement.

4 Experiments and Analysis

The calibration experiment uses the sensors and tools in the following (see Fig. 5):

- (1) Camera: resolution (1280*720)
- (2) 3D LiDAR: Hesai P40
- (3) Chessboard: size (6*4), square (0.07 m*0.07 m), length 0.55 m, width 0.37 m (the length and width of a chessboard should be more than 0.50 m and 0.30 m respectively.)



Fig. 5. The sensors and their locations on the vehicle

Firstly, thirty images from \mathbf{IM}_1 to \mathbf{IM}_{30} and thirty corresponding point clouds from \mathbf{OPC}_1 to \mathbf{OPC}_{30} were used in this experiment. the rough projection transformation matrix is estimated according to the \mathbf{IM}_1 and \mathbf{OPC}_1 . In order to show the result intuitively, the tester and a traffic cone are used to be the objects of reference. As (a) in Fig. 6 shows, a whole point cloud \mathbf{OPC}_i projected into the image according to the rough projection transformation matrix don't coincide with the objects in the corresponding image (the green points represent the tester and the traffic cone).

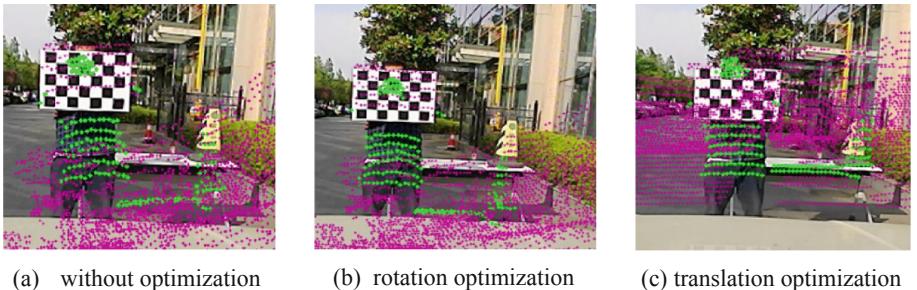


Fig. 6. The result of the calibration with step by step optimizations.

Secondly, new 3D feature points are calculated through the optimization of the rotation. And the projection transformation matrix would be estimated again and the result is showed in (b) of Fig. 6. Although the point cloud projected still don't coincide with the objects. It's obvious that the result is better than the previous result. Finally, the translation optimization in two directions makes the 3D feature points more accurate. The projection transformation matrix would also become more accurate accordingly. The result is much better than the previous result and showed in (c) of Fig. 6. It's easy to see the point cloud projected into the images coincide with the objects in the image very well, which can indicate the extrinsic calibration is successful.

The experiment demonstrated the good result of the extrinsic calibration could be obtained using the method proposed in this paper. The contrast experiment was carried out to demonstrate the effectiveness of the method proposed in this paper. The method proposed by Pusztai [5] makes the extrinsic calibration of the camera and the LiDAR by using the boxes. It makes the lights of the LiDAR hit the three planes of the box. Then three Eqs. (5) of the three planes would be obtained by making plane segmentation which can be used to calculate the 3D coordinate of a vertex of the box. More 3D coordinates can be obtained by using more boxes. Finally, the extrinsic parameter would be estimated by EPnP algorithm. In order to compare the accuracy of two methods. The contrast experiment (see (a) in Fig. 7) was carried out with the same condition.

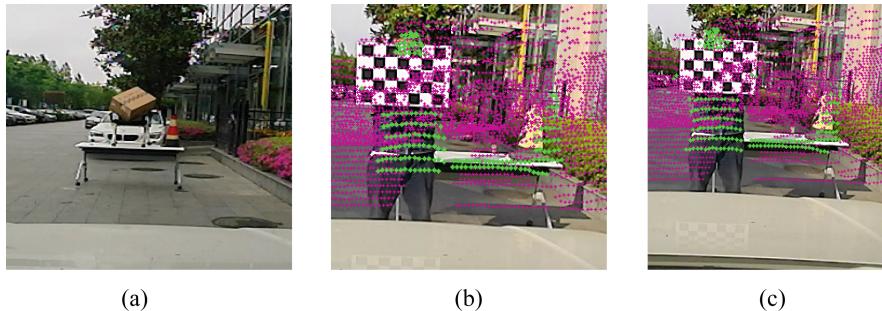


Fig. 7. (a) Is the contrast experiment. (b) Is the result of the calibration using the box. (c) Is the result of the calibration using the method proposed in this paper.

The result of the contrast experiment indicates the method proposed in this paper is more effective compared to the method using the boxes. Point clouds are projected into the images according to the result of the calibration using the method in [5]. The coincidence rate of the points and objects is not higher than the method proposed by this paper (see b and c in Fig. 7).

5 Conclusion

This paper introduces a new method for the extrinsic calibration between the LiDAR and monocular camera. The new method only requires a chessboard with known size. It's not only convenient but also accurate. This approach optimizes the 3D coordinates of the feature points on the chessboard by making full use of many images and corresponding point clouds. It can get over the influence of both the vertical resolution and lateral resolution of the LiDAR. This method is achieved through the algorithm designed by C++ and compared to the state-of-the-art method. The experiments demonstrate the accuracy, convenience and effectiveness of the proposal method.

References

1. Hayat AA et al (2019) A geometric approach for kinematic identification of an industrial robot using a monocular camera. *Robot Comput-Integr Manuf* 57:329–346
2. Dhall A, Chelani K, Radhakrishnan V et al (2017) LiDAR-camera calibration using 3D-3D point correspondences
3. Park Y et al (2014) Calibration between color camera and 3D LIDAR instruments with a polygonal planar board. *Sens (Basel)* 14(3):5333–5353
4. Ha J-E (2015) Improved algorithm for the extrinsic calibration of a camera and laser range finder using 3D-3D correspondences. *Int J Control Autom Syst* 13(5):1272–1276
5. Pusztai Z, Hajder L (2017) Accurate calibration of LiDAR-camera systems using ordinary boxes. In: IEEE international conference on computer vision workshops. IEEE, Venice, pp 394–402

6. Lepetit V et al (2008) EPnP: an accurate O(n) solution to the PnP problem. *Int J Comput Vis* 81(2):155–166
7. Wani SU, Anuar A (2011) Calibration of high resolution digital camera using self-calibration bundle adjustment method. In: IEEE 7th international colloquium on signal processing and its applications. IEEE, Malaysia, pp 137–141
8. Rainer K, Giorgio G et al (2011) A general framework for graph optimization. In: IEEE International conference on robotics and automation (ICRA). IEEE, Shanghai
9. Lourakis MIA (2005) A brief description of the Levenberg-Marquardt algorithm implemented by levmar. *Found Res Technol* 4:1–6
10. Harris C, Stephens M (1988) A combined corner and edge detector. In: Proceedings of the 4th alvey vision conference, Manchester, pp 147–151
11. Yasushi K, Hiroshi K (2004) Detection of planar regions with uncalibrated stereo using distributions of feature points. In: British machine vision conference, London
12. Marquardt Donald W (1963) An algorithm for least-squares estimation of nonlinear parameters. *Soc Ind Appl Math* 11:431–441
13. Mirzaei FM, Kottas DG, Roumeliotis SI (2012) 3D LIDAR-camera intrinsic and extrinsic calibration: identifiability and analytical least-squares-based initialization. *Int J Robot Res* 31(4):452–467
14. Xiang G (2017) Fourteen lectures for the visual SLAM. Electronic Industry Press, Beijing
15. Martin AF, Robert CB (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM* 24(6):381–395
16. Jun W, Wu T, Zheng Z (2016) LIDAR and vision based pedestrian detection and tracking system. In: IEEE international conference on progress in informatics & computing. IEEE, Shanghai, pp 118–122
17. Gerald S, Axel P (2008) Globally optimal o(n) solution to the PnP problem for general camera models. In: BMVC, Leeds University
18. Zheng Y, Kuang Y, Shigeki S, Kalle A, Masatoshi O (2013) Revisiting the PnP problem: a fast, general and optimal solution. In: The IEEE international conference on computer vision (ICCV). IEEE, Sydney, pp 2344–2351
19. Etienne V, Robert L (2001) Detecting planar homographies in an image pair. In: The 2nd international symposium on image and signal processing and analysis. IEEE, Hong Kong, pp 182–187
20. Frederic D, Olivier F (2001) Straight lines have to be straight: automatic calibration and removal of distortion from scenes of structured environments. *Mach Vis Appl* 13(1):14–24



Detection of Load Loss by Using a New Bolt Sensor

Lutao Yan^{1(✉)}, Qi Wang¹, Jinhai Wang², and Bin Zhang¹

¹ School of Automation, Beijing University of Posts and Telecommunications,
Beijing 100876, China

lutaoyan@hotmail.com

² Quality Department, Beijing Aerospace Propulsion Institute, Beijing 100076,
China

Abstract. A new design of bolt sensor is proposed for real time detection of stress relaxation in bolted connections. Full bridge strain-gage circuit is installed in normal bolt and changes in the voltage signal are detected when the bolt is subjected to clamp load loss. The experimental and analytical results are presented and discussed. The proposed bolt sensor sensitivity is defined by series connection with high precision force sensor. The bolt load can be directly measured, and the method is successful in detecting the load loss both in the static and dynamic environments.

Keywords: Bolt sensor · Full bridge strain-gage · Load loss · Dynamic environment

1 Introduction

In many machines and structures, bolted connections are used to connect components composed of dissimilar materials or complicated geometries as well as providing a means of access and maintenance. However, loss of initial clamping load or its decrease often occurs due to creep in bolted joints [1–3]. This may lead to degraded performance or serious injury in machines/structures system, especially in important places [4].

Clearly, Bolted joint loosening is equal to the change of tightening axial tension, and can be detected from the change of tightening axial tension. There are two methods for the axial tension measuring. The first is using an ultrasonic wave and an electromagnetic resonance. Mewer [5] investigated the technique using high frequency input signals and transmittance functions to evaluate changes in bolt tensioning level. Their investigation showed that the techniques were able to distinguish changes in a damage index. Okugawa [6] also proposed a new method of a bolt loosening detection by adopting a smart washer which uses a piezoelectric material with sub-space state space identification algorithm. However, the detection method which used the self-sensing and actuation function to resonance frequency change of the bolted joint is strongly depend on measurement equipment, algorithm stability and calculating precision. Another method, electrical resistance strain gages are mounted on the unthreaded shank of the bolt to measure load [7]. Typically, holes should be drilled in

bolt head to allow the gage connection wires to pass through without interfering with the rest of the joint. However, this structure may result in disturbance in the assemble process. Moreover, it is difficult to arrange full bridge strain-gage circuit on the loaded bolt.

In this paper, a new design of bolt sensor is developed for the early detection of bolt loosening in multi-type bolted joints. In order to obtain amplified signal and temperature compensation, full-bridge strain gage is designed and embedded in a normal bolt. The method for tracking real status of bolts is experimentally validated in both of static and dynamic environments.

2 System Description

Figure 1 presents the schematics of the proposed bolt sensor. A longitudinal bore within both the head and shank portion of the bolt is created. Four strain gauges are placed on the inner surfaces to provide a full bridge strain-gage circuit. Thus, the longitudinal extension or compression of the bolt can be directly revealed. Moreover, this sealed construction is impervious to liquid or vapor and can be made an effective permanent part of a test part. Bolts of diameters from 6 mm to 20 mm are manufactured, as shown in Fig. 1(b).

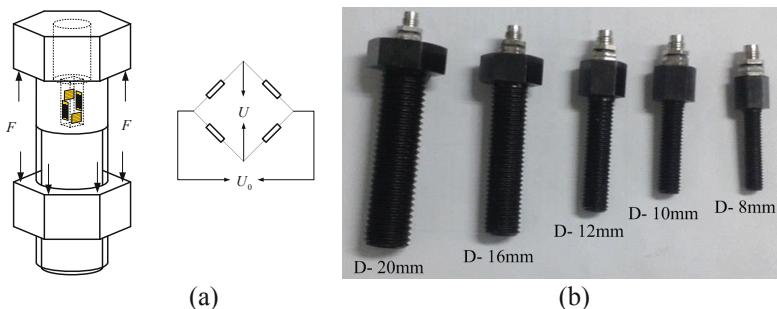


Fig. 1. Bolt sensor (a) Schematic of the bolt sensor (b) Manufactured bolt sensor

Because the proposed bolt sensor should be calibrated and be verified experimentally before it is used in the bolt application, the experimental apparatus for the bolt sensor calibration and force measurement is conducted, as shown in Fig. 2. The sensitive element is series fixed with a high-precision force sensor to obtain same load. External load on the bolted joint can be adjusted by the nut in the structure. Here, bolt with 20 mm diameter is used. Strain in the bolt and force sensor is measured by resistance strain gauge. A mass is fixed on a cantilever in order to simulate dynamic environment. Piezoelectric accelerometer model PCB-20 (100 mV/g) is used to measure the acceleration.

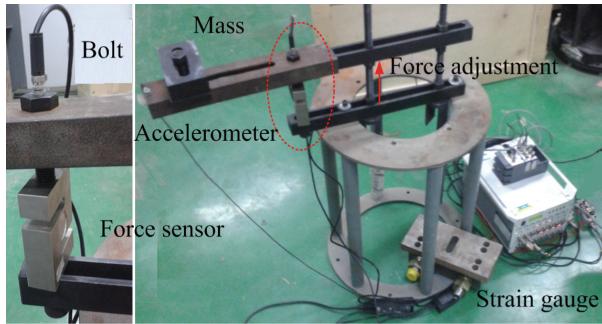


Fig. 2. Experimental set-up

3 Results and Discussion

The calibration plot between voltage output of bolt sensor and force sensor is presented in Fig. 3. As the sensitivity of the force sensor is given ($30\text{kN}/24.5\text{ mV}$), the proposed bolt sensor sensitivity can be defined. Based on repeated experiments, loading and unloading curves are symmetric, which means that the output has a good stability and reproducibility.

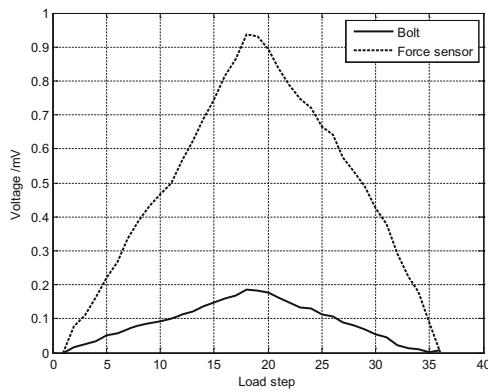


Fig. 3. Voltage output of bolt sensor and force sensor

The measured force curves with time by the force sensor and bolt are shown in Fig. 4. Clearly, the trend is consistent with the result of force sensor, both for the loading and unloading process. Details of over loading can be obtained for each step as seen in the zoomed figure. The maximum error can be seen at the second step of unloading, there are 2.5kN and 2.3kN for force sensor and the bolt. This slight difference could be due to normal variation in the bolt mechanical properties and experimental error.

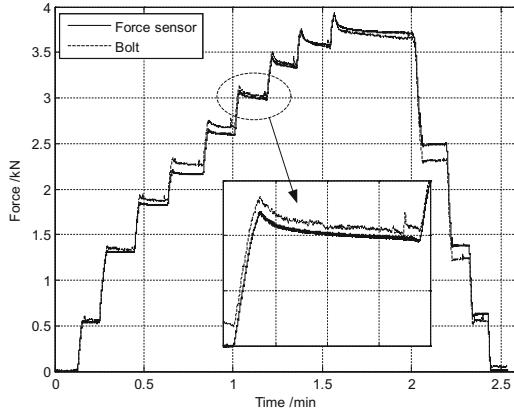


Fig. 4. Load measurement of bolt sensor and force sensor

Under shock load, the force and acceleration respond is achieved in Fig. 5. As expected, good agreement is observed between results of the bolt and force sensor. Moreover, deformation of the beam can be integral calculated from the acceleration, which brings the same trend. The experimental data suggests that the proposed bolt sensor can be used to accurately detect bolt load behavior both in the static and dynamic environments.

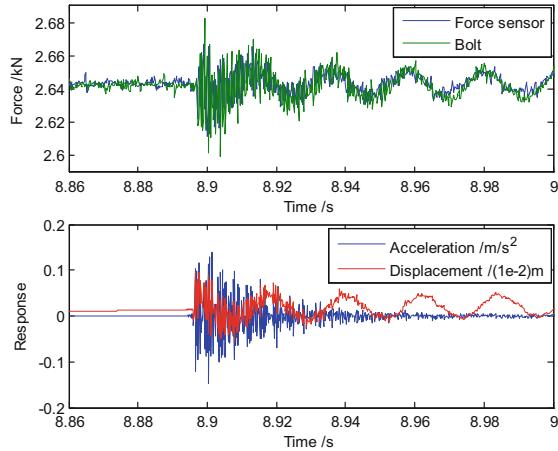


Fig. 5. Load measurement in dynamic environment

4 Conclusions

In this work, a bolt load sensor for load loss detecting in bolted joints is presented. The full bridge strain-gage circuit for measuring the strain state is placed in chamber of the common bolt. The sensitivity of the bolt sensor is first calibrated with a standard force

sensor. According to tests, a good agreement is found between the force sensor and the proposed bolt sensor. The force-sensitive bolt can be used in static and dynamic environments.

Acknowledgments. The authors would like to acknowledge the support of “Fundamental Research Funds for the Central Universities (24820182018RC30)”.

References

1. Zhang Z, Liu M, Liao Y et al (2018) Contact acoustic nonlinearity (CAN)-based continuous monitoring of bolt loosening: hybrid use of high-order harmonics and spectral sidebands. *Mech Syst Signal Process* 103:280–294
2. Hong M, Su Z, Wang Q et al (2014) Modeling nonlinearities of ultrasonic waves for fatigue damage characterization: theory, simulation, and experimental validation. *Ultrasonics* 54:770–778
3. Daouk S, Louf F, Cluzel C et al (2017) Study of the dynamic behavior of a bolted joint under heavy loadings. *J Sound Vib* 392:307–324
4. Jaglinski T, Nimityongskul A, Schmitz R et al (2007) Study of bolt load loss in bolted aluminum joints. *J Eng Mater Technol* 129(1):48–54
5. Caccese V, Mewer R, Vel SS (2004) Detection of bolt load loss in hybrid composite/metal bolted connections. *Eng Struct* 26:895–906
6. Okugawa M (2004) Bolt loosening detection method by using smart washer adopted 4SID. In: AIAA 2004, p. 1981
7. Jaglinski T, Nimityongskul A, Schmitz R et al (2007) Study of bolt load loss in bolted aluminum joints. *J Eng Mater Technol Trans ASME* 129:48–54



Ground Verification of In-Orbit Lunar Relay Communication Satellite

Wenma Jin^(✉)

Beijing Aerospace Control Center, Beijing 100094, China
wenma.jin@sina.com

Abstract. In this paper, we concentrate on the ground verification of in-orbit Queqiao lunar relay communication satellite for the Chang'e-4 mission, which runs in a halo mission orbit around the Earth-Moon L2 point. The lunar relay communication satellite is introduced to overcome the blind zone of control and measurement between ground stations and the spacecraft located at the far side of the moon, for which ground verification is necessary for validating feasibility and performance of the lunar relay communication functionality beforehand. The methods adopted for verifying the in-orbit relay communication satellite are illustrated while the validation procedures are also elaborated. The main verification items include but not limited to calibration of relay communication antenna, the performance evaluation of GNC attitude orientation control, and the functional validation of the relay communication link between the in-orbit lunar relay satellite and the ground simulation equipment for the Chang'e-4 lander and rover. The validation results are presented in terms of performance evaluation and satisfaction for the requirements of design indices.

Keywords: Ground verification · Lunar relay communication satellite · The far side of the moon · Antenna calibration · Relay communication link

1 Introduction

As is known to all, the Chang'e-4 of China achieved soft landing on the far side of the moon successfully early this year, which is for the first time of all human spacecrafts. Subsequently, the lander carried out the in-situ probing while the rover performed cruise exploration. In order to make it, there are many difficulties and obstacles to overcome, the most significant of which is to realize the communication between the spacecraft and the ground [1].

Relay communication satellite, named as Queqiao, is an absolutely important and innovative part of Chang'e-4 mission, which is aimed at overcoming the blind zone of measurement and control between ground stations and the probe located at the far side of the moon. The introduction of lunar relay communication satellite is also the most significant difference between Chang'e-4 and Chang'e-3 mission. It is the world's first communication satellite to adopt a halo mission orbit around Earth-Moon L2 point, which is located at a distance of about 65,000 km from the moon [2]. Repeated orbit maintenance will be performed to keep it in a required halo orbit. Queqiao can provide

relay communication for the lander and the rover operating on the lunar farside to maintain their contacts with Earth.

In order to validate the designed functionality of the in-orbit satellite, the ground verification is generally necessary and in-orbit test platforms are developed [3, 4]. The main purpose for the ground verification is to validate and evaluate several key criterion of control concepts [5]. In particular, ground verification of the in-orbit lunar relay communication is a must before the launch of Chang'e-4 probe so that the relay satellite and the probes on the far side of the moon can establish relay communication link correctly. The main validation items include but not limited to calibration of the relay communication antenna shaped like an umbrella, performance evaluation of the GNC attitude orientation control and validation of the relay communication function between the in-orbit lunar relay satellite and the ground simulation equipment for the Chang'e-4 lander and rover. In-orbit tests were performed and the validation results showed that the performance of the relay communication satellite can meet the mission requirements.

2 Validation Methods

In this section, we introduce the ground verification methods of the in-orbit lunar relay communication satellite. The calibration of the relay antenna shaped like an umbrella is to verify that the antenna orientation could meet the requirement for relay communication. The performance evaluation of the GNC attitude orientation control is to verify that it can point to the earth and moon accurately. Functional validation of the relay communication link is to confirm that the in-orbit satellite could provide relay communication service for the lander and rover in a closed TT&C loop.

2.1 Calibration of the Relay Communication Antenna

The calibration of the relay communication antenna shaped like an umbrella is fundamental to the relay communication. A ground station is required to receive the forward radio frequency signal so that the calibration results can be calculated. The antenna pattern could be numerically measured by the attitude offset of the in-orbit satellite [6, 7]. The calibration methods adopted for the lunar relay satellite are analogical to that used for earth orbit satellite except that the communication distance is much farther [8, 9].

The calibration procedure consists of the following four steps. Firstly, the relay communication antenna should be controlled to point to the given ground station for calibration. The ground station confirms that it can receive the forward radio frequency signal of the relay satellite and then monitor the radio power continuously. Secondly, the orientation calibrations are performed, which consist of cross-shaped scanning with different range of angles and spiral scanning. Thirdly, the orientation calibrations are conducted in several different situations in which the sun, the earth and the moon are in different relative positions separately.

During the scanning process, the selected ground station monitor the power of the forward radio frequency signal continuously. The orientation deviation of the relay

communication antenna could be calculated by comparing with the offset angle in the telemetry data. As is given in the following Table 1, we concentrate on two types of antenna orientation calibration while three different ranges of scanning angles for the cross-shaped scanning are adopted in this paper. Once a round of cross-shaped scanning is completed, the ground calculate and determine to introduce the calibration result into the next round of scanning if or not. Usually, a round of spiral scanning is performed after several rounds of cross-shaped scanning and the calibration procedure concludes.

Table 1. Two types of antenna orientation calibration.

| Calibration types | Ranges of scanning angle | Antenna orientation compensating angle | |
|-----------------------|--------------------------|--|-------------|
| | | Roll angle | Pitch angle |
| Cross-shaped scanning | 1.2° | 0° | -0.6° |
| | | 0° | +0.6° |
| | | -0.6° | 0° |
| | | +0.6° | 0° |
| | 0.8° | 0° | -0.4° |
| | | 0° | +0.4° |
| | | -0.4° | 0° |
| | | +0.4° | 0° |
| | 0.4° | 0° | -0.2° |
| | | 0° | +0.2° |
| | | -0.2° | 0° |
| | | +0.2° | 0° |
| Spiral scanning | - | - | - |

2.2 Performance Evaluation of GNC Attitude Orientation Control

Since the relay communication antenna is directional with a narrow beam, the performance of GNC attitude orientation control should be quantificationally evaluated so that the relay satellite can point to the probe on the far side of the moon accurately. Please refer to [10] for a general review of control strategies for spacecraft attitude and orientation. While the GNC attitude orientation to the earth is mainly used for data transmission and antenna calibration, the GNC attitude orientation to the moon is the normal attitude for in-orbit relay communication. The performance of GNC attitude orientation to the earth could be conducted during the relay antenna calibration process. The GNC attitude orientation precision to the moon should be evaluated separately, in which the orientation to the lunar orbit and the theoretical landing site on the far side of the moon. In addition, the performance of the inertial spatial orientation should also be validated. In each case, the attitude information in the telemetry data is analyzed in detail, from which we can evaluate the attitude orientation stability and precision in three orthogonal axes of the satellite.

2.3 Validation of the Relay Communication Link Function

The most significant procedure of the in-orbit tests is to conduct the functional validation of the relay communication link. As shown in the following Fig. 1, the ground forwarding antenna and simulation equipment for the lander and rover are required in this test, where the ground forwarding antenna and the simulation equipment is connected in a wired way. The relay communication is established between the in-orbit satellite and ground equipment. On one hand, the telemetry data from ground simulation equipment for the lander and rover could be received by the in-orbit relay satellite by which the telemetry data is forwarded to the ground stations and the control center. On the other hand, instructions or injection data for the lander and rover is sent from the control center and ground stations to the in-orbit relay satellite, by which the instructions or injection data is forwarded to the ground simulation equipment for the lander and rover.

In particular, the forward link is defined as the data link from the in-orbit satellite to the ground forwarding antenna while the backward link is defined as the data link from the ground forwarding antenna to the in-orbit satellite. The telemetry from the ground simulation equipment for the lander and rover could be stored and analyzed in the control center. The reception and execution of the instructions or injection data for the lander and rover could also be analyzed in the ground simulation equipment. In this way, the function and performance of the relay communication link could be validated in the ground. However, it should be pointed out that the major distinctions to the real relay communication link between the relay satellite and the probe on the lunar farside are the increase of communication distance and the reduction of signal due to terrestrial atmosphere.

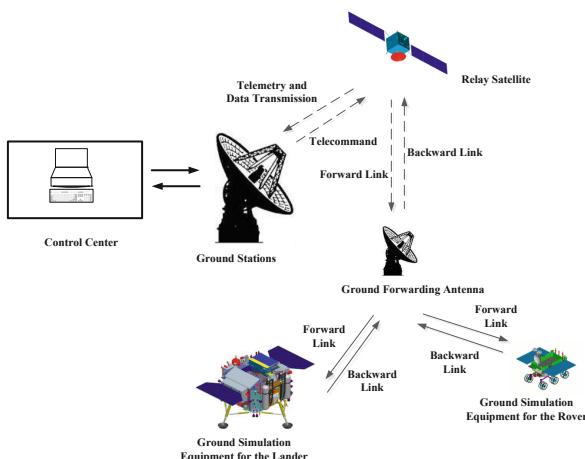


Fig. 1. Illustration for the in-orbit functional validation of the relay communication link

3 Main Results

In this section, we present main results of the in-orbit functional validation of relay communication satellite. The results of the relay communication antenna (RCA) calibration are given in terms of cross-shaped scanning and spiral scanning, respectively. The performance evaluation of the GNC attitude orientation control (GAOC) is given by estimating the intersection angle between +Z-axis and the aimed direction. The functional validation results of the relay communication link (RCL) are presented in several typical working modes.

3.1 RCA Calibration

Due to the limited space, we only report representative calibration results of the relay communication antenna with a specific relative position of the sun, the earth and the moon. Under the ground control, the relay communication antenna shaped like an umbrella points to the ground station with its forward link communication equipment startup, which is realized by injection control data to the satellite.

Calibration results for two kinds of cross-shaped scanning angles, namely $\pm 0.6^\circ$ and 0.4° , are presented in Figs. 2 and 3 respectively. In each figure, the left subfigure gives the error of attitude angle while the right subfigure gives the error of attitude angular velocity. As for the error of the attitude angle and angular velocity, the roll, pitch and yaw component are given simultaneously. From the variation of the attitude angle as shown in Figs. 2 and 3, we can deduce that the antenna acts exactly as the injection data sent from the ground in each case. By statistic analysis of the telemetry data, it can be confirmed that the cross-shaped scanning results satisfy the requirements of design index.

In addition, calibration results for the spiral scanning are also presented, as shown in Fig. 4. It takes about twenty minutes to complete a round of spiral scanning. In the left subfigure of Fig. 4, the spiral scanning angle of X-axis and Y-axis are given at the same time. In the right subfigure of Fig. 4, the spiral scanning angle of Y-axis versus X-axis are given, in which spiral appears. Besides, it is also verified that the spiral scanning results satisfy the requirements of design index.



Fig. 2. The error of attitude angle (the left subfigure) and angular velocity (the right subfigure) for $\pm 0.6^\circ$ cross-shaped scanning

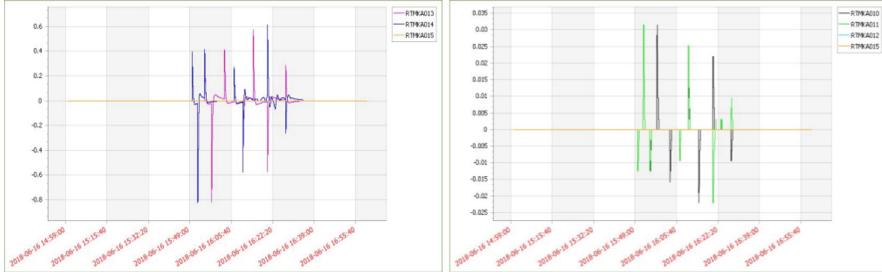


Fig. 3. The error of attitude angle (the left subfigure) and angular velocity (the right subfigure) for $\pm 0.4^\circ$ cross-shaped scanning

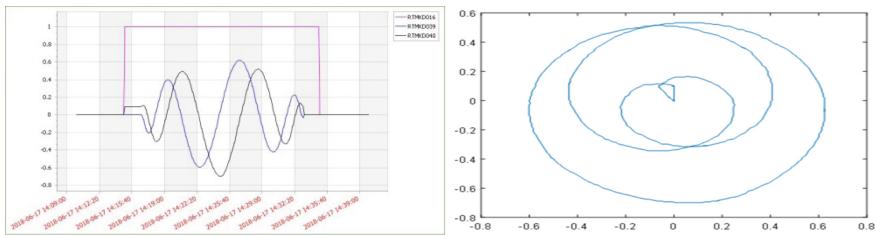


Fig. 4. The variation of attitude angle (the left subfigure) and the trajectory for the spiral scanning (the right subfigure)

3.2 GAOC Estimation

The in-orbit relay satellite could be controlled to direct its +Z-axis to the earth or the moon by injecting orbital data of relay target and setting working mode accordingly. The precision of GNC attitude orientation to both the earth and the moon are numerically evaluated. It should be pointed out that the GAOC precision evaluation process should avoid the momentum wheel unloading. Due to the limited space, the case of inertial spatial orientation is omitted in this paper.

In Table 2, the statistical results of the GAOC evaluation are presented, where the mean value and variance of the intersection angle between +Z-axis and the aimed direction are given at the same time. The mean value of the intersection angle between +Z-axis and the direction pointing to the earth is about 0.006882 while the variance is 6.859688e-6. Besides, the mean value and variance of the intersection angle between +Z-axis and the direction pointing to the moon are 0.007039 and 2.439123e-6, respectively. It should be pointed out that the mean value of the intersection angle between +Z-axis and the direction pointing to the moon is slightly more than the case of pointing to the earth while the variance is slight less. According to the evaluation results in Table 2, it can be deduced that the GAOC precision is better than the designed requirements.

Table 2. The mean value and variance of the intersection angle between +Z-axis and the aimed direction during the given time interval.

| GNC orientation control | Mean Value | Variance |
|--------------------------|------------|-------------|
| Orientation to the earth | 0.006882 | 6.859688e-6 |
| Orientation to the moon | 0.007039 | 2.439123e-6 |

In Fig. 5, the diagram of intersection angle between the +Z-axis and the direction pointing to the earth as well as the direction pointing to the moon is presented, in which the vertical axis denotes the intersection angle while the horizontal axis is time.

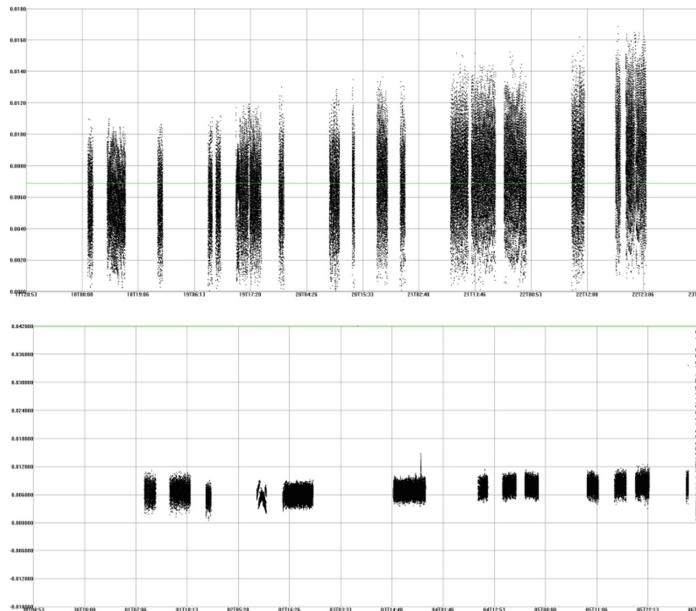


Fig. 5. The intersection angle between the +Z-axis and the direction pointing to the earth (the subfigure above) as well as the direction pointing to the moon (the subfigure below).

3.3 RCL Validation

As mentioned above, validation of the relay communication link (RCL) is fundamental for verifying that the in-orbit relay satellite could supply relay service normally. In general, both the forward link and the backward link between the relay satellite and the ground simulation equipment for the probe should be validated.

As for the validation of the backward link, different types of antenna, frequency and bit rates for the probe should be involved. Five different kinds of working modes for the backward link between the probe and relay satellite are considered. In each

situation, the backward link could be established normally and the ground center could receive data transmission forwarded by the in-orbit relay satellite correctly. In particular, the capability of the two backward link working simultaneously is also validated, in which the backward link between the lander and the relay satellite and the one between the rover and the relay satellite work normally at the same time. In addition, the backward link acquisition time indices are evaluated in each case and the detailed results are omitted here.

The validation of the forward link should cover different TT&C transponders and corresponding frequencies. Two kinds of working modes for the forward link between the probe and relay satellite are considered. In each working mode, the corresponding TT&C transponder of the probe, the forward modulator for the relay satellite and the forward frequency adopted are given. Firstly, we should validate that the relay satellite could forward instructions or injection data to the ground simulation equipment for the lander or the rover normally when only one forward link is established. Secondly, the capability of the two forward link working simultaneously is also validated, in which the forward link between the lander and the lunar relay satellite and the one between the rover and the lunar relay satellite are established at the same time. The validation of the forward link should cover instructions and injection data of different types and code length. In addition, the systemic time delay is also evaluated during validation of the forward link, which is useful for the analysis of the actual relay communication.

4 Discussions and Conclusions

To overcome the communication blind zone on the far side of the moon, the lunar relay communication satellite is introduced into the Chang'e 4 mission. Although it is not the first lunar relay satellite, it is indeed the first lunar relay satellite to adopt a halo mission orbit around the Earth-Moon L2 point, which is key to the achievement of soft landing on the lunar farside for the first time. For the sake of confirming beforehand that the relay communication between the probe and the relay satellite can work well, the ground verification of the in-orbit data relay communication satellite is a must. In this paper, we summarize the major validation contents of the lunar relay satellite and the corresponding functional validation results.

In order to validate that the relay communication antenna can still satisfy the requirements of design index after the launch and longtime flight process from earth to the L2 point, two types of antenna calibration method, namely the cross-shaped scanning and the spiral scanning, are adopted to evaluate the precision of the antenna orientation. Besides, different cross-shaped scanning angles are used for the integrity of the calibration. The reported calibration results validate the functional performance of the relay communication antenna. As for the orientation control capability of the GNC, the precision of the orientation control are obtained in both the earth-directed and the moon-directed case. Moreover, a system for the functional validation of the relay communication link is introduced, in which both the forward link and backward link between the in-orbit relay satellite and the ground simulation equipment for the probe are validated intensively. However, It should be pointed out that there are also some

other in-orbit test items for the lunar relay satellite, which are not all included in this paper due to the space limitations.

In conclusion, based on the first lunar relay communication satellite designed for the Chang'e 4 mission of our country, several important ground verification of in-orbit tests are conducted to validate the feasibility, capability and performance of the lunar relay satellite. Representative functional validation and performance evaluation results are presented. Overall speaking, the ground verification indicate that the actual lunar relay communication function satisfy designed requirements. In addition, some design defects of the spacecraft can be revealed beforehand by proper in-orbit functional validation, which can be vital for the success of the whole mission sometimes.

References

1. Wu W, Wang Q, Tang Y et al (2017) Design of Chang'e-4 lunar farside soft-landing mission. *J Deep Space Explor* 4(2):585–589
2. Gao S, Zhou W, Liang W et al (2017) Trajectory analysis and design for relay satellite at lagrange L2 point of earth-moon system. *J Deep Space Explor* 4(2):122–129
3. Stoll E, Walter U, Artigas J et al (2010) Ground verification of the feasibility of telepresent on-orbit servicing. *J Field Robot* 26(3):287–307
4. Bodin P (2015) PRISMA: an in-orbit test bed for guidance, navigation, and control experiments. *J Spacecr Rocket* 46(3):615–623
5. Caldichoury M (1985) Evaluation of control concepts for a large geostationary data relay satellite. *IFAC Proc Vol* 18(4):33–42
6. Rosell GEG, Teixeira B, Olimpiev A et al (1991) In-orbit RF test of an intelsat VI spacecraft. *COMSAT Tech Rev* 21(2):339–390
7. Xu G, Mao X, He Z et al (2013) Study on in-orbit test methods for antenna coverage of geostationary communication satellites. *J Spacecr TT&C Technol* 32(3):217–223
8. Oswald TH, Macher W, Rucker HO et al (2009) Various methods of calibration of the STEREO/WAVES antennas. *Adv Space Res* 43(3):355–364
9. Sun X, Yang D, Geng Y et al (2004) Antenna pointing control strategy study of tracking and data relay satellite. *Acta Aeronautica Et Astronautica Sinica* 25(4):376–380
10. Hassrizal HB, Rossiter JA (2016) A survey of control strategies for spacecraft attitude and orientation. In: 2016 UKACC 11th international conference on control (CONTROL). IEEE, Belfast, pp 1–6



Trajectory Planning of UAV in Unknown Dynamic Environment with Deep Reinforcement Learning

Jia Wang^(✉), Weihong Wang, and Qian Wu

Beihang University, Beijing 100191, China
beihangwj@buaa.edu.cn

Abstract. Providing a collision-free, safe and efficient optimal trajectory for unmanned aerial vehicles (UAVs) in an unknown dynamic environment is one of the most important issues for researchers. In this paper, a trajectory planning approach for UAV in unknown dynamic environment based on deep reinforcement learning (DRL) is proposed. This study models trajectory planning of UAV as a discrete-time, discrete-action problem, and then proposes an improved deep Q network (IDQN) algorithm to solve it. The IDQN algorithm adds the track angle information of UAV to the reward function to speed up the learning process, furthermore, it also improves the action selection strategy and learning rate setting. Besides, in simulation, the paper considers the trajectory constraints of UAV in order to make the obtained trajectory have better practical availability. Simulation results demonstrate the effectiveness of the IDQN algorithm to implement UAV trajectory planning with constraints in unknown dynamic environments. Meanwhile, comparison with the classical DQN (CDQN) algorithm is conducted to further explore the advantage of the method.

Keywords: UAV · Trajectory planning · Improved DQN · Unknown dynamic environment · Trajectory constraints

1 Introduction

Unmanned Aerial Vehicle (UAV), especially a machine that is not equipped with an operator, but is controlled by ground personnel or programs to achieve autonomous or semi-autonomous flight. Nowadays, with the development of technology, the application of UAVs has expanded from the initial military field to various fields such as agriculture, industry, commerce, and civil use [1].

The autonomous flight of UAVs has become its development trend. Among them, the trajectory planning technology is an important research direction in the autonomous flight of UAV. Many scholars at home and abroad have made in-depth research on A* algorithm, artificial potential field method, genetic algorithm and ant colony algorithm, which has made great achievements in the trajectory planning of UAVs. However, most of above traditional methods in trajectory planning need to be based on known map information, which is insufficient in some unknown environments to cope with changing conditions and unforeseen uncertainty [2]. The development of artificial intelligence (AI) has provided new inspiration for solving this problem.

As an important branch of AI, reinforcement learning (RL) has strong decision-making ability. Through continuous interaction with the environment, the path planning with obstacle avoidance function can be realized. Liang Quan [3] used Q-learning to realize obstacle avoidance of robots in an unknown environment.

With the development of deep neural network, it is possible transfer state value function into neural network by parameters without discretization. DRL combines deep neural networks with reinforcement learning, which enables the agent to learn action strategy from more complex environment and high dimensional observation. The research directions of DRL algorithms widely recognized include deep Q network (DQN) [4, 5] and related improvements, strategy-based DRL algorithms (such as DDPG [6], DPPO [7], etc.) and some other research work (such as A3C [8], UNREAL [9]). Polvara [10] designs two DQNs that control UAV in two delicate phases: landmark detection and vertical descent, applying successfully in autonomous landing problems. Hongji Huang [11] merges the DRL with the UAV navigation issue through massive multiple input multiple output technique, and finally optimizes the UAV navigation by DQN. Ramos [12] designed a versatile Gazebo-based DDPG framework to validate continuous landing task, the method enables the UAV to autonomously landing on a moving platform. Zeng [13] presents a novel two-level hierarchical framework called JPS-IA3C for robot navigation in dynamic environments through continuous controlling signals.

Based on the above content, the study uses the improved DQN algorithm to realize UAV trajectory planning in unknown and dynamic environment. The main contributions of this paper can be summarized in three points:

- (1) The paper systematically proposed a new technical solution to realize UAV trajectory planning: an improved DQN method. And this paper innovatively designed the reward function to better achieve the UAV trajectory planning.
- (2) In simulation, the study builds dynamic complex simulation environments: including static and dynamic obstacles and threats.
- (3) This paper considers the actual trajectory constraints in [14] of the UAV for simulation, making the obtained trajectory have better practical availability.

The rest of the paper is organized as follows: Sect. 2 introduces the related work of trajectory planning. Section 3 introduces the IDQN algorithm and realizes UAV trajectory planning. Simulation settings, simulation results and analysis are shown in Sect. 4. Finally, Sect. 5 gives conclusions.

2 Related Work

2.1 Environment Model

In practice, UAVs are required to fly at a certain altitude due to specific environmental reasons or task requirements, etc. This paper aims to realize the trajectory planning of the UAV in two-dimensional dynamic unknown environments.

The environment model is shown in Fig. 1. This study sets up mobile UAV, obstacles, threat sources and target in the environment. UAV only knows its own

information and target location information in advance, and doesn't know the information of obstacles and threat sources. Every time the UAV only can get information about obstacles and threat sources through sensors within a certain distance.

In Fig. 1, it sets the UAV to target point distance d_{u_t} , UAV to obstacle distance d_{u_o} , UAV to threat source distance d_{u_th} , sets the maximum distance perceived by sensors of UAV $d_sensor = 10$ km, sets the success distance $d_{sa} = 2$ km, the collision distance with the obstacle $d_{ca} = 3$ km, threat range of threat sources $d_{th} = 4.5$ km.

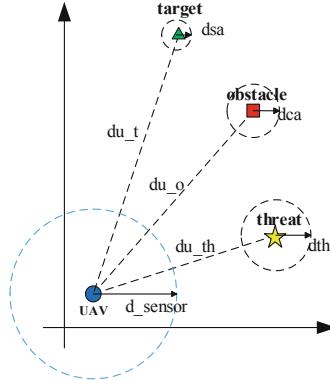


Fig. 1. UAV environment model

2.2 Reward Function

Optimized performance indicators including distance, time, energy consumption. And the paper uses distance as the standard to optimize trajectory. For the obstacle avoidance trajectory planning problem, the reward function includes the negative discrete return function when the collision occurs or enters into the threat zone, the positive discrete return function that successfully reaches the target, and the continuous return function in general. Specifically, the reward function is shown as (1). It is worth noting that the UAV would get -0.01 reward at every time step in order to force it solves the task in minimal number of steps, so that obtain shortest trajectory in distance.

$$r_t = \begin{cases} +10 & d_{u_t}(t) < d_{sa} \\ -10 & (d_{u_o}(t) < d_{ca}) \text{ or } (d_{u_th}(t) < d_{th}) \\ -0.01 - k1 * f_{sh} - k2 * f_{\theta} - k3 * f_{ob} - k4 * f_{th} & \text{otherwise} \end{cases} \quad (1)$$

Where: $k1-k4$ are non-negative weighting coefficients with a sum of 1. In addition, when calculating the reward function, firstly, each evaluation factor f_{sh} , f_{θ} , f_{ob} , f_{th} should be normalized; Their meanings and calculation methods are as follows:

f_{sh} is a reward shaping function designed according to [15] to improve the convergence speed of the algorithm.

$$f_{sh} = d_{u_t}(t) - d_{u_t}(t-1) \quad (2)$$

where: $d_{u_t}(t-1)$ and $d_{u_t}(t)$ respectively represent the Euclidean distance between the current state transition start and stop states and the target point.

f_θ is used to guide the UAV move toward the target's direction by setting the punishment term of the difference between the UAV speed direction angle and the angle target relative to the UAV.

$$f_\theta = |\theta_v(t) - \varphi_{u_t}(t)| \quad (3)$$

where: $\theta_v(t)$ is the direction angle of the UAV speed at time t. $\varphi_{u_t}(t)$ is the direction angle of the target's location relative to the UAV's location.

f_{ob} is the degree of threat to the UAV by the obstacles.

$$f_{ob} = \frac{1}{d_{u_o}(t)} \quad (4)$$

where: $d_{u_o}(t)$ is the distance between the UAV and the nearest obstacle.

f_{th} is the degree of threat to the UAV by the threat sources.

$$f_{th} = \frac{1}{d_{u_th}(t)} \quad (5)$$

where: $d_{u_th}(t)$ is the distance between the UAV and the nearest threat sources.

3 Improved DQN Algorithm

3.1 A New Action Selection Strategy

The performance of DQN algorithm is influenced by the action selection strategy. Designing a reasonable action selection mechanism to control the balance of exploration and exploitation is of great significance, which ensure the neural network quickly converge to the optimal Q-value function. Commonly used action selection strategies in DRL include ϵ -greedy strategy, Boltzmann distribution strategy and so on.

In ϵ -greedy strategy, parameter ϵ is introduced to balance the ratio between randomly and greedy action selection. The ϵ -greedy strategy is shown in (6).

$$a = \begin{cases} \arg \max_{a_i \in A} (Q(s, a_i)) & p < \epsilon \\ \text{rand}(A) & \text{others} \end{cases} \quad (6)$$

Where p is a randomly generated number, and $p \in (0, 1)$.

Compared with ε -greedy strategy, Boltzmann distribution strategy pays more attention to the exploration of suboptimal actions that may become optimal actions. In Boltzmann strategy, the probability p of the agent selecting a_i is shown in (7).

$$p(a_i|s) = \frac{\exp(Q(s, a_i)/T)}{\sum_{a_k \in A} \exp(Q(s, a_k)/T)} \quad (7)$$

Where: T is the temperature parameter, used to control the exploration intensity.

This paper proposes a new action selection strategy which combines the ε -greedy strategy with Boltzmann distribution strategy. The strategy is shown as follows (Table 1):

Table 1. Improved action selection strategy

-
1. Generate p_1 randomly, $p_1 \in (0,1)$;
 2. **If** $p_1 < \varepsilon_1$;
 3. $a = \arg \max_{a_i \in A} (Q(s, a_i))$;
 4. **Else**
 5. Generate p_2 randomly, $p_2 \in (0,1)$;
 6. **If** $p_2 < \varepsilon_2$;
 7. $a = \text{Boltzmann}(A)$;
 8. **Else**
 9. $a = \text{rand}(A)$;
 10. **End if**
 - 11.**End if**
-

In the new action selection strategy, at the beginning of the training, the parameters ε_1 and ε_2 are set to small value, and then as the number of iteration steps increases, ε_1 and ε_2 gradually increase to the maximum value. By adjusting the values of ε_1 , ε_2 , $\Delta\varepsilon_1$ and $\Delta\varepsilon_2$ realizing the balance of exploration and exploitation better.

3.2 Learning Rate Setting Improvement

Learning rate is a key hyperparameter in the optimization process of neural networks, which indicates the magnitude of each network parameter change. If the learning rate is too large, the parameter changes greatly, which may cause it to move around the optimal solution. However, too low learning rate leads to low training speed. In order to solve the problem, the paper uses the exponential decay method as shown in (8) to adjust the learning rate. In this way, during training, the parameters can be quickly optimized to the vicinity of the optimal solution, and after that, the variation of parameters is smaller and slowly approaches the optimal solution.

$$l = L * d^{\frac{s_g}{s_l}} \quad (8)$$

Where l is the learning rate, L is the initial learning rate, d is the scaling coefficient for each decay, s_g is the global number of steps, and s_l is the decay rate.

3.3 Improved DQN Algorithm

The above new action selection strategy and learning rate setting improvement method are applied to the classical DQN (CDQN) algorithm to obtain the improved DQN (IDQN) algorithm. The IDQN algorithm is as follows (Table 2):

Table 2. IDQN algorithm

Algorithm1. IDQN algorithm for UAV trajectory planning

1. **Initialization:** the capacity of the target memory unit D is N ; the initial weight of the Q network is θ ; the initial weight of the target Q network is $\theta^- = \theta$; max episode number is MAX_EPISODES; max step of every episode is MAX_EP_STEPS.
 2. **For** episode in range (MAX_EPISODES):
 3. Initial state sequence is s_1 , step = 0
 4. **For** step in range (MAX_EP_STEPS):
 5. Use the new action selection strategy to select action a_t
 6. Perform action a_t , and get the reward r_t and the next state s_{t+1} from environment.
 7. Store transfer sequence (s_t, a_t, r_t, s_{t+1}) into memory unit D
 8. **If** the memory unit D is full:
 9. Randomly sample a fixed number of transfer samples from D. Start training.
 10. **If** reach the termination:
 11. set the optimization goal $y_j = r_j$
 12. **Else:**
 13. $y_j = r_j + \gamma \max_a Q(s_{j+1}, a, \theta^-)$
 14. Update Q network parameter θ using loss function $L = E[(y_j - Q(s_j, a_j, \theta))^2]$
 15. Reset parameters of target Q network $\theta^- = \theta$ every C time steps
 16. $s_t = s_{t+1}$
 17. **If** reach the termination:
 18. **break**
 19. **End for**
 20. **If** episode reaches a certain set value:
 21. Reset the learning rate using exponential decay method
 22. **End for**
-

3.4 UAV Trajectory Planning Implementation

Interact the IDQN above with the constructed environment model according to Fig. 2.

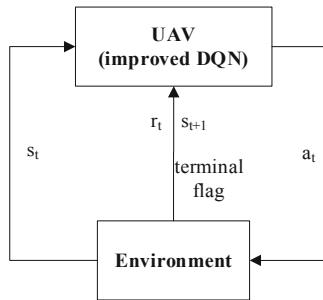


Fig. 2. Interaction

The IDQN algorithm is used to explore, train and exploit in the UAV's trajectory planning process, and finally the deep neural network converges to the optimal Q-value function. The final well-trained network model and parameters are saved for testing. Finally, the test algorithm for the trajectory planning is as follows (Table 3):

Table 3. Testing algorithm

Algorithm2. Testing method based on well-trained IDQN for UAV trajectory planning

Input: Environment model, well-trained IDQN.

1. Load the well-trained IDQN framework and parameters
 2. Receive initial state sequence s_1 from the environment model; step = 0
 3. **For** step in range (MAX_EP_STEPS):
 4. Select action $a_t = \arg \max_{a_i \in A} (Q(s_t, a_i, \theta))$
 5. Perform action a_t in the environment model, and get reward r_t , the next state s_{t+1} , and the terminal flag.
 6. Save the location coordinates of the UAV
 7. **If** the terminal flag is True:
 8. **break**
 9. $s_t = s_{t+1}$
 10. **End for**
-

4 Simulation

The trajectory constraints derived from the actual UAV engine parameters are given in [14], among them, the given constraints of the UAV1 are considered in the simulation (Table 4).

Table 4. UAV trajectory constraints

| Maximum turning angle (β_{max}) | Air-range (R) | Duration of flight (t_d) | Minimum track length (l_{min}) |
|---|-------------------|------------------------------|------------------------------------|
| 47.34° | 285 km | 2.5 h | 1.2 km |

According to the constraints above, the UAV trajectory planning simulation settings are as follows. The flight area is set to 80 km × 80 km. The UAV starting point is (0,0) and the target point is (70, 70). Considering β_{max} is 47.34°, the UAV action is set to {0, -30°, +30°}. Considering R and t_d , the speed of the UAV is set to 0.03 km/s. Considering l_{min} , the study sets the length of the UAV to be 1.5 km per step, so the time step is set to 50 s. In addition, the study sets 6 obstacles and 3 threat sources.

The simulation code is written in python, running 600 episodes in total, and the maximum number of steps per episode is 200. The neural network is constructed using a three-layer fully connected layer, which has 50, 50, and 3 neurons respectively.

4.1 The Training Process

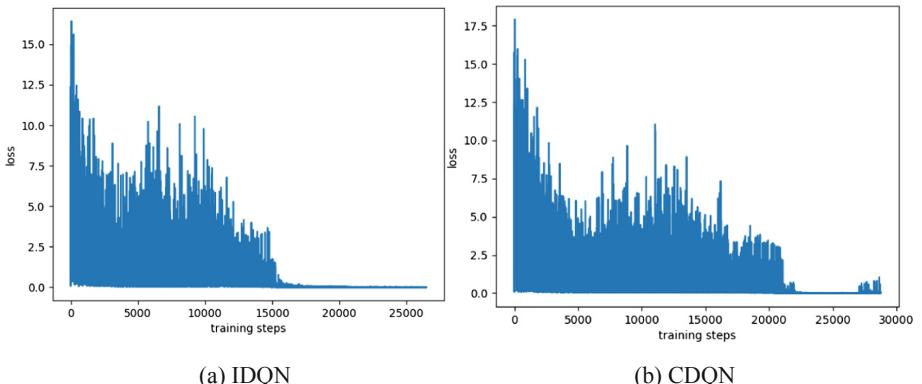


Fig. 3. Loss function curve during training with IDQN and CDQN

Figure 3(a) shows that with the number of training steps increases, the loss function exhibits a decaying trend, eventually decaying to zero, and so the IDQN algorithm converges. This proves that the IDQN algorithm is effective in training process.

In addition, comparing Fig. 3(a) and (b), it can be seen that the loss function of the IDQN is attenuated to near 0 faster, about 6000 steps ahead of the CDQN. Therefore, the IDQN algorithm improves the convergence speed of training process.

4.2 UAV Optimal Trajectory Obtained by Training

One of the UAV shortest trajectories obtained by training the IDQN and the CDQN are shown in Fig. 4(a) and (b). Among them, green represents the target, blue represents obstacles, and yellow represents threats, the red curve is the UAV's path.

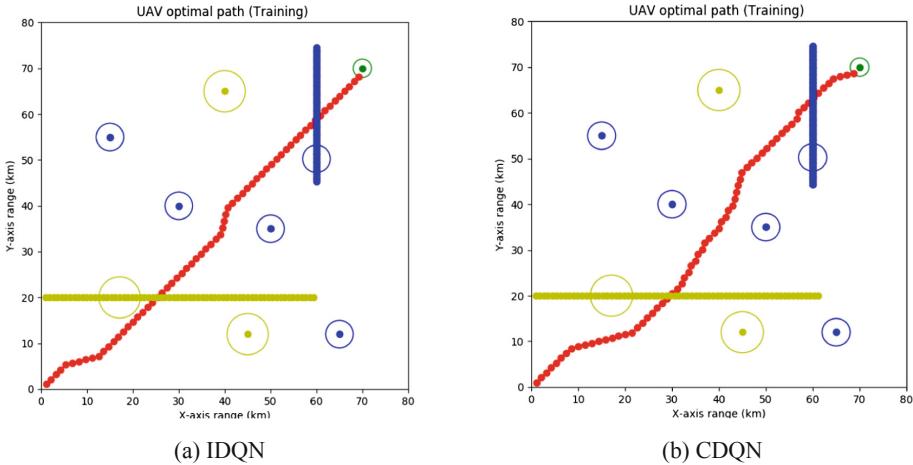


Fig. 4. One of the UAV optimal trajectories obtained by training IDQN and CDQN

Figure 4 show that the both algorithms achieve UAV from the starting point without collision and no threat to reach the target point. In Fig. 4(a), the training shortest trajectories using the IDQN is 66 steps, and the CDQN is 68 steps in Fig. 4(b). Therefore, the optimal trajectory obtained by training with IDQN algorithm is shorter.

4.3 UAV Trajectory Obtained by Testing

Figure 5 shows that both algorithms can achieve UAV trajectory planning finally by testing. In Fig. 5(a), the trajectory obtained by testing the IDQN is 66 steps. And the CDQN is 68 steps in Fig. 5(b). Therefore, the final test results indicate the trajectory planned by the IDQN is shorter, and so the IDQN optimizes trajectory in distance.

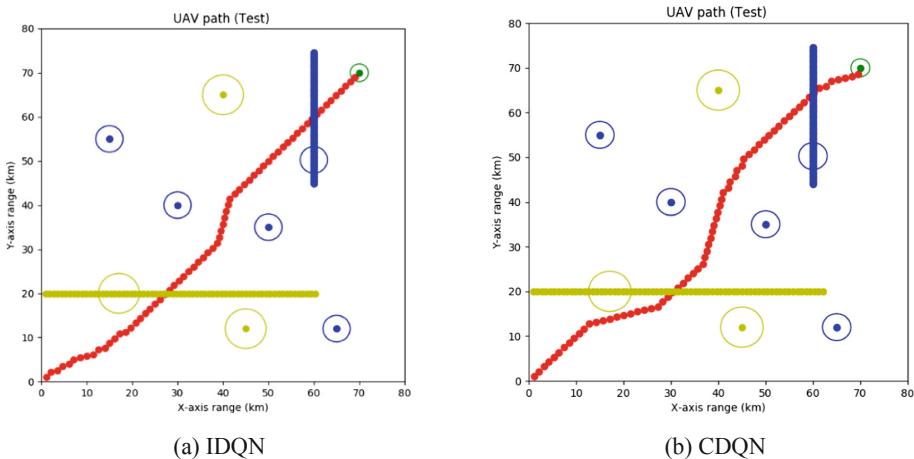


Fig. 5. The UAV trajectory by testing well-trained IDQN and CDQN

5 Conclusions

In the paper, the trajectory planning of UAV with trajectory constraints in unknown dynamic environment is considered, and deep reinforcement learning method is employed to solve the problem. Exactly, the IDQN algorithm is proposed to solve it.

Through performing Python programming simulation on the IDQN algorithm and comparing the simulation results with the CDQN algorithm, finally concluding that the IDQN algorithm can effectively achieve the UAV's trajectory planning with trajectory constraints in unknown dynamic environment, and the IDQN algorithm has a faster convergence speed and shorter trajectory in training than the CDQN algorithm. The final test results also show that the IDQN algorithm get a better performance than the CDQN algorithm, so the IDQN algorithm can plan the UAV trajectory better.

References

1. Xu Z, Wei R, Zhang Q et al (2016) Obstacle avoidance algorithm for UAVs in unknown environment based on distributional perception and decision making. In: IEEE Chinese Guidance, Navigation and Control Conference (CGNCC). IEEE, pp 1072–1075
2. Zhao Y, Zheng Z, Liu Y (2018) Survey on computational-intelligence-based UAV path planning. Knowl-Based Syst 158:54–64
3. Quan L (2012) Mobile robot path planning based on reinforcement learning in an unknown environment. Diss
4. Mnih V, Kavukcuoglu K, Silver D et al (2013) Playing atari with deep reinforcement learning. arXiv preprint [arXiv:1312.5602](https://arxiv.org/abs/1312.5602)
5. Mnih V, Kavukcuoglu K, Silver D et al (2015) Human-level control through deep reinforcement learning. Nature 518(7540):529
6. Lillicrap TP, Hunt JJ, Pritzel A et al (2015) Continuous control with deep reinforcement learning. arXiv preprint [arXiv:1509.02971](https://arxiv.org/abs/1509.02971)

7. Heess N, Sriram S, Lemmon J et al (2017) Emergence of locomotion behaviours in rich environments. arXiv preprint [arXiv:1707.02286](https://arxiv.org/abs/1707.02286)
8. Mnih V, Badia AP, Mirza M et al (2016) Asynchronous methods for deep reinforcement learning. In: International conference on machine learning, pp 1928–1937
9. Jaderberg M, Mnih V, Czarnecki WM et al (2016) Reinforcement learning with unsupervised auxiliary tasks. arXiv preprint [arXiv:1611.05397](https://arxiv.org/abs/1611.05397)
10. Polvara R, Patacchiola M, Sharma S et al (2017) Autonomous quadrotor landing using deep reinforcement learning. arXiv preprint [arXiv:1709.03339](https://arxiv.org/abs/1709.03339)
11. Huang H, Yang Y, Gui G et al (2019) Deep reinforcement learning for UAV Navigation through Massive MIMO. arXiv preprint [arXiv:1901.10832](https://arxiv.org/abs/1901.10832)
12. Rodriguez-Ramos A, Sampedro C, Bavle H et al (2019) A deep reinforcement learning strategy for UAV autonomous landing on a moving platform. J Intell Robot Syst 93(1–2):351–366
13. Zeng J, Qin L, Hu Y et al (2019) Integrating a path planner and an adaptive motion controller for navigation in dynamic environments. Appl Sci 9(7):1384
14. Wang X, Jiang Z, Xie R, Chen H (2012) Engine constraints in UAV route planning (Doctoral Dissertation)
15. Ng AY, Jordan MI (2003) Shaping and policy search in reinforcement learning. University of California, Berkeley



Search-Based Motion Planning for Aggressive MAV with Limited Sensors FOV

Yu Zheng^{1(✉)}, Bowen Li², Zihe Zhao¹, Xu Zeng³,
and Xuedong Zhao³

¹ Beihang University, Beijing 100191, China
yzheng@buaa.edu.cn

² Stevens Institute of Technology, Hoboken 07030, USA

³ Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Abstract. This paper presents a search-based motion planning method to generate resolution-complete, collision-free, dynamically feasible trajectories for Micro Aerial Vehicles (MAVs) aggressive flight in an obstacle-cluttered environment. This method searches for smooth trajectories by exploring the map using a set of short-duration motion primitives. The primitives are generated by solving an optimal control problem and induce a finite lattice discretization on the state space which can be explored using a graph-search algorithm. Finally, we extend the search-based planning approach for Aggressive MAV with constraints on Field-of-View (FOV) while navigating in an unknown, obstacle-cluttered environment, and the MAV uses its onboard depth sensor (RGB-D or event camera) with FOV to detect obstacles, and perform real-time trajectory re-planning. Simulation results demonstrate the feasibility of the generated trajectories in an obstacle-cluttered environment.

Keywords: MAVs · Motion planning · Unknown environment · Limited FOV

1 Introduction

Micro Aerial Vehicles (MAVs) are small multi-rotor helicopters that are able to freely aggressive fly in an unknown, obstacle-cluttered environment. It has been shown in [1] that the MAV dynamics are differential flat which implies that the control inputs can be computed as function of the flat outputs and their derivatives. Many works [1–4] show precise control of MAVs through trajectories represented as time-parameterized piecewise polynomials with respect to these flat outputs. However, in order to consider that they have to be dynamically feasible, collision-free, resolution-complete and optimal, generating those trajectories is a challenging task. Moreover, navigation of the MAV in unknown, obstacle-cluttered environments requires fast re-planning for avoiding new obstacles. Some recent works [5–7] try to apply optimization algorithms that satisfy these complex requirements, however, these approaches are not complete and require heuristic based initial guess to set up the optimization problem. In addition, in more complex situation where limited sensors field-of-view (FOV) such as range sensor (RGB-D or event camera), it is much harder to find the optimal result using the optimization-based approaches.

Our contribution in this work is to provide solution to the limited field-of-view problem using the search-based planning framework. Vision-based state estimation and the limited FOV of sensors to detect obstacles require the MAV to travel with constraints on the yaw angle. We propose a way to find the desired yaw profile along the trajectory that obeys this constraint.

2 General Optimal Control Problem Formulation

Similar to [1], we selected flat output of the MAV as $\sigma = [xyz\varphi]$. Let $x(t) \in \chi \subset \mathbb{R}^{3n}$ is the state of the differential flat system, consisting of the three-dimensional position and its derivatives (speed, acceleration, jerk, etc.). Let $\chi^{free} \subset \chi$ denote free region of state space, P^{free} denote the obstacle-free positions, D^{free} denote the system's dynamics (maximum velocity v_{max} , maximum acceleration a_{max} , maximum jerk j_{max} , etc.), and χ^{free} is used to capture P^{free} and D^{free} . Thus, $\chi^{free} := P^{free} \times D^{free} = P^{free} \times [-v_{max}, v_{max}]^3 \times [-a_{max}, a_{max}]^3 \times \dots$, and $\chi^{obs} := \chi \setminus \chi^{free}$ denote the obstacle region.

The polynomial state trajectories are denoted as

$$x(t) := \left[p_D(t)^T, \dot{p}_D(t)^T, \dots, p_D^{(n-1)}(t)^T \right]^T \quad (1)$$

where

$$p_D(t) := \sum_{k=0}^K d_k \frac{t^k}{k!} = d_K \frac{t^K}{K!} + \dots + d_1 t + d_0 \in \mathbb{R}^3 \quad (2)$$

and $D := [d_0, \dots, d_K] \in \mathbb{R}^{3 \times (K+1)}$. To simplify the notation for the following representations, let $v(t) := \dot{p}_D(t)^T$, $a(t) := \ddot{p}_D(t)^T$, $j(t) := \dddot{p}_D(t)^T$. The polynomial trajectory in Eq. (2) can be generated by considering a linear time-invariant dynamic system $p_D^{(n)}(t) = u(t)$.

The expression of the dynamical system is obtained:

$$\dot{x}(t) = Ax(t) + Bu(t) \quad A = \begin{bmatrix} 0 & I_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & I_3 \\ 0 & \cdots & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I_3 \end{bmatrix} \quad (3)$$

and the control input is $u(t) \in U := [-u_{max}, u_{max}]^3 \in \mathbb{R}^3$. The smoothness or effort of a trajectory is defined as a square L_2 norm of the control input:

$$J(D) := \int_0^T \|u(t)\|^2 dt = \int_0^T \|p_D^{(n)}(t)\|^2 dt \quad (4)$$

Problem 1: Given an initial state $x_0 \in \chi^{free}$ and a goal region $\chi^{goal} \subset \chi^{free}$, find a polynomial trajectory $D \in \mathbb{R}^{3 \times (K+1)}$ and a time $T \geq 0$ such that:

$$\begin{aligned} & \min_{D,T} J(D) + \rho T \\ & \text{s.t. } \dot{x}(t) = Ax(t) + Bu(t), \forall t \in [0, T] \\ & x(0) = x_0, x(T) \in \chi^{goal} \\ & x(t) \in \chi^{free}, u(t) \in U, \forall t \in [0, T] \end{aligned} \quad (5)$$

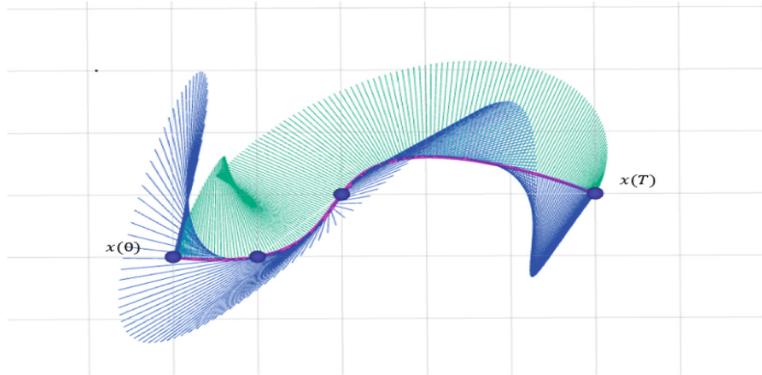


Fig. 1. The jerk trajectory start from $x(0)$ to $x(T)$

The optimal cost from an initial state x_0 to a goal region χ^{goal} is denoted as $C^*(x_0, \chi^{goal})$, such as Fig. 1, the optimal cost of the jerk trajectory is 7.86, and blue and green rays indicate the magnitude of velocity and acceleration along trajectories, respectively. This problem is a Linear Quadratic Minimum-Time problem with state constraints $x(t) \in \chi^{free}$ and input constraints $u(t) \in U$. In this paper, these safety constraints $x(t) \in \chi^{free}$ and $u(t) \in U$ can be handled by converting the problem to a deterministic shortest path problem with a 3 dimensional input state U and $3n$ dimensional state space χ .

3 Aggressive Optimal Trajectory Planning

3.1 Motion Primitive

In this section, we show that the problem 1 can be transformed from the optimal control problem to the graph search problem by constructing the motion primitives of the dynamical system in Eq. (3). And we consider a lattice discretization $U_M := \{u_1, \dots, u_M\} \subset U$, where each control vector $u_m \subset \mathbb{R}^3$ will define a motion of short duration for the system. Given an initial state $x_0 := [p_0^T, v_0^T, a_0^T, \dots]^T$, we generate a motion primitive of duration $\tau > 0$ and apply a constant control input $u(t) \equiv u_m \in U_M$ for $t \in [0, \tau]$ so that,

$$u(t) = p_D^{(n)}(t) = \sum_{k=0}^{K-n} d_{k+n} \frac{t^k}{k!} \equiv u_m \quad (6)$$

Because control input is constant, it applies that all coefficients that involve time need to zero, $d_{(n+1):K} = 0 \Rightarrow u_m = d_n$. And integrating the control input $u(t) = u_m$ with an initial condition x_0 results in

$$p_D(t) = u_m \frac{t^n}{n!} + \dots + a_0 \frac{t^2}{2} + v_0 t + p_0 \quad (7)$$

So the resulting trajectory of the linear time-invariant system in Eq. (3) is:

$$x(t) = \underbrace{e^{At}}_{F(t)} x_0 + \underbrace{\left[\int_0^t e^{A(t-\sigma)} B d\sigma \right]}_{G(t)} u_m \quad (8)$$

Because of the duration τ and control input are fixed, the cost of motion primitive according to Eq. (5) is $(\|u_m\|^2 + \rho)\tau$.

3.2 Induced Space Discretization

Define graph $G(S, \xi)$, where S is the discrete set of reachable system states and ξ is the set of edges that connect states in the graph, each defined by a motion primitive $e := (u_m, \tau)$. Let s_0 be the state corresponding to x_0 , $s_0 := [p_0^T, v_0^T, a_0^T, \dots]^T$. We use Algorithm 1 to explore free state space χ^{free} and build the connected graph, and the motion primitive $u_{ij} \in U_M$ which connects two consecutive states $s_i, s_j \in S$ with $s_j = F(\tau)s_i + G(\tau)u_{ij}$ is optimal according to the cost function in Eq. (5).

Algorithm 1. Given $s \in S$ and a motion primitive set U_M with duration τ , find the states $R(s)$ that are reachable from s in one step and their associated costs $C(s)$

```

1 : function GETSUCCESSORS( $s, U_M, \tau$ )
2 :    $R(s) \leftarrow \emptyset, C(s) \leftarrow \emptyset$ 
3 :   for all  $u_m \in U_M$  do
4 :      $e_m(t) \leftarrow F(t)s + G(t)u_m, \quad t \in [0, \tau]$ 
5 :     if  $e_m(t) \subset \chi^{free}$  then
6 :        $s_m \leftarrow e_m(\tau)$ 
7 :        $R(s) \leftarrow R(s) \cup \{s_m\}$ 
8 :        $C(s) \leftarrow C(s) \cup \{(\|u_m\|^2 + \rho)\tau\}$ 
9 :   return  $R(s), C(s)$ 

```

3.3 Deterministic Shortest Trajectory

Problem 1 can be re-formulated as a graph-search problem by constructing the set of motion primitive and inducted space discretization. This can be done by introduced additional constraints that the control input in Eq. (5) is piecewise-constant over intervals of duration τ . By introducing additional variable $N \in \mathbb{Z}^+$, such that $T = N\tau$, and $u_k \in U_M$ for $k = 0, \dots, N - 1$ and a constraint in Eq. (5):

$$u(t) = \sum_{k=0}^{N-1} u_k \text{sign}_{\{\tilde{t} \in [k\tau, (k+1)\tau]\}} \quad (9)$$

The forced control trajectory becomes a combination of motion primitives in U_M , which generates the following deterministic shortest path problem.

Problem 2: Given an initial state $x_0 \in \chi^{free}$ and a goal region $\chi^{goal} \subset \chi^{free}$, and a finite set of motion primitive U_M with duration $\tau > 0$, choose a sequence of motion primitives $u_{0:N-1}$ of length N such that:

$$\begin{aligned} & \min_{N, u_{0:N-1}} \left(\sum_{k=0}^{N-1} \|u_k\|^2 + \rho N \right) \tau \\ & \text{s.t. } x_k(\tilde{t}) = F(\tilde{t})s_k + G(\tilde{t})u_k \subset \chi^{free} \\ & x_k(\tilde{t}) \subset \chi^{free} \quad \forall k \in \{0, \dots, N-1\}, \tilde{t} \in [0, \tau] \\ & s_{k+1} = x_k(\tau) \quad \forall k \in \{0, \dots, N-1\} \\ & s_0 = x_0, \quad s_N \in \chi^{goal} \\ & u_k \in U_M, \quad \forall k \in \{0, \dots, N-1\} \end{aligned} \quad (10)$$

The optimal cost of Problem 2 is the upper limit of the optimal cost of Problem 1, because Problem 2 is only the constraint of Problem 1. However, this is an efficient solution by re-formulation to discrete control and state spaces. And combining with collision detection, the safety, dynamic feasible trajectory can be obtained.

3.4 Collision Checking

For an edge $e(t) = [p(t)^T, v(t)^T, a(t)^T, \dots]^T$, we need to check if $e(t) \subset \chi^{free}$ for $t \in [0, \tau]$. We check collisions in the geometric space $P^{free} \subset \mathbb{R}^3$ separately from enforcing the dynamic constraints $D^{free} \subset \mathbb{R}^{3(n-1)}$. The edge is valid only if its geometric shape $p(t) \subset P^{free}$ and derivatives $(v(t), a(t), \dots) \subset D^{free}$, i.e.

$$\begin{aligned} & \|v\|_\infty \leq v_{\max}, \quad \forall t \in [0, \tau] \\ & (v, a, \dots) \subset D^{free} \Leftrightarrow \|a\|_\infty \leq a_{\max}, \quad \forall t \in [0, \tau] \\ & \vdots \end{aligned} \quad (11)$$

Since the derivatives v, a are polynomials, their extremes over time $[0, \tau]$ are calculated and compared to the maximum limits of velocity, acceleration, and so on.

The more challenging part is to detect collisions in the geometric space P^{free} . In this paper, the model P is indicated as a grid map. Let $P := \{p(t_i) | t_i \in [0, \tau], i = 0, \dots, I\}$ be a set of positions that dynamical system traverses along the trajectory $p(t)$. Given a polynomial $p(t), t \in [0, \tau]$, the positions $p(t_i)$ are sampled by defining:

$$t_i := \frac{i}{I} \tau, \text{ such that } \frac{\tau}{I} v_{\max} \geq R \quad (12)$$

Here R is the occupied grid resolution. This condition ensures that the maximum distance between two consecutive samples does not exceed the map resolution.

3.5 Planning with Limited Sensors FOV

We solve the planning problem with FOV constraints by adding two constraints: a soft constraint that minimizes the difference between φ and ξ ; and a hard constraint that enforces the moving direction ξ to be inside the FOV of the range sensor.

A soft constrained additional cost term representing the FOV is defined as the integral of the square of the angular difference between the velocity direction and the desired yaw:

$$J_\varphi(\Phi) = \int_0^T [\varphi(t) - \xi(t)]^2 dt \quad (13)$$

While the hard constraints can be expressed by the absolute angular difference and the horizontal FOV of the sensor:

$$|\varphi(t) - \xi(t)| < \frac{\theta}{2} \quad (14)$$

Problem 3: Given an initial state $x_0 \in \chi^{free}$, a goal region $\chi^{goal} \subset \chi^{free}$, and a sensor FOV θ , find a polynomial trajectory Φ such that:

$$\begin{aligned} & \arg \min_{\Phi} J_q + \rho_T T + \rho_\varphi J_\varphi \\ & x(0) = x_0, x(T) \in \chi^{goal} \\ & x(t) \in \chi^{free}, u(t) \in U \\ & |\varphi(t) - \xi(t)| < \frac{\theta}{2} \end{aligned} \quad (15)$$

where the weights $\rho_T, \rho_\varphi \geq 0$ determine the relative importance of the trajectory duration T , the yaw cost J_φ and its smoothness J_q .

4 Analysis of Simulation Results

We simulate minimum snap trajectories with *FOV* constraints (Figs. 2, 3 and 4) and without *FOV* constraints (Figs. 5, 6 and 7) from start s to goal g in 2D and 3D scenes, respectively. The blue (green, red) lines shows the velocity (acceleration, jerk) along trajectory respectively. The purple trajectories are the resulting minimum snap feasible trajectories. And the small triangle is the desired yaw at the corresponding position. Environmental information is initially unknown, and MAVs can use depth sensors with *FOV* to detect obstacles and the surrounding environment.

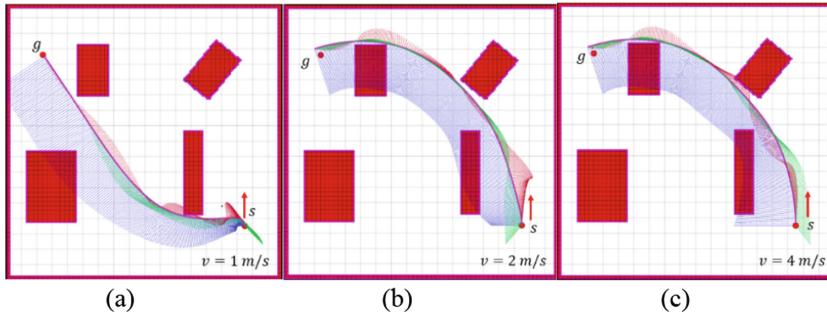


Fig. 2. The minimum snap trajectories from start s to goal g with initial velocity.

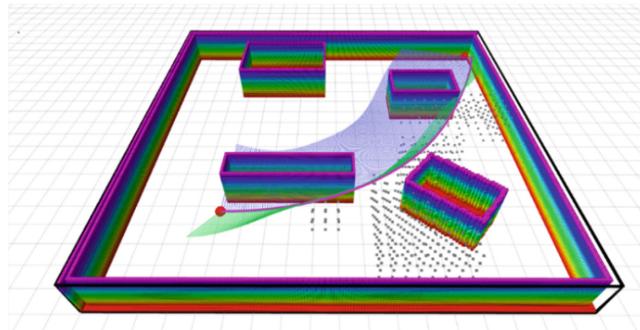


Fig. 3. Minimum snap trajectory without limited *FOV* θ in 3D complex scene

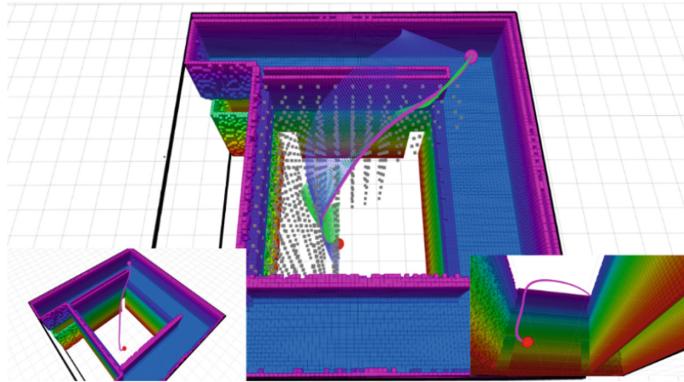


Fig. 4. Minimum snap trajectory without limited $FOV\theta$ in 3D indoor multi-layer scene

The simulations show that the optimal trajectory can be planned in a reasonable time, especially for the dynamic feasible trajectory of the differential flat system, which can make the MAVs reach the target region quickly and safely. In the case of obstacles, since the trajectory of the execution is constrained within sensor FOV , real-time local re-planning can be performed to ensure flight safety.

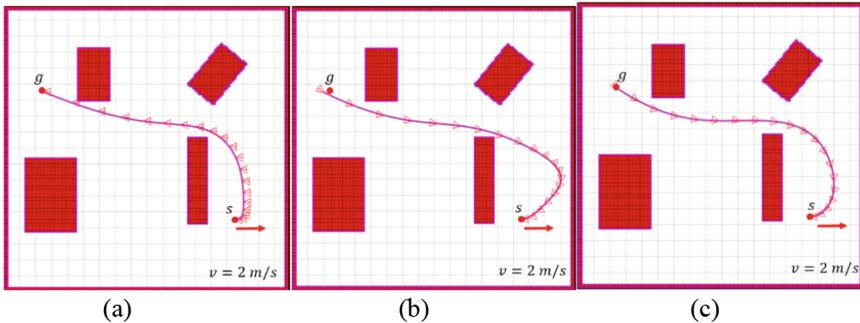


Fig. 5. Minimum snap trajectory with limited $FOV\theta$ with initial velocity

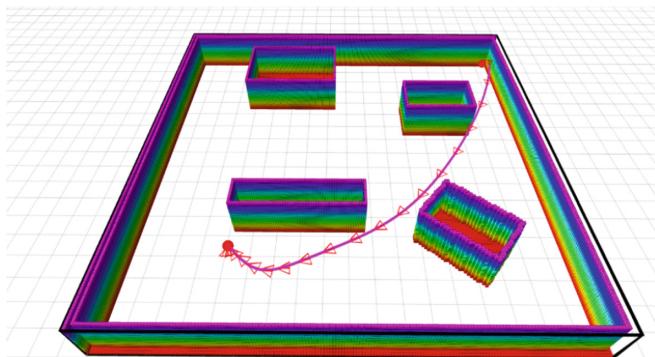


Fig. 6. Minimum snap trajectory with limited $FOV\theta$ in 3D complex scenes

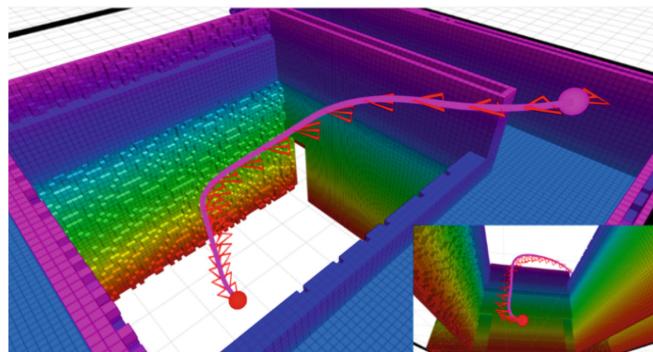


Fig. 7. Minimum snap trajectory with limited $FOV\theta$ in 3D indoor multi-layer scene

5 Conclusion

In conclusion, we proposed a way to use a search-based motion planning method with motion primitives to solve planning problem with FOV constraints in 2D and 3D scenes. The simulation shows that the proposed planner is efficient and resolution complete. We believe the basic approach proposed in this paper is variable for planning optimal trajectories for any system that is differential flat. Moreover, we also believe the proposed methodology has the potential to be used in a board class of MAV navigation problems such as uncertainty environments and dynamically changing environments.

References

1. Mellinger D, Kumar V (2011) Minimum snap trajectory generation and control for quadrotors. In: Proceedings of the 2011 IEEE international conference on robotics and automation (ICRA)
2. Hehn M, D'Andrea R (2011) Quadrocopter trajectory generation and control. IFAC Proc 44(1):1485–1491
3. Mueller M, Hehn M, D'Andrea R (2015) A computationally efficient motion primitive for quadrocopter trajectory generation. IEEE Trans Robot (T-RO) 31(6):1294–1310
4. Richter C, Bry A, Roy N (2016) Polynomial trajectory planning for aggressive quadrotor flight in dense indoor environments. In: Robotics research. Springer, Heidelberg, pp 649–666
5. Oleynikova H, Burri M, Taylor Z, Nieto J, Siegwart R, Galceran E (2016) Continuous-time trajectory optimization for online uav replanning. In: 2016 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, pp 5332–5339
6. Gao F, Lin Y, Shen S (2017) Gradient-based online safe trajectory generation for quadrotor flight in complex environments. In: 2017 IEEE/RSJ international conference on intelligent robots and systems (IROS), September 2017, pp 3681–3688
7. Liu S, Watterson M, Mohta K, Sun K, Bhattacharya S, Taylor CJ, Kumar V (2017) Planning dynamically feasible trajectories for quadrotors using safe flight corridors in 3-D complex environments. IEEE Robot Autom Lett 2(3):1688–1695



An Improvement of Gravitational Search Algorithm

Chaochang Zheng, Hua Li, and Lin Wang^(✉)

School of Electrical Engineering, Inner Mongolia University of Technology,
Huhhot 010080, China
lwang@imut.edu.cn

Abstract. The gravitational search algorithm (GSA) has the advantages of strong exploitation performance and fast convergence speed. However, the GSA is easy to appear premature phenomenon and get into the local optimum during the search process because the particle diversity declines during the optimization process and the particle swarm optimization information is not shared. Therefore, an improved GSA (IGSA) is proposed, which keeps the particle diversity by adjusting the gravitational constant and enhances particle swarm information sharing ability by imposing global optimal information into the search position of each particle. The proposed IGSA has been evaluated on 9 nonlinear benchmark functions and compared with standard GSA and particle swarm optimization (PSO). The obtained results confirm that the convergence accuracy of IGSA is several orders of magnitude higher than that of GSA, and the search speed is also increased by more than 2 times. In addition, a case study of optimizing generator operating costs of IEEE9 is carried, and the IGSA algorithm achieves lower operating costs and less network loss than that of GSA.

Keywords: Improve-GSA · Gravitational constant · Fitness value · Optimal power flow

1 Introduction

Normally, it is difficult to solve the problem of strong nonlinear and high dimensional optimization in the traditional methods. Therefore, heuristic algorithm has been developed rapidly, it includes particle swarm optimization (PSO), simulated annealing (SA), artificial bee colony (ABC), ant colony optimization (ACO), etc. Gravitational search algorithm (GSA), proposed by Rashedi et al. [1], is inspired by the Newton theory of gravity and has simple calculation structure and be easy to implement.

GSA has fast convergence when solving nonlinear and high-dimensional functions. However, this feature may also lead to premature convergence and then trapped in local optimum. In order to overcome its weakness, many researchers have tried to apply periodicity and disorder of chaos to combine with the GSA algorithm [2, 3]. Li et al. proposed a hybrid algorithm, namely chaotic gravitational search algorithm, which has two kinds of searches, using GSA for large-scale search to get a better solution, and chaotic search based on the better solution to jump out of local solutions [2]. But this improved method is only suitable for low-dimensional objective function cases, when it

is used for the high-dimensional objective function, not only the time of calculation increases, but also the results are prone to failure. Yang applied niche technology to the GSA optimization process, adjusting the fitness value of each particle by the similarity between particles, thereby increasing the diversity of optimized particles [4]. Kumar et al. used the nonlinear method to dynamically adjust the gravitational constant, improving the ability of exploration and exploitation [5]. Ju et al. integrated the support vector machine (SVM) method to optimize the parameters in GSA, and the parameters are dynamically adjusted in each search step to make the algorithm perform global search [6]. However, this algorithm needs to choose a SVM model, and the data of processing capacity is large. S. Sarafrazi et al. introduced the interrupt operator to improve the GSA's ability of exploration and exploitation [7].

In the above, the improvement methods of GSA are developed from different aspects, but it usually causes a heavier computational load after the improvement, or more complexity of the algorithm due to introducing other models. In this paper, the advantages and disadvantages of the GSA are firstly analyzed, then the gravitational constant is modified to ensure maintaining the particle swarm diversity. To enhance reliability, a sharing mechanism is also introduced. So the optimal information in the particle swarm during each search step can be shared to every particle. This paper is organized as follows: The GSA model, principles and convergence are introduced in Sect. 2, including the formulation of GSA, proof of convergence and advantages and disadvantages analysis. The improvement way is proposed in Sect. 3. Algorithm steps are given in Sect. 4. A numerical example is presented in Sect. 5. IEEE9 optimal power flow is solved by IGSA in Sect. 6 and conclusions are discussed in Sect. 7.

2 Gravitational Search Algorithm

2.1 Brief Introduction of GSA

GSA is based on the law of universal gravitation and Newton's second law. It regards the search space as a space with gravitational field, the particle mass in space is determined by the current fitness value. Definitely, the closer to the optimal solution,

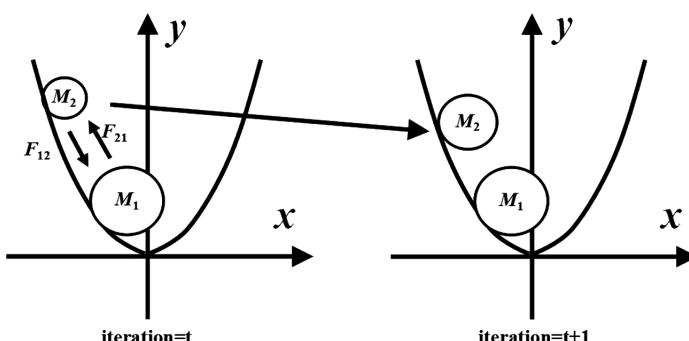


Fig. 1. The motion of particles in GSA during two adjacent searches.

the higher the particle mass, and other particles move to the position of the optimal solution under its action. In order to preferably illustrate the search process of the GSA algorithm, the search principle of the algorithm is explained by using a unary quadratic function ($y = x^2$), from Fig. 1, in the iteration t , M_1 is a better solution and therefore the mass is larger, and M_2 is a particle around the better solution, so the quality is smaller. Known by the law of universal gravitation and Newton's third law $F_{12} = F_{21}$, the acceleration of M_2 is greater than the acceleration of M_1 , and M_2 moves closer to M_1 .

2.2 Model of GSA

Assumed: there are N agents (masses) in D -dimensional search space, and the position of the i th agent is $X_i = (x_i^1, x_i^2 \dots x_i^K \dots x_i^D)$, where $i = 1 \dots N$.

According Newton gravitation theory, at time t , the force on the i th agent from the j th agent in dimension K is defined as Eq. (1).

$$F_{ij}^k(t) = G(t) \frac{M_{ii}(t) \times M_{jj}(t)}{R_{ij}(t) + \varepsilon} (x_j^k(t) - x_i^k(t)) \quad (1)$$

Where M_{ii} and M_{jj} are masses of agents at time t , $G(t)$ is the gravitational constant at time t . $R_{ij}(t)$ denotes the Euclid geometric distance between i th agent and j th agent. ε is a particularly small constant.

In order to increase the diversity in the particle search process, random (rand) numbers are introduced. Then the sum of the forces putting on from other agents in a dimension K is defined as Eq. (2).

$$F_i^k(t) = \sum_{j=1, j \neq i}^N \text{rand}_j F_{ij}^k(t) \quad (2)$$

Where $F_{ij}^k(t)$ denotes the force between i th agent and j th agent at time t , rand_j is a random number in the interval $[0,1]$.

Based on the law of motion, the acceleration of the i -th agent at time t in a dimension K is defined as Eq. (3).

$$a_i^k(t) = \frac{F_i^k(t)}{M_{ii}(t)} \quad (3)$$

Then, the next velocity and position of the agent i are defined as Eqs. (4) and (5):

$$v_i^k(t+1) = \text{rand}_i \times v_i^k(t) + a_i^k(t) \quad (4)$$

$$x_i^k(t+1) = x_i^k(t) + v_i^k(t) \quad (5)$$

2.3 Analysis of the Advantages and Disadvantages of GSA

By analyzing the GSA optimization principle, it can be found that the direction and step size of particle optimization are determined by the force. In addition, some particles with small mass produce greater acceleration, and some particles with higher mass produce less acceleration from Eq. (3). Therefore, in a search process, these particles with small mass have larger sizes, which is beneficial to the global search. On the contrary, the larger mass is smaller, which is beneficial to the maintenance of the current optimal value. However, [8, 9] point out GSA has a faster convergence speed, but it is easy to converge prematurely and fall into local optimum. Next, global convergence criterion of stochastic optimization algorithm that is raised by Solis and Wets is used to prove the convergence of GSA [10].

Hypothesis 1: $f(D(z, \xi)) \leq f(Z)$, where $\xi \in S$, if $f(D(z, \xi)) \leq f(\xi)$, then it can be got that Random algorithm is global convergence. In other words, $\{f(z_k)\}_{k=1}^{\infty}$ converge on ϕ .

Hypothesis 2: For any Borel subset A of S , if its measure $v(A) > 0$, then it satisfies $\prod_{k=0}^{\infty} (1 - \mu_k(A)) = 0$ where $\mu_k(A)$ is the probability of A obtained by the measure μ_k .

Theorem 1: GSA satisfies hypothesis 1.

Proves: Define the function $D(\cdot)$ as Eq. (6).

$$D(p_{g,k}, x_{i,k}) = \begin{cases} p_{g,k} & , f(g(x_{i,k})) \geq f(p_{g,k}) \\ g(x_{i,k}) & , f(g(x_{i,k})) < f(p_{g,k}) \end{cases} \quad (6)$$

where $g(x_{i,k})$ denotes algorithm update equation, which can be written by Eq. (7).

$$x_{i,k+1} = g(x_{i,k}) = g_1(x_{i,k}) + g_2(x_{i,k}) + g_3(x_{i,k}) \quad (7)$$

where $g_1(x_{i,k}) = x_{i,k} + \omega v_{i,k}$, $g_2(x_{i,k}) = C_1 U G_1$ and $g_3(x_{i,k}) = C_2 U G_2$, C_1 and C_2 are learning factors, $x_{i,k}$ indicates the position of the mass point at the k -th generation.

According to the above definitions, it can be proven GSA satisfies hypothesis 1.

Theorem 2: Arbitrarily $\varepsilon > 0$, existing $N \geq 1$ makes for any $n \geq N$, if choosing ω , Φ_1 , Φ_2 , it exist $\max(\|\alpha\|, \|\beta\|) < 1$, then it can be proven $\|g^n(x_{i,k}) - g^{n+1}(x_{i,k})\| < \varepsilon$.

Proves: According to the $\lim_{t \rightarrow \infty} (X(t)) = \lim_{t \rightarrow \infty} (k_1 + k_2 \alpha^t + k_3 \beta^t) = \frac{\varphi_1 p + \varphi_2 p_g}{\varphi_1 + \varphi_2}$ and $X(t+1) - X(t) = V(t+1)$, when $\max(\|\alpha\|, \|\beta\|) < 1$, Eq. (8) is

$$\lim_{t \rightarrow \infty} (V(t+1)) = \lim_{t \rightarrow \infty} (X(t+1) - X(t)) = \lim_{t \rightarrow \infty} (k_2 \alpha^t (\alpha - 1) + k_3 \beta^t (\beta - 1)) = 0 \quad (8)$$

According to Eq. (8), $\lim_{t \rightarrow \infty} X(t+1) = \lim_{t \rightarrow \infty} X(t)$ and there is a presence at Eq. (9).

$$X(t+1) = X(t) + V(t+1) = X(t) + \omega V(t) - X(t)(\varphi_1 + \varphi_2) + \varphi_1 p + \varphi_2 p_g \quad (9)$$

Simplify Eq. (9).

$$X(t+1) = X(t) + \omega V(t) - X(t)(\varphi_1 + \varphi_2) + \varphi_1 p + \varphi_2 p_g \quad (10)$$

According to $\lim_{t \rightarrow \infty}(V(t)) = 0$, $\lim_{t \rightarrow \infty}(X(t+1)) = \lim_{t \rightarrow \infty}(X(t))$ and assuming $p = p_g + \Delta p$, it is rewritten as Eq. (11).

$$\begin{aligned} \omega V(t) &= X(t)(\varphi_1 + \varphi_2) - \varphi_1(p_g + \Delta p) - \varphi_2 p_g \\ &= X(t)(\varphi_1 + \varphi_2) - \varphi_1 p_g - \varphi_2 p_g - \varphi_1 \Delta p \end{aligned} \quad (11)$$

Simplify Eqs. (11), and (12) is got.

$$-\omega V(t) = (p_g - X(t))(\varphi_1 + \varphi_2) + \varphi_1 \Delta p \quad (12)$$

According to Eq. (12), $X(t)$ eventually converges to P_g . At this point, the particles velocity ($V(t)$) is 0. But if P_g is local optimal solution. GSA will fall into local optimum.

The above problems created by GSA in the search process are mainly caused by two reasons. Firstly, the particle swarm diversity is decline rapidly. From Eqs. (3)–(5), it can be found the search direction and step size of the particle depend on the particle mass, and the current position of the particle is set by the fitness value. In the early iteration, all particles are randomly distributed in the search space, but if one particle finds a better solution (not global optimal solution), all other particles move to this location, each particle has the same mass when the particles are in one position, and it will be lead to the attraction between particles is almost the same, according to Eq. (12), GSA is stagnated. Secondly, no information sharing and transmission among the optimized particles, from Eqs. (4) and (5), it can be found that the particle only uses own state information to update velocity and position, and does not consider global optimal information. The multimodal function problem is easily raised due to the local optimum caused by two reasons.

3 Improvements on GSA

Based on the above analysis, GSA is an algorithm with sharp decline in diversity and lack of memory. In this paper, the algorithm is improved from the following two aspects.

3.1 Gravitational Constant G Adjustment Strategy

According to the Eq. (1), it can be noticed the determinants of the interaction force among particles include the mass of the gravitational force of the particle M , the mutual position R among the particles, and the gravitational constant G . Furthermore, the first two factors (M and R) are determined by the spatial positions of the particles, but the gravitational constant G is an artificially introduced variable that is independent to the search process.

Thus, the gravitational constant G directly determines the magnitude of the force during the search step, and then determines the speed and acceleration of the particle movement, and it can affect the step size and direction. Therefore, gravitational constant G has a direct impact on algorithm performance. [1] gives the definition of $G(t)$ as Eq. (13).

$$G(t) = G_0 \times e^{-\alpha t/T} \quad (13)$$

Where G_0 is the initial value, t is the current iterations, T is the maximum iterations, α is attenuation factor, and according to [1], GSA has better performance when $G_0 = 100$, $\alpha = 20$.

According to Eq. (13), it can be shown $G(t)$ is an exponential decay function, and the decay rate of $G(t)$ will be fast under the action of the attenuation factor α . Although this effectively avoids the oscillation of the algorithm, it reduces the diversity of the particle swarm, which leads to the algorithm falling into the local optimum. To this end, this paper improves the gravitational constant adjustment strategy that is to design $G(t)$ as component function as Eq. (14).

$$\Delta G(t) = \begin{cases} G_0 \times e^{-\alpha t/T} & G > \lambda \\ \lambda - \lambda/(T-t+1) & G \leq \lambda \end{cases} \quad (14)$$

Where: λ denotes the segmentation values, and other parameter is same to Eq. (13).

It can be obtained that the gravitation constant does not change at the initial stage of the iteration, so that the algorithm guarantees the original global search ability. Gravitational constant exhibits inverse proportional function attenuation when $G(t)$ less than λ . From Fig. 2, it can be found that the improved gravitational constant decreases the gravitational decay rate compared to the gravitational constant former. This adjustment is also beneficial to the particles jumping out of the local optimum.

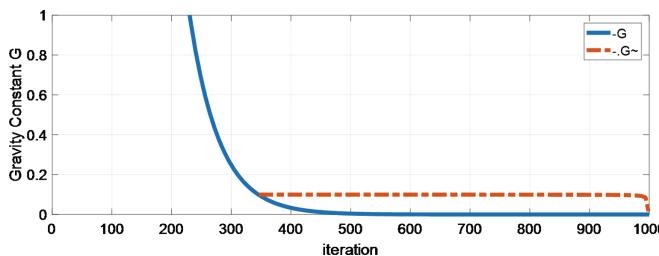


Fig. 2. Gravitational constant change law ($\lambda = 0.1$)

3.2 Introducing Particle Swarm Information Sharing Mechanism

According to Eqs. (4) and (5), GSA only uses its own information to search the best position, which is not conducive to global optimal search. So, if it can share global optimal information to each particle, then the global search capabilities can be

enhanced. For GSA, The optimal solution in each optimization iteration represents the information of the current optimized particle swarm in the particle swarm. Based on the above analysis, the new particle swarm update strategy is defined as Eq. (15).

$$x_i^k(t+1) = x_i^k(t) + v_i^k(t+1) + \omega \cdot v_{best}^k(t+1) \quad (15)$$

Where: $v_{best}^k(t) = \min(fit_i^k)$ (for a minimization problem) or $v_{best}^k(t) = \max(fit_i^k)$ (for a maximization problem), and fit_i^k denotes the optimal solution of the i -th particle in the k -th optimization process. ω denotes weight coefficient.

According to the proof of GSA convergence in the previous section, the main reason that the algorithm falls into local optimum is that the local optimal solution causes the particle velocity to be zero, so the algorithm falls into a stagnant state. Equation (16) is gotten by Eqs. (12) and (15).

$$-\omega V(t) + V^{best}(t+1) = (p_g - X(t))(\varphi_1 + \varphi_2) + \varphi_1 \Delta p \quad (16)$$

According to the Eq. (16). When the particle $X(t)$ runs to the local best P_g , however, the particle swarm will not stop searching as the existence of $V^{best}(t+1)$. Only having $V(t) = 0$ and $V^{best}(t+1) = 0$ are established at the same time, that is the global optimal solution, the particle will stop at the global optimal solution.

4 Steps of IGSA

Step 1: Initialize n particles, randomly distribute the particles throughout the search space, and give an initial velocity to each particle.

Step 2: Calculate the n particle fitness values, record the current particle group optimal value using $v_{best}^k(t) = \min(fit_i^k)$ (for a minimization problem) or $v_{best}^k(t) = \max(fit_i^k)$ (for a maximization problem).

Step 3: Update particles positions according to the Eqs. (5) and (15), respectively.

Step 4: Preserve the better results for fitness values by using above two particle movement results.

Step 5: Output optimal value when the loop reaches the maximum number of iterations.

5 Case Study I: Nonlinear Benchmark Functions

To evaluate the performance of IGSA, we applied it to 9 standard benchmarks. The test results are also compared with traditional GSA and PSO. These benchmark functions are presented in Table 1. The first three functions (F_1 to F_3) are unimodal high-dimensional functions, F_4 to F_6 are multimodal high-dimensional functions, F_7 to F_9 are multimodal low-dimensional functions.

Table 1. Test functions

| Function | Dim | Range |
|---|-----|------------------|
| $F_1(x) = \sum_{i=1}^n x_i^2$ | 30 | [-100,100] |
| $F_2(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2$ | 30 | [-100,100] |
| $F_3(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2) + (x_i - 1)^2]$ | 30 | [-30,30] |
| $F_4(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$ | 30 | [-5.12,5.12] |
| $F_5(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$ | 30 | [-600,600] |
| $F_6(x) = \frac{\pi}{n} \left\{ \sum_{i=1}^{n-1} (y_1 - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})] + (y_n - 1)^2 \right\}$ + $\frac{\pi}{n} 10 \sin(\pi y_1) + \sum_{i=1}^n u(x_i, 10, 100, 4)$ | 30 | [-50,50] |
| $F_7(x) = \left(\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^{25} (x_i - a_{ij})^6} \right)^{-1}$ | 2 | [-65.536,65.536] |
| $F_8(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_x - 4x_2^2 + 4x_2^4$ | 2 | [-5,5] |
| $F_9(x) = -\sum_{i=1}^4 c_i \exp\left(-\sum_{j=1}^3 a_{ij}(x_j - p_{ij})^2\right)$ | 3 | [1, 3] |

5.1 Optimal Parameters in IGSA

According to the improvement methods in the previous section Eqs. (14) and (15), which include parameters λ and ω . Some tests are designed to examine the effect of the different values of λ and ω on the performance of IGSA. Four different types benchmark functions from Tables 1 selected in these calculations. The parameters for setting GSA and IGSA are 30 particle size, the maximum iterations is 1000, and the gravitational attenuation factor α is 20. In addition, for IGSA, the segmentation values λ are respectively 0, 0.05, 0.1 and 1.2, respectively. The weighting factors ω are 0, 0.8, 1 and 1.2. The results are averaged over 20 independent runs, and Table 2 is results with different λ ($\omega = 1$) search, Table 3 is results with different ω ($\lambda = 0.1$).

Table 2. Comparison of optimization results with different λ ($\omega = 1$)

| | F_2 | F_3 | F_4 | F_7 |
|------------------|-----------|---------|---------|-------|
| $\lambda = 0$ | 2.4898 | 13.0208 | 22.8840 | 0.998 |
| $\lambda = 0.05$ | 0.0342 | 3.3823 | 15.9197 | 0.998 |
| $\lambda = 0.1$ | 2.2528e-5 | 0.0991 | 16.9160 | 0.998 |
| $\lambda = 1.2$ | 0.0175 | 0.6013 | 19.9016 | 0.998 |

In Table 2, the result of taking the piecewise function is better than the result of mono attenuation function ($\lambda = 0$), in addition, IGSA has the strongest search ability when λ is 0.1. In Table 3, the algorithm which introducing the particle information sharing mechanism is better than the previous algorithm results ($\omega = 0$), similarly, the

Table 3. Comparison of optimization results with different ω ($\lambda = 0.1$)

| | F_2 | F_3 | F_4 | F_7 |
|----------------|-----------|--------|---------|--------|
| $\omega = 0$ | 160.6137 | 8.2005 | 19.9010 | 1.1621 |
| $\omega = 0.8$ | 3.4586e-5 | 0.3445 | 16.9161 | 0.9980 |
| $\omega = 1$ | 2.2528e-5 | 0.0991 | 16.9160 | 0.9980 |
| $\omega = 1.2$ | 2.9459e-5 | 0.1225 | 21.8909 | 0.9980 |

strength of the information sharing mechanism does not show a linear relationship with the optimization results, but the algorithm has the best performance when ω is 1. In summary, IGSA shows the best exploration and exploitation capabilities when the segmentation values λ is 0.1 and the weighting factors ω is 1.

5.2 Comparative Study: IGSA, GSA and PSO

The effectiveness of IGSA is tested by comparisons among IGSA, GSA and PSO. The parameters of GSA and IGSA are set as follows: population size is 30, iteration number is 1000, $G_0 = 100$, α is 20, λ is 0.1, ω is 1. For PSO, the swarm size is 30, iteration number is 1000, C_1 is 2, C_2 is 2, ω is decreased linearly from 0.9 to 2. The stop criteria is cycle index attained its maximum iteration. For F_7 to F_9 , their dimensions are 2, 2 and 3, respectively. The results are averaged over 30 runs and the best solutions are reported in Table 4.

Table 4. Minimization results of benchmark functions

| Function | BEST | Std.PSO | | Std.GSA | | IGSA | |
|----------|---------|-----------------|----------------|-----------------|----------------|------------------|-------------------|
| | | Ave | Best | Ave | Best | Ave | Best |
| F_1 | 0 | 2.83e-04 | 8.75e-06 | 1.45e-16 | 1.52e-16 | 8.11e-06 | 1.07e-05 |
| F_2 | 0 | 5.19e+3 | 1.91e+3 | 506.5918 | 459.3045 | 5.08e-2 | 2.2528e-5 |
| F_3 | 0 | 201.665 | 15.5933 | 35.0076 | 26.2566 | 3.602e-01 | 9.91e-02 |
| F_4 | 0 | 72.9581 | 55.7182 | 31.1185 | 24.1444 | 29.1542 | 16.9160 |
| F_5 | 0 | 5.43e-03 | 9.96e-07 | 10.0896 | 12.5974 | 7.43e-02 | 2.3128e-07 |
| F_6 | 0 | 2.29 | 1.06e-1 | 3.30e-03 | 1.1e-02 | 2.07e-2 | 3.898e-08 |
| F_7 | 0.998 | 0.998 | 0.998 | 3.14 | 1.0003 | 0.998 | 0.998 |
| F_8 | -1.0316 | -1.0316 | -1.0316 | -1.0316 | -1.0316 | -1.0316 | -1.0316 |
| F_9 | -3.8628 | -3.8628 | -3.8628 | -3.8625 | -3.7756 | -3.8628 | -3.8628 |

As Table 4 illustrates, in unimodal high-dimensional F_2 and F_3 , IGSA has the ability to find the optimum, GSA provides much better results than IGSA and PSO. The average value of IGSA's optimization results for F_2 is four orders of magnitude higher than that of GSA and is seven orders of magnitude higher than PSO's because it is more robust. Particularly, for F_3 (Rosenbrock function) that some literatures [11, 12] improved the GSA algorithm from various ways, but they did not achieve good results. However, the improved strategy adopted in this paper makes IGSA has stronger search ability than standard GSA and PSO or test-GSA and BPO-GSA. IGSA's optimization

results for F_3 are improved by four orders of magnitude compared to these algorithms, and the results are closer to the optimal value.

In F_1 , IGSA does not skillfully explore the search space to find the optimum. For the multimodal high-dimensional functions F_4 , IGSA has not been significantly improved. Whereas for function F_5 or F_6 , IGSA has not only a very powerful ability to jump out of the local optimum, but also has a high convergence rate. It shows that IGSA algorithm has strong optimization ability in multi-peak optimization state. For multimodal low-dimensional F_7 - F_9 , IGSA and PSO both succeeded in find the optimal solution but the GSA failed to find the solution. In summary, IGSA has a high adaptability for benchmark functions of different types.

In order to visualize the optimization process of IGSA and GSA, the optimization curves of three functions (F_2 , F_6 , F_7) are selected from the above nine test functions in this paper. From Fig. 3, it is obvious that the F_2 optimization results of IGSA are significantly better than GSA, and IGSA's optimization speed (Iteration = 300) is twice as fast as GSA optimization (Iteration = 600). From Fig. 4, the optimal solution of F_6 under IGSA is also better than the optimal solution under GSA method, and the convergence speed is increased by about five times. According to the IGSA optimization curve (dashed line) in Fig. 5, there are four turns, and the GSA optimization curve has only one turn. It can be explained that IGSA jumps out of the local optimal solution multiple times in the optimization process for global optimization. Therefore, according to the change of the optimization curve, it can be concluded that IGSA has not only stronger search ability than GSA, but also the convergence speed is also improved.

In summary. By comparing the optimization results of the nine test functions under different algorithms, it can be seen that although the IGSA fails to improve the optimization for the F_1 function, the optimal solution and the convergence speed for other test functions are significantly improved compared with the standard GSA and PSO. Moreover, the Rosenbrock function optimization result is a step closer to the optimal solution, which fully demonstrates the superiority and feasibility of IGSA.

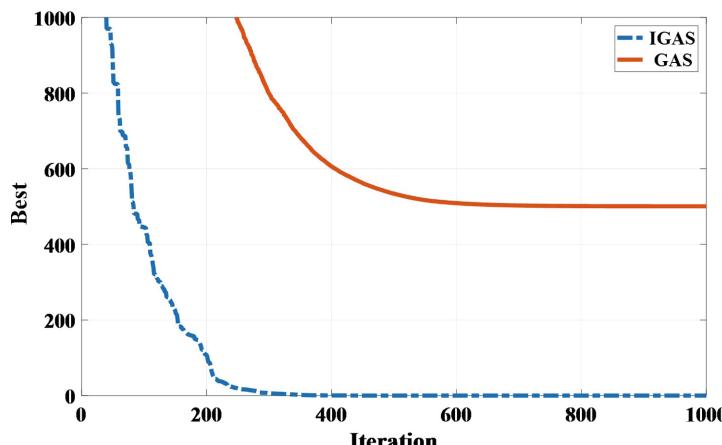


Fig. 3. Comparison of performance of GSA and IGSA for minimization of F_2 .

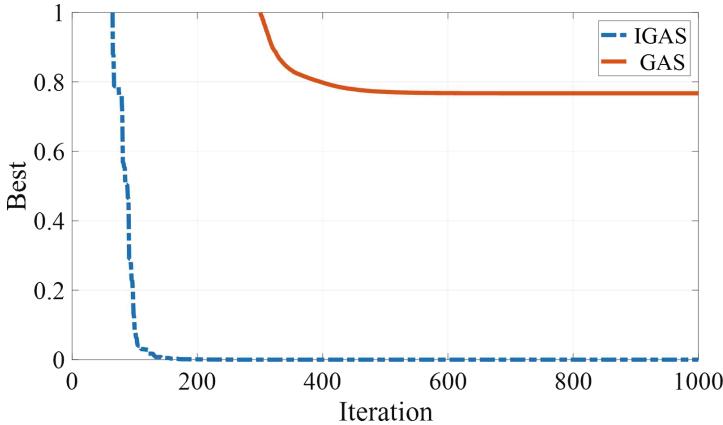


Fig. 4. Comparison of performance of GSA and IGSA for minimization of F_6 .

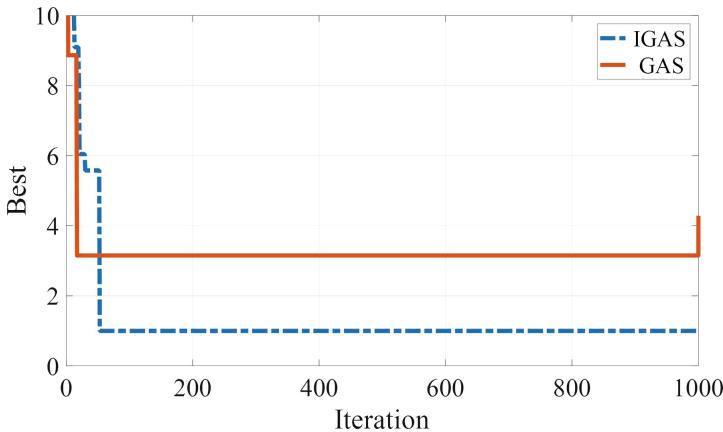


Fig. 5. Comparison of performance of GSA and IGSA for minimization of F_7 .

6 Case Study II: Power Flow in IEEE 9

6.1 Objective Function of the Problem

The optimal power flow of the power system is to meet the safe and reliable operation mode, and to adjust the output of the generator or some control devices, so that one or more performance indicators (such as power generation cost, network loss, etc.) of the system achieve the optimal operation. Such objectives are generally nonlinear, non-convex. Therefore, it is necessary to introduce an intelligent algorithm to find the optimal power flow at the lowest cost of power generation.

The objective function of this case study is the lowest operation cost of the generator. This requires that the generator's active power is optimally distributed under the

premise of satisfying the constraints and inequality constraints of the power system and the unit. Its model can be defined as Eq. (17).

$$F_{\text{Cost}} = \sum_{i=1}^m f_i(P_{Gi}) = \sum_{i=1}^m (a_i + b_i P_{Gi} + c_i P_{Gi}^2) \quad (17)$$

Where P_i denotes the active power of generator i -th, a_i , b_i and c_i denote the cost coefficient of the i th generator.

The equality constraint guarantees the safe and reliable operation of the power system. It consists of the active power and reactive power balance equations. They can be defined as Eqs. (18) and (19).

$$P_{Gi} = P_{Di} + U_i \sum_{j=1}^n U_j (G_{ij} \cos \delta_{ij} - B_{ij} \sin \delta_{ij}) \quad (18)$$

$$Q_{Gi} = Q_{Di} + U_i \sum_{j=1}^n U_j (G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}) \quad (19)$$

Where P_{gi} (Q_{gi}) denotes the active (reactive) output of i th bus, P_{Di} (Q_{Di}) denotes the consumption of load active (reactive) power at the i th node. G_{ij} (B_{ij}) denotes the conductance (susceptance) between i th bus and j th bus, δ_{ij} denotes the voltage phase angle difference between i th bus and j th bus.

The inequality constraint guarantees the qualification of the power quality of the power system, usually includes the node voltage amplitude, the generator active power and the reactive power output. They are defined as Eqs. (20)–(22).

$$P_{Gi,\min} < P_{Gi} < P_{Gi,\max} \quad (20)$$

$$Q_{Gi,\min} < Q_{Gi} < Q_{Gi,\max} \quad (21)$$

$$U_{Gi,\min} < U_{Gi} < U_{Gi,\max} \quad (22)$$

6.2 Solutions and Discussions

In order to verify the feasibility and effectiveness of IGSA to solve industrial problems, IGSA is tested on the IEEE 3-machine 9-bus system. The system consists of 3 generators, 7 lines, with a total load of 3.15 (Base capacity $S_B = 100$ MVA). The cost function of generators G_1 , G_2 and G_3 is a quadratic function as given in Table 5. Bus 5, 6, and 8 have constant power loads, for the parameters, it can be found in power flow calculation package [13].

Table 5. Generator cost parameter

| Generator | P_{\min} | P_{\max} | Cost parameter | | |
|-----------|------------|------------|----------------|-----|-----|
| | | | a | b | c |
| G_1 | 0.1 | 2.5 | 0.011 | 5 | 150 |
| G_2 | 0.1 | 3.0 | 0.0085 | 1.2 | 600 |
| G_3 | 0.1 | 2.7 | 0.1225 | 1 | 335 |

According to the Eq. (17), the model is multimodal high-dimensional functions. Therefore, using traditional algorithm or GSA may get a local optimal solution. IGSA is applied to solve this problem, maximum number of iterations for IGSA and GSA is set to 50. Population size is set to 30, and the values meet all of inequality constraints. Other parameters are consistent with Sect. 5.1. In this case, G_1 is a slack bus, the power generation cost is the lowest by optimizing the active output of G_2 and G_3 .

In order to further illustrate the superiority of IGSA, the optimal power flow results under the GSA algorithm and the independent incremental transmission losses (IITL) which is given on [14] are compared. From Table 6, it can be found the IEEE 9 system generates a power generation cost of \$1095.0 per hour and a network loss of 0.04 MW under the traditional power flow method (N-R method). If the system is carried out by using the algorithm IGSA, the optimization result is generator power generation costs \$1088.7 per hour, saving \$6.3 per hour compared with the N-R method. The power generation will save a considerable amount of money for a continuous operation mode. Moreover, network loss is only 3.43×10^{-2} MW, reducing 5.7×10^{-3} MW compared with the N-R method. Therefore, the optimal power flow results based on IGSA are results of win-win between the machine side and the network side. Meanwhile, IGSA has better results than the above two optimization algorithms, GSA and IITL, which indicates that IGSA has high practical value for solving industrial problems.

Table 6. Comparison of algorithm results

| Algorithm | Cost (\$/h) | Network loss (MW) |
|-----------|-------------|-------------------|
| N-R | 1095.0 | 0.04 |
| IITL | 1092.4 | 0.0407 |
| GSA | 1089.3 | 0.0377 |
| IGSA | 1088.7 | 0.0343 |

7 Conclusions

This paper introduces the fundamental of GSA from the theory, and analyzes the causes of GSA trapping in the local optimal solution by using the global convergence criterion of stochastic optimization algorithm. In the IGSA algorithm, there is no other stochastic optimization models or parameters are introduced, only the GSA's own information is used to improve the performance of GSA. The IGSA effectively improves the

convergence speed and accuracy. Especially when the algorithm falls into local optimum, it can effectively jump out of the local optimal solution and perform global optimal search.

In order to evaluate the IGSA algorithm, 9 functions of three different types are examined. The result has greater advantages than PSO and GSA, showing strong search capabilities. In solving industrial problems, satisfactory results were obtained by calculating the optimal power flow for the IEEE 9-node power system. From the above descriptions, IGSA has potential practical value in solving nonlinear, non-convex, multi-peak programming problems.

In the future, IGSA will be used to solve multi-objective functions, and its convergence speed will be tested by IEEE 33-node power system.

Acknowledgments. This work was supported by NSFC-21868019 and NMGSFC-2018MS06022.

References

- Rashedi E, Nezamabadi H, Saryazdi S (2009) GSA: a gravitational search algorithm. *Inf Sci* 179(13):2232–2248
- Li C, Zhou J, Xiao J, Xiao H (2012) Parameters identification of chaotic system by chaotic gravitational search algorithm. *Chaos, Solitons Fractals* 45:539–547
- Gaos S, Vairappan C, Wang U et al (2014) Gravitational search algorithm combined with chaos for unconstrained numerical optimization. *Appl Math Comput* 231(11):48–62
- Yang Y (2018) Improvement of gravitational search algorithm. *Syst Simul Technol* 14 (1):78–82 (in Chinese)
- Kumar JV, Kumar DMV, Edukondaluk K (2013) Strategic bidding using fuzzy adaptive gravitational search algorithm in a pool based electricity market. *Appl Soft Comput* 13 (5):2445–2455
- Ju F, Hong W (2013) Application of seasonal SVR with chaotic gravitational search algorithm in electricity forecasting. *Appl Math Modell* 37:9643–9651
- Sarafrazi S, Nezamabadi H, Saryazdi S (2011) Disruption: a new operator in gravitational search algorithm. *Scientia Iranica* 18(3):539–548
- Khajehzadeh M, Taha MR, El-Shafie A et al (2012) A modified gravitational search algorithm for slope stability analysis. *Eng Appl Artif Intell* 25(8):1589–1597
- Rashedi E, Nezamabadi H, Saryazdi S (2012) GSA: a gravitational search Algorithm. *Intell Inf Manag* 4(6):390–395
- Solis F, Wets R (1981) Minimization by random search techniques. *Math Oper Res* 6:19–30
- Zhang N, Zhao Z, Bao X et al (2019) The gravitational search algorithm based on improved tent chaos. *Control Decis* 1:1–8 (in Chinese)
- Mirjalili S, Hashin S (2010) A new hybrid PSOGSA algorithm for function optimization. In: 2010 International conference computer and information application (ICCIA), Tianjin, pp 374–377
- Zimmerman R, Gan D. Matpower: a MATLAB power system simulation package. <http://www.pserc.cornell.edu/matpower>
- Gao L (2007) Energy-saving generation dispatching based on independent incremental transmission losses and environmental cost, Dissertation, Hunan University. (in Chinese)



Specular Reflection Separation for a Single Image Using IR-SIP

Yongqi Guo, Jianlin Wang^(✉), Kepeng Qiu, Liqiang Zhao,
Xuesong Fu, and Rutong Wang

Beijing University of Chemical Technology, Beijing 100029, China
wangjl@mail.buct.edu.cn

Abstract. Specular reflection separation can improve the accuracy of edge detection, target recognition, and feature matching. Existing methods can't completely separate the specular component for a single image. We propose an effective method to separate specularity for a single image using IR-SIP in this paper. Firstly, the pixels of the textured image are clustered in pseudo-chromaticity space. Secondly, we estimate the intensity ratio (IR) of diffuse pixels in each cluster based on series importance point (SIP). Finally, the specular component is separated according to the dichromatic reflection model. The experiments on the simulation image and natural images show the effectiveness of the proposed method.

Keywords: Dichromatic reflection model · Specular reflection separation · Intensity ratio · Series importance point

1 Introduction

In the real world, the specular reflection is widespread on the surface of various materials. As one of the most important problems in the field of machine vision, the specular reflection can increase the error of edge extraction, target detection, feature matching, and pattern recognition. Therefore, many methods for separating specular reflection have been proposed [1]. These methods are divided into two categories: methods for multiple images and methods for a single image.

Mallick et al. [2] used partial differential equations and color information between nearby image points to separate specular and diffuse components in the continuous domain; Shah et al. [3] exploited information in neighboring frames from video footage to suppress the effects of specularity. However, these methods for separating specular component using multiple images seem impractical in many application scenarios.

Some researchers placed additional polarization filters in front of the imaging sensors to separate specularity [4]. While these methods can separate specularity using a single image, utilizing additional polarization is difficult for various constraints.

Tan et al. [5] generated the specular-free (SF) image and used intensity logarithmic differentiation of the SF image and the normalized image to remove specular reflection iteratively; Shen et al. [6–8] analyzed the error of chromaticity, introduced the modified specular-free (MSF) image and presented the concept of intensity ratio (IR). Then specular reflection is separated by solving the least-squares problem of the dichromatic

reflection model or estimating the IR of diffuse pixels in each cluster; Yang et al. [9, 10] observed that the maximum diffuse chromaticity changes smoothly in local patches and utilized a low-pass filter to remove the noise contributed by the specular reflection; Nguyen et al. [11, 12] utilized tensor voting to obtain reflectance distribution and dominant color of diffuse reflection components, and specular components were removed from specular pixels based on the painting technique; Ren et al. [13] clustered image pixels using invariant variables to obtain color lines and estimate the illumination chromaticity to separate specularity. Most methods for separating specularity utilize iterative or solve linear (or nonlinear) systems, which is inappropriate in real-time applications. Although Yang's method and Shen's method have achieved good results, they require manual tuning of parameters and the IR of diffuse pixels may be inaccurate when the number of pure diffuse pixels and specular pixels in each cluster is not equal in Shen's method.

Inspired by Shen's method [8], we propose a method for removing specular reflection using IR-SIP. Firstly, the image pixels are clustered using the K-means algorithm in pseudo-chromaticity space; then the IR of pixels that contain pure diffuse reflection in each cluster are estimated based on series importance point (SIP) [14]. The specular component is separated effectively according to the dichromatic reflection model.

2 Reflection Model

The RGB color vector $\mathbf{J}(x) = [J_r(\mathbf{x}), J_g(\mathbf{x}), J_b(\mathbf{x})]^T$ is a linear combination of diffuse component and specular component at position \mathbf{x} . It can be expressed as

$$\mathbf{J}(\mathbf{x}) = \mathbf{J}^D(\mathbf{x}) + \mathbf{J}^S(\mathbf{x}) = m_d(\mathbf{x})\Lambda(\mathbf{x}) + m_s(\mathbf{x})\Gamma \quad (1)$$

where $m_d(\mathbf{x})$ and $m_s(\mathbf{x})$ are the coefficients of diffuse and specular reflection, which depend on the geometry of object.

Diffuse chromaticity and specular chromaticity can be defined as follows:

$$\Lambda_c = \frac{J_c^D}{\sum_{c \in \{r,g,b\}} J_c^D} \quad (2)$$

$$\Gamma_c = \frac{J_c^S}{\sum_{c \in \{r,g,b\}} J_c^S} \quad (3)$$

The reflectance properties of various materials determine the diffuse chromaticity $\Lambda(\mathbf{x}) = [\Lambda_r, \Lambda_g, \Lambda_b]$, which is identical for the uniform colored surface. The specular chromaticity $\Gamma = [\Gamma_r, \Gamma_g, \Gamma_b]$ can be assumed to be equal to the illumination chromaticity, which can be measured by imaging a whiteboard or estimated from the image, and we can easily obtain $\Gamma_r = \Gamma_g = \Gamma_b = 1/3$ and $J_r^S = J_g^S = J_b^S = J^S$.

3 Separation Method

In this section, we introduce the method for separating the specular component using IR [8]. Then, we analyze the proportion of diffuse pixels and propose the method for estimating the IR of diffuse pixels based on SIP. Finally, we summarize the algorithm of the proposed method.

3.1 Separation Method Using IR

The MSF image can be calculated as:

$$J_c^{\text{MSF}}(\mathbf{x}) = J_c(\mathbf{x}) - J_{\min}(\mathbf{x}) + \bar{J}_{\min} \quad (4)$$

where $J_{\min}(\mathbf{x}) = \min\{J_r(\mathbf{x}), J_g(\mathbf{x}), J_b(\mathbf{x})\} = m_d(\mathbf{x})\Lambda_{\min} + m_s(\mathbf{x})\Gamma$ is the minimum value of the RGB channels, and $\bar{J}_{\min} = \sum J_{\min}(\mathbf{x})/N$ is the mean of $J_{\min}(\mathbf{x})$ for all pixels.

The chromaticity of the MSF image can be defined as

$$\Lambda_c^{\text{MSF}}(\mathbf{x}) = \frac{J_c^{\text{MSF}}(\mathbf{x})}{J_r^{\text{MSF}}(\mathbf{x}) + J_g^{\text{MSF}}(\mathbf{x}) + J_b^{\text{MSF}}(\mathbf{x})} \quad (5)$$

The pseudo-chromaticity space is constructed as

$$\Lambda^P(\mathbf{x}) = [\Lambda_{\min}^{\text{MSF}}(\mathbf{x}), \Lambda_{\max}^{\text{MSF}}(\mathbf{x})]^T \quad (6)$$

where $\Lambda_{\min}^{\text{MSF}}(\mathbf{x}) = \min\{\Lambda_r^{\text{MSF}}(\mathbf{x}), \Lambda_g^{\text{MSF}}(\mathbf{x}), \Lambda_b^{\text{MSF}}(\mathbf{x})\}$ is the minimum chromaticity and $\Lambda_{\max}^{\text{MSF}}(\mathbf{x}) = \max\{\Lambda_r^{\text{MSF}}(\mathbf{x}), \Lambda_g^{\text{MSF}}(\mathbf{x}), \Lambda_b^{\text{MSF}}(\mathbf{x})\}$ is the maximum chromaticity. The pixels can be clustered based on the chromaticity distance.

In each cluster, the maximum intensity $J_{\max}(\mathbf{x}) = \max\{J_r(\mathbf{x}), J_g(\mathbf{x}), J_b(\mathbf{x})\}$ can be expressed as $J_{\max}(\mathbf{x}) = m_d(\mathbf{x})\Lambda_{\max} + m_s(\mathbf{x})\Gamma$. Consequently, the minimum and maximum intensities have the same specular component $m_s(\mathbf{x})\Gamma$. We can obtain the range intensity as

$$J_{\text{ran}}(\mathbf{x}) = J_{\max}(\mathbf{x}) - J_{\min}(\mathbf{x}) = m_d(\mathbf{x})(\Lambda_{\max} - \Lambda_{\min}) \quad (7)$$

which eliminates the specular component $m_s(\mathbf{x})\Gamma$.

The IR of the maximum to range intensity can be defined as

$$Q(\mathbf{x}) = \frac{J_{\max}(\mathbf{x})}{J_{\text{ran}}(\mathbf{x})} \quad (8)$$

for a pixel that contains pure diffuse reflection, the IR is

$$Q_d(\mathbf{x}) = \frac{\Lambda_{\max}}{\Lambda_{\max} - \Lambda_{\min}} \quad (9)$$

clearly, it is independent of surface geometry, and the IR of the specular pixel is

$$Q_s(\mathbf{x}) = \frac{\Lambda_{\max}}{\Lambda_{\max} - \Lambda_{\min}} + \frac{m_s(\mathbf{x})\Gamma}{m_d(x)(\Lambda_{\max} - \Lambda_{\min})} \quad (10)$$

which is relevant to $m_d(\mathbf{x})$ and $m_s(\mathbf{x})$.

It is clear that the IR of specular pixel is larger than that of diffuse pixel and the IRs of diffuse pixels is approximately constant from Eqs. (9) and (10). We can denote the constant as Q_d . If we assume that the proportion of diffuse pixels in each cluster is less than a percentile threshold T_p , then the \hat{Q}_d can be defined as $\hat{Q}_d = \text{rank}(Q_d)_{[T_p L]}$, where the operator $[\cdot]$ denotes rounding to nearest integer and L is the number of pixels. From Eqs. (7) and (9), the diffuse component $m_d(x)\Lambda_{\max} \approx Q_d I_{\text{ran}}$. Therefore, the specular component can be computed as

$$J^S(\mathbf{x}) = \max\{J_{\max}(\mathbf{x}) - \hat{Q}_d J_{\text{ran}}, 0\} \quad (11)$$

where the $\max\{\cdot\}$ operator ensures that the specular component is not smaller than zero. Then the diffuse component is computed as

$$\mathbf{J}^D(\mathbf{x}) = \mathbf{J}(\mathbf{x}) - \mathbf{J}^S(\mathbf{x}) \quad (12)$$

3.2 Separation Method Using IR-SIP

Separation method using IR choose the fixed value as the proportion of diffuse pixels. However, the IR of diffuse pixels may be inaccurate when the proportion of diffuse pixels and specular pixels in each cluster is not fixed. Figure 1 shows the clustering result of the toy image using threshold $T_C = 0.3$ and the IRs of pixels in the first cluster. It is observed that the IR is 1.26 when the proportion of diffuse pixels is 50% in the first cluster; however, the expected threshold is about 1.6, and the proportion of diffuse pixels may be not same in different clusters. Therefore, the fixed value as the proportion of diffuse pixels is improper in different clusters.

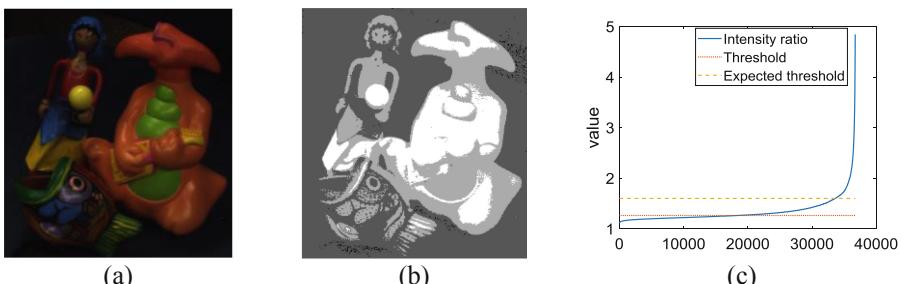


Fig. 1. Clusters of the toy image. (a) original image, (b) ratio map, and (c) the IRs in the first cluster.

Let $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots, \mathbf{x}_L\}$ be a dataset that collects data points in time sequence, where \mathbf{x}_i can be defined as $\mathbf{x}_i = (t_i, v_i)$. The data segment consisting of any two data points have three trends, such as the rising stage, the holding stage, and the falling stage (Fig. 2). The dividing point of different trends is SIP. The evaluation methods of the series point importance include the Euclidean distance (ED), the perpendicular distance (PD), and the vertical distance (VD). Figure 3 shows the different distances. We can calculate these distances as

$$ED = |AC| + |CB| = \sqrt{(t_1 - t_3)^2 + (v_1 - v_3)^2} + \sqrt{(t_2 - t_3)^2 + (v_2 - v_3)^2} \quad (13)$$

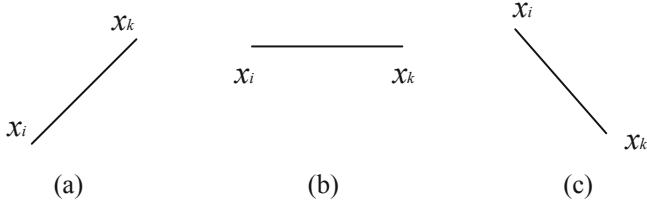


Fig. 2. Trends in data sequences. (a) the rising stage, (b) the holding stage, and (c) the falling stage.

$$PD = |CD| = \sqrt{(t_D - t_3)^2 + (v_D - v_3)^2} \quad (14)$$

$$VD = |CE| = |y_E - y_3| = \left| (y_1 + (y_2 - y_1) \frac{x_E - x_1}{x_2 - x_1}) - y_3 \right| \quad (15)$$

Therefore, the IRs of pixels in each cluster can be considered as data series included the rising stage and the holding stage. The IRs of diffuse pixels is in the holding stage, and the IRs of specular pixels is in the rising stage. Chung et al. [14] showed both PD and VD obtain the same results for evaluating series point importance, while ED got different behaviors. We can get the SIP of the IRs of pixels in each cluster based on PD. Then the y-coordinate of the SIP can be regarded as the IR of diffuse pixels. It is suitable to estimate the IR of diffuse pixels when the proportion of diffuse pixels and specular pixels is not fixed in each cluster.

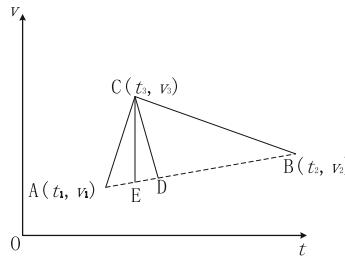


Fig. 3. The evaluation methods of the series point importance.

3.3 Outline of the Proposed Method

We summarize the algorithm of specular reflection separation in Algorithm 1 based on the above discussion.

Algorithm 1 IR-SIP

Input: Original image $\mathbf{J}(\mathbf{x})$, threshold T_c

While $i < N$

- Calculate distances of x_i to all cluster center
- Find the label with minimum distance $l_{\min} = \arg \min_l d_l(x_i)$
- If** minimum distance $d_{\min}(x_i) \leq T_c$ **then**

 - $\text{label}(x_i) \leftarrow l_{\min}$
 - Update cluster center of cluster l
 - else**

 - $L=L+1$
 - end**

- repeat-until** $\text{label}(x_i)$ remain unchanged
- for** $l=0; l < L; l++$

 - Calculate intensity ratio $Q_l(x_i)$
 - Sort the intensity ratio $Q_l(x_i)$
 - Find the SIP
 - Estimate the \hat{Q}_d
 - $J_c^D(\mathbf{x}) = J_c(\mathbf{x}) - \max\{J_{\max}(\mathbf{x}) - \hat{Q}_d J_{\text{ran}}, 0\}$

- end**

4 Experiment

In this section, we compare the separation results of the proposed method with the methods proposed by Shen [8] and Yang [10] on a synthetic image and the standard dataset built by Shen [8] or Yang [10]. The peak signal to noise ratio (PSNR) and the Structural Similarity Image Measurement (SSIM) [15] are used to evaluate the performance of different methods.

Figure 4 shows the synthetic image which is comprised of blue and red rendered hemispheres and the diffuse images obtained by different methods. We can observe that the bottom of each image is the PSNR value. The Shen's method [8] fail to remove all the specularity and produces low PSNR value. The diffuse image obtained by Yang's method [10] has a bad edge between the two hemispheres. The diffuse image obtained by our method has a better visual ability and higher PSNR value than the diffuse image obtained by Yang's method does. Figures 5 and 6 show the diffuse images of the masks image and the cups image. All the methods have almost the same performance from the visual appearance. But our method performs better than other state-of-the-art methods under PSNR quantitative measure. Figure 7 shows the diffuse images of the fruits

image. Yang's method produces the lowest PSNR value. The diffuse images obtained by Shen's method and our method have almost equal PSNR value.

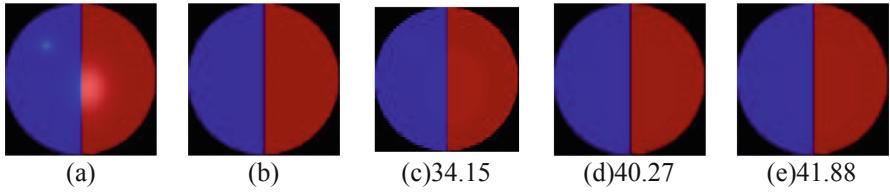


Fig. 4. Diffuse images of the synthetic image. (a) synthetic image, (b) ground truth, (c) diffuse image obtained by Shen's method [8], (d) diffuse image obtained by Yang's method [10], and (e) diffuse image obtained by our method.



Fig. 5. Diffuse images of the masks image. (a) image contains specular pixels, (b) ground truth, (c) diffuse image obtained by Shen's method [8], (d) diffuse image obtained by Yang's method [10], and (e) diffuse image obtained by our method.



Fig. 6. Diffuse images of the cups image. (a) image contains specular pixels, (b) ground truth, (c) diffuse image obtained by Shen's method [8], (d) diffuse image obtained by Yang's method [10], and (e) diffuse image obtained by our method.



Fig. 7. Diffuse images of the fruits image. (a) image contains specular pixels, (b) ground truth, (c) diffuse image obtained by Shen's method [8], (d) diffuse image obtained by Yang's method [10], and (e) diffuse image obtained by our method.

We also use SSIM to evaluate the performance of different methods. SSIM is a perceptual metric, which quantifies the structural similarity between the ground truth image and the original image. Table 1 shows that the image obtained by different methods has relatively high SSIM values. It is clear that all the methods can recover the main structures of the images. Our method performs better on the masks, cups, and fruits image in the perspective of SSIM. Though the image obtained by Yang’s method [8] has a higher SSIM value than that by our method on the ball image, it obtains low PSNR values because of incomplete separation. The quantitative comparison shows that our method can preserve edge information and tiny details well while separating the highlights.

Table 1. Comparison of different methods

| Method | Ball | | Masks | | Cups | | Fruits | |
|------------------|--------------|---------------|--------------|---------------|--------------|---------------|--------------|---------------|
| | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |
| Shen et al. [8] | 34.15 | 0.9516 | 36.88 | 0.9550 | 41.71 | 0.9783 | 39.57 | 0.9866 |
| Yang et al. [10] | 40.27 | 0.9823 | 33.99 | 0.9468 | 38.66 | 0.9488 | 37.17 | 0.9677 |
| Ours | 41.88 | 0.9812 | 37.26 | 0.9550 | 42.15 | 0.9784 | 39.26 | 0.9878 |

5 Conclusions

This paper proposes a method to separate specularity for a single image using IR-SIP. We cluster all pixels of the MSF image in chromaticity space and estimate the IR of diffuse pixels in each cluster based on SIP. The experimental results validate that the proposed method performs better in PSNR and SSIM. However, the proposed method cannot remove specularity of the achromatic region. Our future work is to solve these limitations to apply in general-purpose applications.

Acknowledgments. This work was supported by the National Key Research and Development Program of China [No. 2017YFF0107303].

References

1. An D et al (2015) Illumination estimation from specular highlight in a multi-spectral image. *Opt Exp* 23(13):17008–17023
2. Mallick SP et al (2006) Specularity removal in images and videos: a PDE approach. In: European conference on computer vision. Springer, Heidelberg
3. Shah SMZA, Marshall S, Murray P (2017) Removal of specular reflections from image sequences using feature correspondences. *Mach Vis Appl* 28(3–4):409–420
4. Kong N, Tai Y-W, Shin SY (2011) High-quality reflection separation using polarized images. *IEEE Trans Image Process* 20(12):3393–3405
5. Tan RT, Ikeuchi K (2005) Separating reflection components of textured surfaces using a single image. *IEEE Trans Pattern Anal Mach Intell* 27(2):178–193
6. Shen H-L et al (2008) Chromaticity-based separation of reflection components in a single image. *Pattern Recogn* 41(8):2461–2469

7. Shen H-L, Cai Q-Y (2009) Simple and efficient method for specularity removal in an image. *Appl Opt* 48(14):2711–2719
8. Shen H-L, Zheng Z-H (2013) Real-time highlight removal using intensity ratio. *Appl Opt* 52(19):4483–4493
9. Yang Q, Wang S, Ahuja N (2010) Real-time specular highlight removal using bilateral filtering. In: European conference on computer vision. Springer, Heidelberg
10. Yang Q, Tang J, Ahuja N (2014) Efficient and robust specular highlight removal. *IEEE Trans Pattern Anal Mach Intell* 37(6):1304–1311
11. Nguyen T et al (2014) A novel and effective method for specular detection and removal by tensor voting. In: 2014 IEEE international conference on image processing (ICIP). IEEE
12. Nguyen T et al (2015) Color structure recovering in strong specular text regions. In: 2015 13th international conference on document analysis and recognition (ICDAR). IEEE
13. Ren W, Tian J, Tang Y (2017) Specular reflection separation with color-lines constraint. *IEEE Trans Image Process* 26(5):2327–2337
14. Fu T et al (2008) Representing financial time series based on data point importance. *Eng Appl Artif Intell* 21(2):277–300
15. Wang Z et al (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612



A Calibration Method for 6-UPS Stewart Platform

Xuedong Jing, Yisheng Fang^(✉), and Ziteng Wang

Shanghai Institute of Technology, Shanghai 201418, China
etlon1995@gmail.com

Abstract. This paper focus on the calibration for 6-UPS Stewart platform. By the kinematic analysis, it is found that the error of final position and posture is mainly influenced by radius and 6 angles of joint of movable platform. Therefore, this paper proposes a calibration method which compensates the radius and 6 angles. According to the error analysis, a method of compensating the parameters is proposed, which applied interior-point algorithm. A simulation is set to verify this method. It indicates that the method could compensate simulated error and apparently reduce the error of 6 axis lengths.

Keywords: Stewart platform · Error analysis · Calibration

1 Introduction

The Stewart platform was firstly designed in 1954, and has been applied for many years. It is superior in maturity and stability to the other parallel robot.

The robots are designed to complete varies of tasks, which all require the robot to be fast, precise and stable. To control the robot precisely, plenty of details of axis structure require proper designing. The axis structure of Stewart platform is conducted by 3 main kinematic pairs. There included sphere pair(S), Hook joint(U), column pair(C), prismatic pair(P), revolving pair(R), planar contact pair(E) and helical pair(H). Some common Stewart platforms are using structures of SPS, RSS, PSS and so on. We designed a Stewart platform for spinal surgery in UPS structure, which is Synonymous with SPS. The reason why change sphere pair into Hook joint is the limit of machining precision in sphere pair.

However, the UPS structure is more convenient to control machining accuracy. The machining errors are results of the design, manufacture and assembly, which means unavoidable. To make the Stewart platform operating precisely, calibration before formal operation is significant, which calls a serial model analysis for the Stewart platform.

2 Model Description

2.1 Kinematic Analysis

To control the error of the machine structure, kinematic analysis would help to build an error model to find the critical parameters. The Stewart platform would be analyzed is shown in Fig. 1.

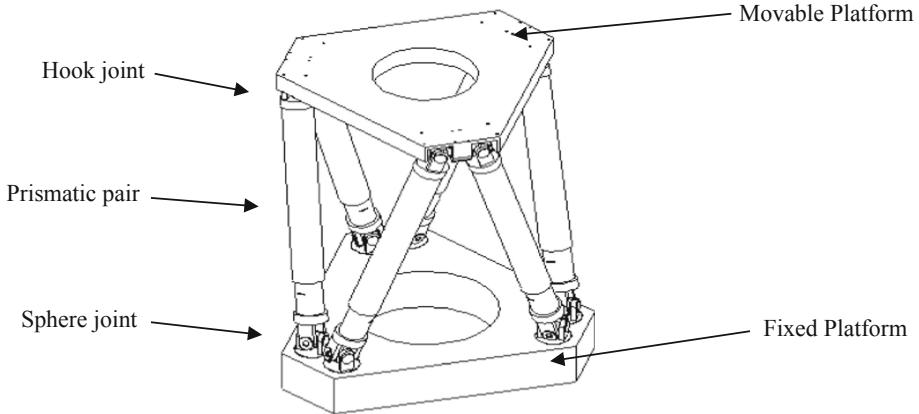


Fig. 1. The Stewart platform (6-UPS) established for spinal surgery

For kinematic analysis, the structure of Stewart platform could be simplified as Fig. 2. The frame $\{P\}$ is fixed on the centroid of movable platform, while frame $\{B\}$ is fixed on the centroid of fixed platform. Both platforms are semi-symmetric hexagon. The angle between P_0 and P_1 is α , while the angle between B_0 and B_1 is β [1, 2, 5]. The x-axis of $\{P\}$ is determined by the line of origin and P_6 , the angle between P_1 and x-axis is α . The x-axis of $\{B\}$ is similar with $\{P\}$. The length of 6 axis is given by

$$L = [l_1, l_2, l_3, l_4, l_5, l_6] \quad (1)$$

The coordinate of P_i in frame $\{P\}$ is

$$P_i = (r^P \cos \theta_i, r^P \sin \theta_i, 0)^T \quad (2)$$

where r^P is the radius of movable platform, θ_i is the angle of joint P_i . Similarly, the coordinate of B_i is

$$B_i = (r^B \cos \varphi_i, r^B \sin \varphi_i, 0)^T \quad (3)$$

where r^B is the radius of movable platform, φ_i is the angle of joint B_i . The transition matrix from $\{P\}$ to $\{B\}$ is

$${}^P_B T = \begin{bmatrix} {}^B_R & {}^B_P \\ 0 & 1 \end{bmatrix} \quad (4)$$

The L could be calculated by function

$$L = P^{\{B\}} - B^{\{B\}} = {}^P_B T \cdot P^{\{P\}} - B^{\{B\}} \quad (5)$$

According to the function (5), the L is unique when the position and posture of movable platform is confirmed.

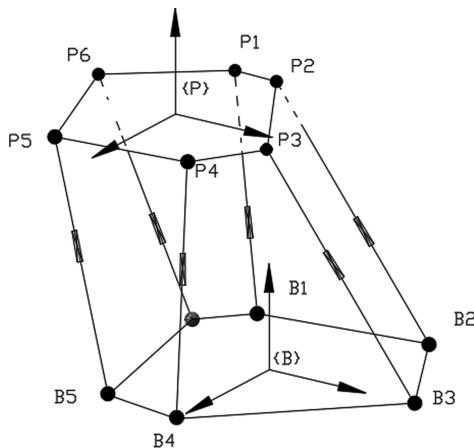


Fig. 2. Simplified Stewart Platform kinematic model

2.2 Error Analysis

By the function of inverse kinematic, we can find where the error comes from [2]. Assumed that the sphere joint and hook joint are perfect, the error model could be defined as $\Delta P_i, \Delta B_i$. It could set l_{off} which refers to offset value of axis length to calibrate, or reset the value of each parameter. The previous one is called software offset, the last one is called hardware offset. For software offset, W. Khalil and S. Besnard present a good method [3].

Once the position and posture of movable platform confirmed, the l is confirmed. By using precise measuring instrument, we can acquire the actual position and posture of movable platform, and build the measured position into transformation matrix as ${}^B_P T_m$ [4].

$$L_m = {}^B_P T_m \cdot P^{\{P\}} - B^{\{B\}} \quad (6)$$

The simplified ${}^B_P T$ and ${}^B_P T_m$ could be assumed as

$$\begin{bmatrix} n_x & o_x & a_x & x \\ n_y & o_y & a_y & y \\ n_z & o_z & a_z & z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} n_{mx} & o_{mx} & a_{mx} & x_m \\ n_{my} & o_{my} & a_{my} & y_m \\ n_{mz} & o_{mz} & a_{mz} & z_m \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The coordinate of vector l_i and l_{mi} is

$$l_i = \begin{pmatrix} n_x r^P \cos \theta_i + o_x r^P \sin \theta_i - r^B \cos \varphi_i \\ n_y r^P \cos \theta_i + o_y r^P \sin \theta_i - r^B \sin \varphi_i \\ n_z r^P \cos \theta_i + o_z r^P \sin \theta_i \end{pmatrix} \quad (7)$$

$$l_{mi} = \begin{pmatrix} n_{mx} r^P \cos \theta_i + o_{mx} r^P \sin \theta_i - r^B \cos \varphi_i \\ n_{my} r^P \cos \theta_i + o_{my} r^P \sin \theta_i - r^B \sin \varphi_i \\ n_{mz} r^P \cos \theta_i + o_{mz} r^P \sin \theta_i \end{pmatrix} \quad (8)$$

The difference between l_i and l_{mi} is

$$l_{ei} = \begin{pmatrix} (n_x - n_{mx}) r^P \cos \theta_i + (o_x - o_{mx}) r^P \sin \theta_i \\ (n_y - n_{my}) r^P \cos \theta_i + (o_y - o_{my}) r^P \sin \theta_i \\ (n_z - n_{mz}) r^P \cos \theta_i + (o_z - o_{mz}) r^P \sin \theta_i \end{pmatrix} \quad (9)$$

$$|l_{ei}| = r^P \sqrt{A \cos^2 \theta_i + B \cos \theta_i \sin \theta_i + C \sin^2 \theta_i} \quad (10)$$

Where A is $\sum_{i=x}^z (n_i - n_{mi})^2$, B is $\sum_{i=x}^z 2(n_i - n_{mi})(o_i - o_{mi})$, C is $\sum_{i=x}^z (o_i - o_{mi})^2$.

By this function, it could conclude that only one parameters of platform need to be calibrated.

2.3 Calibration Method

After measurement, there are n groups of data as

$$L_{ei} = [|^1 l_{ei}|, |^2 l_{ei}|, \dots, |^n l_{ei}|] \quad (11)$$

Then, the calibration of one axis is to find the solution of

$$\min Z = |^1 l_{ei}| + |^2 l_{ei}| + \dots + |^n l_{ei}|, (i = 1, 2, \dots, 6) \quad (12)$$

Considering the conditions of real machining error, the range of $r^{P'}$ and θ'_i is limited from 99% to 101% of itself.

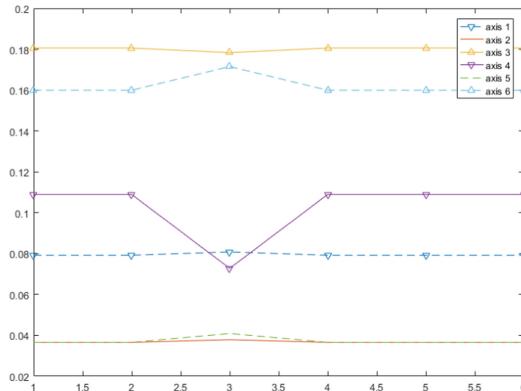
3 Simulation and Result

To verify the calibration method, we can set a random error (less than 0.1%) simulation, which is displayed in Table 1.

Table 1. Random error

| Parameters | Nominal value | Random error |
|------------|---------------|---------------|
| r^P | 60 | +0.03 |
| θ_1 | 25° | -0.07° |
| θ_2 | 120° | -0.02° |
| θ_3 | 145° | -0.17° |
| θ_4 | 240° | $+0.10^\circ$ |
| θ_5 | 265° | $+0.02^\circ$ |
| θ_6 | 360° | $+0.15^\circ$ |

Then, it simulates 6 kinds of position and posture as $(\frac{\pi}{6}, 0, 0, 0, 0, 0)$, $(0, \frac{\pi}{6}, 0, 0, 0, 0)$, $(0, 0, \frac{\pi}{6}, 0, 0, 0)$, $(0, 0, 0, 10, 0, 0)$, $(0, 0, 0, 0, 10, 0)$, $(0, 0, 0, 0, 0, 10)$. It calculates two groups of L , with random error or without random error, and got error data as Fig. 3.

**Fig. 3.** Error of 6 axis length

Because lacking the measured data, the simulation makes a substitution in function which would not influence the optimization. The function (9) changed into

$$l_{ei} = \begin{pmatrix} n_x(r^P \cos \theta_i - r_e^P \cos \theta_{ei}) + o_x(r^P \sin \theta_i - r_e^P \sin \theta_{ei}) \\ n_y(r^P \cos \theta_i - r_e^P \cos \theta_{ei}) + o_y(r^P \sin \theta_i - r_e^P \sin \theta_{ei}) \\ n_z(r^P \cos \theta_i - r_e^P \cos \theta_{ei}) + o_z(r^P \sin \theta_i - r_e^P \sin \theta_{ei}) \end{pmatrix} \quad (13)$$

Where r_e^P refers to radius of movable platform with error, θ_{ei} refers to angle of axis with error.

By using the MATLAB optimization toolbox to program, which contains interior point algorithm, the optimization could find a local optimal solution. After optimizing the parameters, the compensation value is almost equal to random error value, which could be concluded from Fig. 4. The error of length is also apparently reduced from

0.1 mm average to 2×10^{-7} average in Fig. 5. There are still many algorithms could be applied for calibration, which may prefer than interior point algorithm.

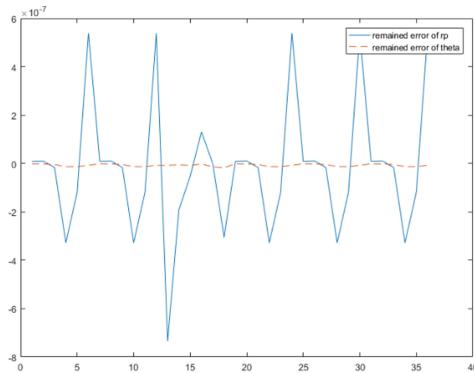


Fig. 4. Remained error of r^P and θ_i

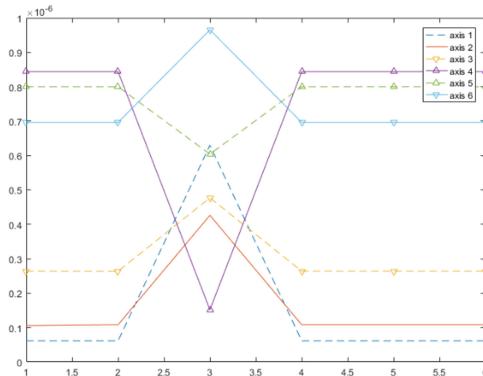


Fig. 5. Error of 6 axis length after calibration

Acknowledgments. This study was supported by ‘Study on Key technologies of Prarallel Robot for Minimally Invasive Spine Surgery’, Scientific Research Project of Shanghai Municipal Science and Technology Commission, (Projection No. 16090503700).

References

1. Besnard S, Khalil W (2001) Identifiable parameters for parallel robots kinematic calibration. In: Proceedings 2001 ICRA, IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164), vol 3. IEEE, pp 2859–2866
2. Yao R, Zhu W, Huang P (2013) Accuracy analysis of Stewart platform based on interval analysis method. Chin J Mech Eng 26(1):29–34

3. Khalil W, Besnard S (1999) Self calibration of Stewart-Gough parallel robots without extra sensors. *IEEE Trans Robot Autom* 15(6):1116–1121
4. Borm JH, Meng CH (1991) Determination of optimal measurement configurations for robot calibration based on observability measure. *Int J Robot Res* 10(1):51–63
5. Tsai LW (2000) Solving the inverse dynamics of a Stewart-Gough manipulator by the principle of virtual work. *J Mech Des* 122(1):3–9



Image of Burden Point Cloud Based on Kmeans-Bayesian Segmentation with Energy Estimation

Jiangxue Wu, Xianzhong Chen^(✉), Qingwen Hou, Yang Liu,
and Yue Meng

University of Science and Technology, Beijing 100083, China
cxz@ustb.edu.cn

Abstract. Harsh environment inside blast furnace (BF) makes it hard to monitoring burden surface shape with heterogeneous fluidization characteristic, which is an important factor affecting the production efficiency and safety of the iron-making process. In this paper, based on the imaging principle of remote sensing Synthetic Aperture Radar (SAR) radar, we design the industrial scanning radar, multiplying the density of the sample points of the surface. By analyzing the characteristics of spectrum of BF radial echo signal, this paper proposes a new intelligent image algorithm that improves the Kmeans segmentation rectified by the minimum error Bayes decision (MEBD) to extract a strip of surface point cloud. Energy centrobaric correction method is used to estimate the distance frequency of the blast surface and the results of it are used to reconstruct 3D burden surface model. Compared with other typical methods with measured data, the presented method is verified to be more effective and robust.

Keywords: Burden surface · Point cloud · Kmeans · Minimum error Bayes decision · Energy centrobaric correction method

1 Introduction

Real-time monitoring of BF burden surface distribution plays a significant role in the energy efficiency and safety of steel industry. However, the harsh environment inside BF, such as high pressure and temperature, heavy dust, strong corrosion, and the inhomogeneous gas-solid fluidization characteristic of rough burden surface complicate observation system. Gui et al. designed a parallel low-light-loss backlight high-temperature industrial endoscope to observe burden surface image [1], but its fixed viewing angle made the surface image lack depth information. Chen developed a kind of frequency modulated continuous wave (FMCW) industrial high temperature radar, which was based on pointwise measurement method to achieve continuous monitoring of the burden surface height in industrial field [2, 3].

Radar echo signal can reflect inner state of BF, but contains complex noises, such as multiple physical and electromagnetic disturbances from BF and microwave near-field antenna corrosion and adhesion, Rayleigh clutter and fluidization degree of the material surface, and gas flow distribution in core area. Especially, when the material surface presents fluidized or even sprayed, the measured signal fluctuates drastically.

Considering Hawking-Penrose singularity theorem conditions [4], the traditional peak-seeking method would generate incorrect shape of the material surface and also overlook other information.

Two kinds of signal processing methods are applied on radar signals. The one improves signals to noise ratio by filtering out interference signals. W. Liang et al. used signal bandpass filtering, time domain decomposition and other operations as denoising methods and improve the effective signal components [5]. Li et al. proposed an automatic target detection algorithm based on discrete wavelet transform (DWT) to suppress radar signal noise and enhance target information [6]. But they are weak in generalization ability and overlook continuity and correlation of radar signal in time dimension. The other focuses on correction of radar signals. Wei et al. developed a modified Kalman filter and anomaly detection model to coordinate the tracking of the blast furnace charge depth [7]. Zhu et al. proposed a segment integrated correlation detection (SICD) method for wide-band radar moving target detection [8]. But they usually characterize for poor adaptability and complex calculation.

Simulated the SAR imaging technology, this paper designs a continuous swing radar that can obtain radial high-density continuous point cloud echo signal of surface. In this paper, firstly, we analyze the burden surface echo signal spectrum in detail. Secondly, we design an intelligent segmentation based on Kmeans which separates background and foreground firstly, and then judges whether there is fixed interference in foreground, and furtherly removes the fixed interference by MEBA based on K-means. After this step, in light of the complex situation of the high and low energy of the material in the blast furnace, the energy center of gravity principle is used to extract energy points of the burden surface and the 3D point cloud model is generated. At last, we verify the performance of the algorithm with measured signals of some steels and compare segmentation and burden line results with other methods.

2 Background

Radar system installation and distance diagram is shown in Fig. 1. The industrial SAR scanning system obtains high density sequential point cloud of blast surface by continuous swing Frequency Modulated Carrier Wave (FMCW) radar with ranging resolution K is 0.0915 [9]. A single-sweep of radar along the radius direction of BF burden surface samples a set of N burden surface points. And the system operates time domain signal with Fast Fourier Transform (FFT) into frequency domain for convenience of analysis and processing.

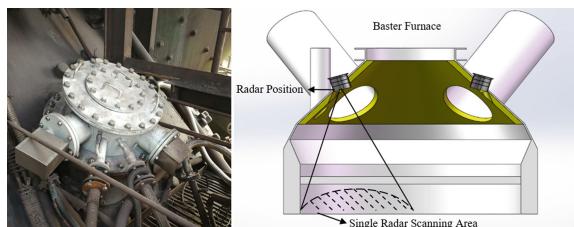


Fig. 1. Radar installation and distance diagram.

Due to burden surface distance in limited range, the first 120 spectrum numbers represent all valid information. We generate an image matrix $A_{120 \times N}$. X_i is denoted as spectrum data of sample points. Then we project $A_{120 \times N}$ into standard 64-order image, defined as 2D spectrum of radial echo signals of blast surface and shown in (a) and (c) of Fig. 2(b) and (d) of Fig. 2 are the radial echo signal that projected into sector with absolute distance by introducing angle information.

$$A = [X_1 \ X_2 \ \cdots \ X_N] \quad (1)$$

$$X_i = [x_{1,i} \ x_{2,i} \ \cdots \ x_{120,i}]^T, \ i = 1, \dots, 120 \quad (2)$$

Figure 2 shows marked difference between BF burden surface and noises. The shape of surface is a fluctuant and vague transitional belt, which is caused by its unsteady rough fluidization interface characteristics and splashing particles. Traditional peak-seeking methods fail to find burden surface distance. The foreground value is usually larger and usually contain target and fixed jamming information while the background value is usually smaller and represent random jamming information. To our dataset, there are 3 problems that require the method with intelligence and strong generalization. First, fixed jamming shows in the foreground of low SNR radar echo signal. Second, the fixed value is usually approximate to burden surface value, or even larger, which needs position feature to judge Third, the position of noise and object, and distance between them are uncertain, and sometimes noise is quite close to object region.

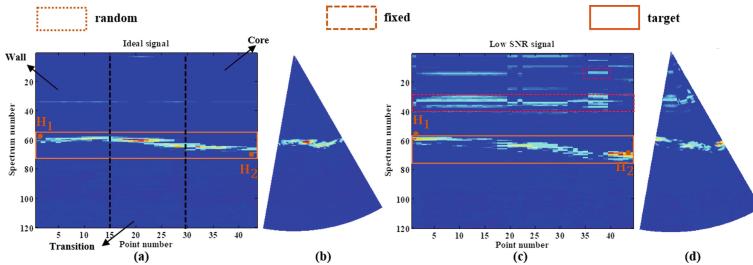


Fig. 2. Comparison of 2D original spectrum of radial echo signal of blast surface

3 Burden Surface Region Segmentation and Burden Line Extraction

This section is aim at extracting real target burden surface belt from random and fixed jamming region and extract burden line. This paper designs an algorithm based on K-means to obtain foreground, and then using coordination information to judge whether it contains fixed noise. When the fixed noise is detected, we use improved K-means based on EMBD to find the real burden surface region. After that, Energy Centrobaric Correction method is taken to select points that best represent the burden surface and Gaussian function is adopted to fitting a burden line.

Since the burden line distance is greater than 4 m and there are many noises in A , the first 40 rows in A can be removed, and then we need filter its inside salt-and-pepper noise by Median Filtering, defined as $F_m(\cdot)$, and enhance edge details by Sobel Operator, defined as $F_s(\cdot)$. Then, we obtain B which can be expressed as:

$$B = F_s[F_m(A)] \quad (3)$$

3.1 Foreground Segmentation and Judgement

The Kmeans [10] theory performances well in following aspect: center cluster control, fast computation speed and simple interpretability. Based on the principle of Kmeans, we take every element in B as input, and set $K = 2$. But there can be some spots in B after processed, which needs to remove small areas. K-mean calculation is shown by Algorithm 1.

Algorithm 1 Kmeans

Input: n object and number of clusters K

Output: Cluster $C_i (i=1, \dots, K)$

- 1: Initialization of each cluster sample center
 - 2: Assigning data points into cluster according to optimization criterion function
 - 3: Updating the cluster center
 - 4: Repeating and testing condition to stop
-

After removing background, we use the morphological character of foreground to detect whether it has fixed noise. According to features of burden line radar echo signal in Fig. 3, the spectrum can be divided into three parts equally: the furnace wall region where the burden line appears as a thick horizontal straight line with strong and concentrated energy, the transition region where the line descends slowly from the wall to the core, and the furnace core region where the shape of line is uncertain with low and dispersed echo energy. However, the fixed noise appears as a clear horizontal straight line in the whole area above the burden line. So we can take the difference between the spectral line number of point H_1 and H_2 as a basis of judging whether there is interference or not. With prior knowledge, the critical value of $|\Delta H_1 H_2|$ is usually set as 3. If $|\Delta H_1 H_2|$ is smaller than 3, there is fixed noise in foreground, otherwise there is only burden surface.

3.2 Disposal of Fixed Jamming by Minimum Error Bayes Decision Based on Kmeans

In this section, we design a segmentation method by combining minimum error Bayes decision [11] with K-means to remove fixed jamming in foreground. Figure 3 shows the typical distribution of sum energy value per row in different regions. The jamming noise region energy is lower than that in burden surface region, and there are three peaks of which the first one belongs to noise and others are in burden surface region.

But there are usually two peaks in a single region, and the gap between noise and burden surface in each region are different.

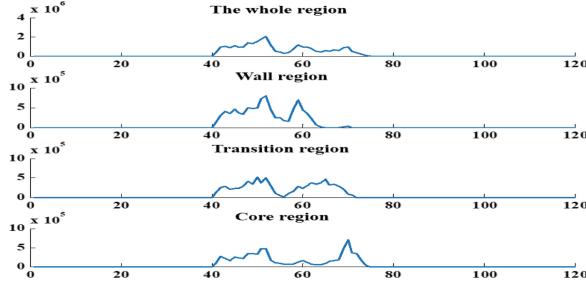


Fig. 3. Distribution of sum energy value per row in different regions.

According to these features, this paper removes noise by regions, and defines sum energy value of per row as S , the feature vector. By inputting S into the Kmeans and setting $K = 2$, the pre-classification result, the row numbers of fixed jamming are C_1 while the class of the row numbers of material surface are C_2 , it can be obtained.

$$S = [s_1, s_2, \dots, s_n]^T \quad (4)$$

Because Kmeans is sensitive to noise and highly depends on the initialization of the cluster center, there will be excessive segmentation. To reduce the error rate, we introduce minimum error Bayes decision. In binary classification, error rate is supposed to be as small as possible [12, 13]. MEDB can be the rule of the minimum error. The probability of state of cluster can be estimated. The priori probability is known as $P(C_1)$, $P(C_2)$ and the sum is 1. The feature d-dimensional observation vector is $x = [x_1, x_2, \dots, x_d]^T$ and $p(x|C_i)$ is the conditional probability of the cluster x that obtained from observing the state of cluster C_i . Bayes function is the formula (5). The obtained conditional probability $p(C_i|x)$ is called posterior probability of state. If state of each cluster obeys normal distribution, each probability is formula (6).

$$P(C_i|x) = \frac{p(x|C_i)P(C_i)}{\sum_{j=1}^2 p(x|C_j)P(C_j)} \quad (5)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6)$$

in which μ is mathematical expectation of normal distribution, $\mu = \bar{x}$ can be thought approximation and σ^2 is variance value of normal distribution and can be thought approximation to the formula (7). $p(x|C_i)$ can be defined as formula (8). We can make the decision that: If $p(C_i|x) > p(C_j|x)$, $i \neq j$, x belong to C_i . When x_0 satisfies formula

(9), the error rate is minimized. Assuming the conditional probability density distribution sum energy value per row of two classes obey normal distribution, $P(C_i)$ can be estimated by the ratio of C_i and parameters μ and σ^2 can be estimated by the pre-classification result. Then, we weighed average result of MEDB and Kmeans as the final result, which realizes the isolation of burden surface belt.

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (7)$$

$$p(x|\omega_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}} \quad (8)$$

$$p(C_i|x_0) = p(C_j|x_0) \quad (9)$$

3.3 Energy Centrobaric Correction Method for Burden Line Extraction

Due to the gushing of particles above the material surface, and under the influence of Gaussian white noise and shot noise, even after filtering and sharpening peak ridge processing, some interference signals are existed around the actual position of the material surface. Consequently, this paper processes the material surface point cloud according to the energy centrobaric correction [14] to extract the material line information and filter out the redundant point cloud data.

The energy centrobaric correction method is suitable for spectrum correction, and the effect is good when $n = 3$. In the context of Gaussian white noise, the power spectrum satisfies the following distribution. Under the influence of Gauss white noise, the normalized frequency of the energy centrobaric correction satisfies:

$$P_k = S_w^2(k) \sim N(X_w^2(k), \frac{2X_w^2(k)P_w\sigma^2}{N}) \quad (10)$$

$$\hat{f}_{k_r}^1 = \frac{\sum_{i=-n}^n (k_r + i)P_{k_r+1}}{\sum_{i=-n}^n P_{k_r+1}}, n = \infty \quad (11)$$

$$k_0 = k_1 + \frac{\sum_{i=-3}^3 i \cdot P_{k_1+i}}{\sum_{i=-3}^3 P_{k_1+i}} \quad (12)$$

Where $k_r + 1$ is the maximum spectral line of power spectrum; P_{k_r+i} is the $k_r + 1$ power spectral line. Therefore, the central coordinate k_0 of the main lobe is (k_1 is the maximum spectral line). After extracting the main energy point from the point cloud data of single point material surface by the energy centrobaric correction method, Gaussian fitting is carried out to extract a smooth fluctuating material line. After eliminating outliers according to the energy centrobaric theory, the point cloud data needs to be fitted to obtain an ideal smooth material line.

4 Experiments

This paper takes the measured echo signal from No.3 BF in Nan steel plant to validate and analyze the model performance. The whole process is shown in Algorithm 2. Figure 4 shows comparison of effects between our KBA method and the Variance and Iterative Threshold Filtering (VITF) [15] for four low quality signals. In a conclusion, the KBA filtered many kinds of interference and kept the burden surface area more intact

Algorithm 2 Burden surface visualization

Input: A 2D spectrum of radial echo signal

Output: Burden surface visualization

1: Radar signal data pre-processing

2: Foreground segmentation with K-means

3: Judgment the existence of fixed:

If $|\Delta H_1 H_2| > 3$ then

4: Isolation of burden surface belt with KBA

5: Extraction of burden line based on energy centrobaric correction and Gaussian fitting method

6: Reconstruction of 3D burden surface

In Fig. 5, we compared the burden line points extracted by our method with two other methods. The one is Single-Tone algorithm that is a frequency estimator based on Maximum likelihood estimator and Weighted average estimator [16]. The other is VITF. The results show that our method can filter outliers automatically, and the points extracted are smoother and more closer to material surface shape. According to these point cloud data, the 3D point cloud and smooth burden surface are reconstructed. The 3D model is more close to the real condition and directly reflects the fluidization character of burden surface. The results of ideal signal and low SNR signal are shown in Fig. 6. By randomly selecting 800 groups of scanning signals from April and August 2018 No. 3 BF of Nan-Steel, and comparing with the original signal spectrum, the accuracy rate of material surface extraction by imaging algorithm in this paper is calculated. The extraction accuracy rate reaches 82.5%, which overcomes the deficiency of anti-interference ability of traditional signal processing methods.

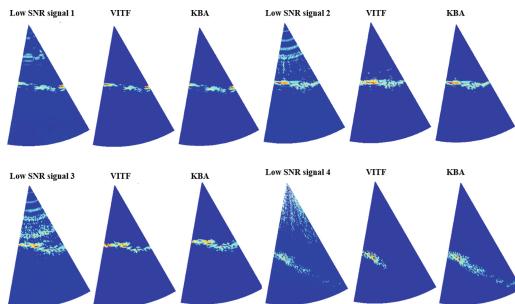


Fig. 4. Two filtering model effects for four low SNR signals

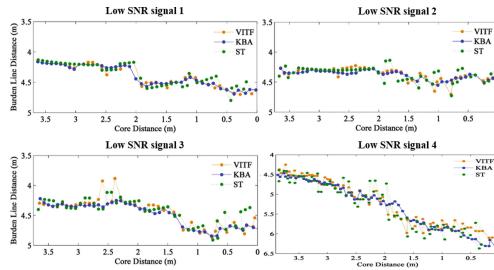


Fig. 5. Three burden line extraction methods effects for four low SNR signals

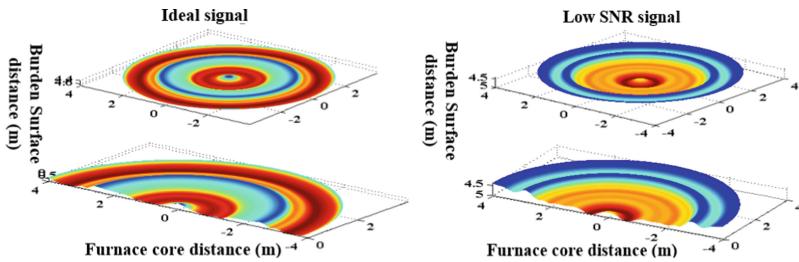


Fig. 6. Imaging effect of 3D surface point cloud

5 Conclusion

Based on the imaging principle of remote sensing SAR, this paper analyses the characters of measured signal in detail and designs the Kmeans with minimum error Bayes decision algorithm (KBA) and energy estimate to extract the burden surface point cloud, which can deal with different quality signals intelligently and accurately, and especially performance well to low-SNR signals. The method in this paper has been applied into Nan-Steel practical production and is significant to improve the efficiency and safety of BF. Because of the limited life of industrial radar antenna, there will be a large number of low SNR signals when the antenna is polluted at the end of its life. At this time, the image features are blurred and the phenomenon of missing and submerged is aggravated. At present, more intelligent image extraction methods and signal processing algorithms are not perfect, so this kind of algorithm is the focus of future research work and needs to be followed up.

Acknowledgments. This research was partially supported by the National Natural Science Foundation of China (NSFC Grant No. 61671054 and No. 61333002), Beijing Municipal Natural Science Foundation (Grant No. 4182038) and the China Scholarship Council (CSC No. 201606460036).

References

1. Chen ZP, Jiang ZH, Gui WH et al (2016) A novel device for optical imaging of blast furnace burden surface: parallel low-light-loss backlight high-temperature industrial endoscope. *IEEE Sens J* 16(17):6703–6717
2. Chen XZ, Liu FM, Hou QW, et al (2009) Industrial high-temperature radar and imaging technology in blast furnace burden distribution monitoring process. In: International conference on electronic measurement & instruments IEEE, vol 1, pp 621–625
3. Clerk Maxwell J (1982) A treatise on electricity and magnetism, vol 2, 3rd edn. Clarendon, Oxford, pp 68–73
4. William HS, Penrose R (1970) The singularities of gravitational collapse and cosmology. *Proc R Soc A* 314(1519):529–548
5. Liang W, Wang Y (2018) A new signal processing algorithm for ultra-wideband radar life detection. In: Progress in electromagnetics research symposium-spring. IEEE, pp 560–564
6. Li QZ, Zhang WD, Li M (2017) Automatic detection of ship targets based on wavelet transform for HF surface wavelet radar. *IEEE Geosci Remote Sens Lett* 14(5):714–718
7. Wei JD, Chen XZ, Kelly JR et al (2015) Blast furnace stockline measurement using radar. *Ironmak Steelmak* 42(7):533–541
8. Zhu H, Chen Y, Wang N (2015) A novel method of wideband radar signal detection. In: International congress on image and signal processing. IEEE, pp 847–851
9. Qingwen H, Xianzhong C, Xiaopan W, Yixin Y, Xiaoli L (2010) Improved FMCW signal weighted compensated correction phase difference method. *J Instrum* 04:721–726
10. Hartigan JA, Wong MA (1979) Algorithm as 136: a k-means clustering algorithm. *J R Stat Soc C* 28(1):100–108
11. Juang BH, Katagiri S (1992) Discriminative learning for minimum error classification [pattern recognition]. *IEEE Trans Sig Process* 40(12):3043–3054
12. Bao X, Wang Y (2006) Apple image segmentation based on the minimum error Bayes decision. *Trans CS AE* 22(5):122–124
13. Kurihara K, Welling M (2009) Bayesian k-Means as a “maximization-expectation” algorithm. *Neural Comput* 21(4):1145–1172
14. Ding K, Zheng C, Yang Z (2010) Precision analysis and improvement of frequency correction by the center of gravity method of discrete spectrum energy. *J Mech Eng* 46(5):43–48
15. Wang Q, Wu J, Hou Q, Chen X (2019) Sharpness Image of burden point cloud based on cramer-rao lower bounds unbiased energy estimation. *Acta Automatica Sinica*, 1–9. <https://doi.org/10.16383/j.aas.c180683>
16. Zhivomirov H, Nedelchev I, Vasilev R (2016) A method for single-tone frequency estimation. *Rom J Acoust Vib*



An Efficient Loop Closure Detection Based on Convolutional Autoencoder Neural Network for Autonomous Robots

Jian Wang¹ and Sen Zhang^{2(✉)}

¹ School of Automation and Electrical Engineering,
University of Science and Technology Beijing, Beijing 100083, China
² 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
zhangsen@ustb.edu.cn

Abstract. Loop closure detection is the core part of the visual simultaneous localization and mapping (VSLAM) for autonomous robots. In the dynamic environments, loop closure detection turns to be a very difficult problem compared to the static scenario. This paper proposes a novel approach based on convolutional autoencoder neural network (CAENN) architecture for extracting images features, and then uses Euclidean loss for minimizing the difference between the extracted feature and the gist feature which has the ability of scene recognition. In order to improve the accuracy and recall rate for loop closure detection in the dynamic scenario, the perspective transformation and dynamic object are applied in the process of the construction of the training set. And by calculating the Manhattan distance between the two image feature vectors, the loop closure is accepted when the distance is smaller than the threshold. The experimental results demonstrate that the proposed method obtains better accuracy and recall rate compared to the commonly used gist feature method and has a lower cost on time and space compared to the BOW method and AlexNet method in loop closure detection.

Keywords: Visual simultaneous localization and mapping · Loop closure detection · Convolutional autoencoder neural network · Gist feature

1 Introduction

The solution to the VSLAM problem is the basis for mobile robots to achieve autonomous navigation in large-scale, unstructured, dynamic, and unknown real-world environments [1], which is also the key to achieving autonomous location. A very well-known VSLAM framework, ORB-SLAM [2, 3] decomposes VSLAM into three threads, including tracking, mapping, and loop closure detection. Loop closure detection is an application of scene recognition in VSLAM [4]. This technology not only allows a robot to detect whether it has returned to a previously visited place, but also provides a retargeting application for mobile robot tracking failures.

At present, the widely used loop detection method is called the bag of words (BOW) [5] method, whose two implementations are best known as FAB-MAP2.0 [6] and DBOW [7]. This method firstly needs to extract the local hand-crafted features, such as SIFT [8], SURF [9] and ORB [10], and then classify these features to create a visual feature dictionary, which is mainly used to accelerate the feature searching and matching. Finally, the feature extracted from the image is described as a word vector, and the similarity between different images is judged by comparing the distance of word vectors of different images. However, the hand-crafted features, which may not really represent the image itself. On the one hand, the existence of the visual dictionary accelerates the feature searching and matching, and on the other hand, it is affected by the residual from the different descriptors assigned to the same word, which would make the mismatch of different features.

The BOW method needs to extract local features (including finding key points and calculating descriptor of key points). These hand-crafted local descriptors may not represent the essential information of the image, and deep learning method has been proven to be better used to characterize images than the traditional methods. Shin et al. [11] uses the DCGAN network to learn local patches centred on the SURF feature, each local patch is characterized by a 512-dimensional vector which aims to replace the SURF descriptor as a word vector. Hou et al. [12] handed over feature extraction and feature description to the CNN model, and verified the validity of the CONV3 and POOL5 features of the trained CNN model for the loop closure detection task. Xia et al. [13] compared the effects of current popular PCANet, CaffeNet, AlexNet, and Google LeNet deep neural networks as feature extractors in loop detection. However, these models were not originally designed for loop closure detection purposes, so some people propose their own trained models for loop closure detection tasks. Gao et al. [14] designed a modified stacked denoising autoencoder architecture and showed a better result than FAB- MAP2.0. Inspired by VLAD, Arandjelovic et al. [15] added a new VLAD layer behind the CNN neural network and built the NetVLAD architecture. However, its training data needs to be manually labelled, and each image needs to have a set of positives and negatives. Liu et al. [16] designed an end-to-end siamese convolutional neural network to directly obtain the similarity between two different images of the input network, but it is highly dependent on the environment and needs to be pre-trained in the actual scene.

The gist feature [17] is designed as a global image descriptor in scene recognition tasks, and in present some people have applied it to the loop closure detection [18]. There are five spatial envelop properties used in the gist feature including the degree of naturalness, the degree of openness, the degree of the roughness, the degree of expansion, and the degree of ruggedness, but the original purpose of designing the gist feature is used in scene recognition which is a classification problem. And in the loop closure detection task, the performance of the gist feature is lower than the method based on BOW method. Therefore, we design a convolution autoencoder neural network to learn the gist feature and fine-tune the model with a homemade loop closure training set to optimize its performance.

Therefore, we attempt to solve these problems by proposing a new approach in the following of the paper and help robots more effectively and efficiently to realize loop

closure detection. We make the following contributions: (1) We propose a model based on CAENN which can achieve the vectorized representation of images quickly and usefully. (2) We propose a method to construct a training dataset which consists of paired image like a real loop closure. (3) Compared to the gist feature method, our method gets a better performance. And compared to the BOW method and AlexNet method, our method has a lower cost on time and space.

The paper is organized as the follows. In Sect. 2, the overall framework is introduced. In Sect. 3, the process of training set preparation is introduced with two image processing method and the pipeline of training set preparation. In Sect. 4, we present the proposed method on the aspect of the network architecture, and the usage for loop closure detection. Section 5 shows experimental results on three datasets to compare the performance between our method and the state-of-the-art algorithms. And the proposed approach is tested in a real-time loop closure Scenes. Finally, we conclude the paper in Sect. 6.

2 Overall Framework

The essence of loop closure detection is to compare the similarity between two scenes. The stable feature vector of the image scene is obtained by feature extracting method, then the distance between the feature vectors is calculated to determine whether it is a loop closure. In this paper, a neural network model based on CAENN is designed for feature extraction of images. One of the most difficult things for deep learning is obtaining a suitable training set. As far as we know, there is no training set for loop closure detection, therefore, two methods are used to build a loop closure training set, including perspective transformation processing and image fusion processing. By using these methods, a paired image which looks like a loop closure could be generated from an image, just like the left side of the Fig. 1. And then one of the paired images is input into the proposed neural network, and the other is encoded into a gist feature as the label. The reason for designing this is that the proposed network can easily encode an image with spatial structure feature by learning the gist feature which can realize the compressed quantization coding of an image. Finally, the image captured by the robot is transformed into a descriptor vector through the proposed network. By calculating the similarity between the different descriptor vectors, it is determined whether the scene passed by the robot is a loop closure.

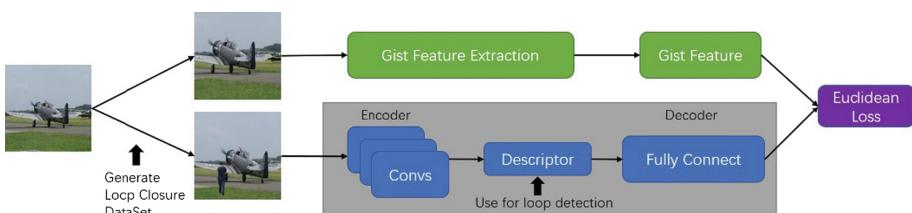


Fig. 1. The framework of the proposed approach.

3 Training Data Preparation

In this section, two image processing methods used in the proposed scheme are described for generating paired images of the original image. The two methods include perspective transformation processing and image fusion processing which are proposed in Subsects. 3.1 and 3.2 respectively.

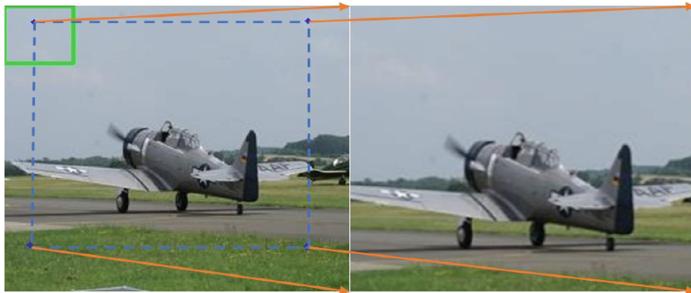


Fig. 2. The sketch map of perspective transformation processing

3.1 Perspective Transformation Processing

For a given image I , in order to generate the perspective transformation image I_{pp} paired with it, firstly, the homography matrix $H_p \in R^{3 \times 3}$ for projection transformation is acquired, and then apply Eq. 1 to calculate the transformed image.

$$I_{pp} = H_p I$$

$$\text{Where } I_{pp} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}, H_p = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}, I = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

To solve the matrix H_p , it is assumed $H_{33} = 1$ because the homography matrix is a reversible matrix with rank 3. Thus, we can get

$$\begin{aligned} x'(h_{31}x + h_{32}y + 1) &= h_{11}x + h_{12}y + h_{13} \\ y'(h_{31}x + h_{32}y + 1) &= h_{21}x + h_{22}y + h_{23} \end{aligned} \quad (2)$$

Equation 2 can be written as

$$\begin{aligned} h_{11}x + h_{12}y + h_{13} - h_{31}xx' - h_{32}yx' &= x' \\ h_{21}x + h_{22}y + h_{23} - h_{31}xy' - h_{32}yy' &= y' \end{aligned} \quad (3)$$

From Eq. 3, we can know that it is a least square problem and matrix H_p need 4 groups of (x, y) value to solve. Therefore, we set a border of $H/5 \times W/5$ at the upper

left of the image, where H is the height of the image and W is the width of the image, just like the green corner in left image of the Fig. 2. And then we randomly select 1 value of (x,y) just like the blue point in the green border, and calculate the other three points to ensure that the length and width of the new image is 80% of the original image by taking the blue point in the green border as reference. By using this method, the intersection over Union (IOU) of the new images generated from a same image is between 40% and 100%. And the loop closure also has a little rotation, to this end, we add a 3 pixels randomness for each point to simulate this rotation, just like the blue point in the left image. Finally, the homography matrix H_p is calculated through the four pairs value. With the raw Image I and the homography matrix H_p , the image with perspective transformation I_{pp} is obtained by using Eq. 1.

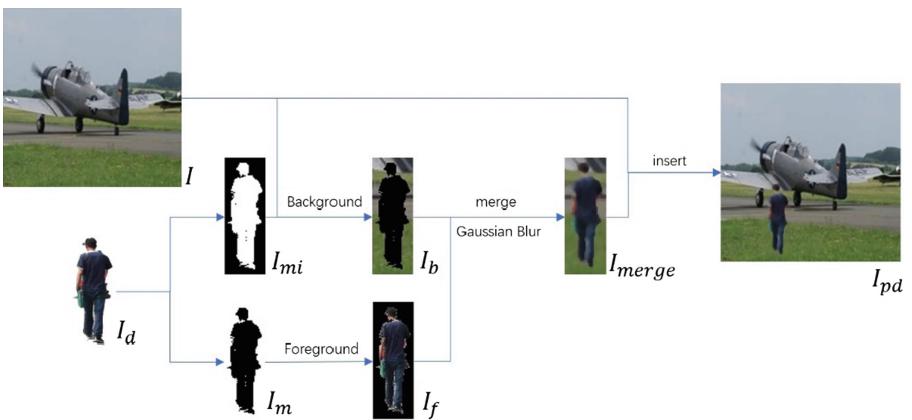


Fig. 3. The procedure of image fusion processing

3.2 Image Fusion Processing

For a given image I , in order to generate an image I_{pd} that is paired with the dynamic object, the dynamic objects to be joined by randomly selecting some images including pedestrians and vehicles were download from the Internet. Then pedestrians and vehicles were extracted from the background via photoshop. When directly adding a dynamic object image with a solid color background to the original image, a rectangular box would be added on each generated image, which do not match the actual scene of the loop closure. Therefore, the binarization method is used to generate the mask image I_m and the inverted mask image I_{mi} of the dynamic object image I_d . Foreground image I_f is obtained by using bitwise and operation between Image I_d and mask image I_m , so does the background image I_b with image I and mask image I_{mi} , and then the fusion image I_{merge} is obtained by merging the foreground image I_f and background images I_b with adding Gaussian blur to reduce the influence of the fusion edge. Finally, the image I_{pd} is generated by putting the image I_{merge} and the raw image I together. All procedure is shown as Fig. 3. Of course, the position and size of the dynamic object are randomly generated within a certain range. Considering that the

position of the dynamic object in the loop detection does not generally exceed 2/3 of the image height, and the size does not exceed half of the image area, so the reasonable limits on the size and location of dynamic objects were made.

4 Loop Closure Detection Implementation

In this section, we present the architecture and the implementation details of the proposed neural network based on CAENN. The proposed network is designed to compress the image into a feature vector which saves the scene information of the image. And the loop closure detection is calculated by the similarity between different feature vectors.

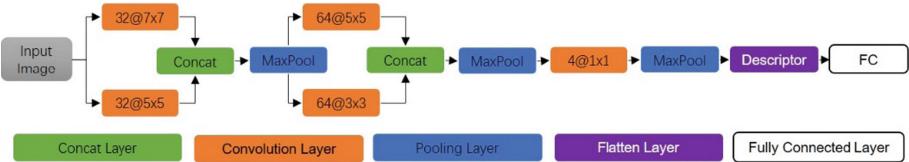


Fig. 4. The architecture of the proposed network.

4.1 Network Architecture

The CAENN designed in this paper is introduced in detail in Fig. 4. It consists of five convolution layers, two concat layers, three pooling layers, one flatten layer and one fully-connected layer. Firstly, in order to ensure that the output dimension of each layer is fixed, the image should be transformed into grey scale and scaled to 256×256 before it is input into the network. Secondly, to extract the information of different scales of the image, two convolutional layers are used with different kernel size which are 7×7 and 5×5 . Then rectified linear unit (ReLU) activation function is used to add the non-linear ability of our model, and concat layer is used to integrate the output of the ReLU layer from above. The local-response-normalization (LRN) layer which is following the max-pooling layer is added to increase the generalization ability of the model. The next part is similar to the described above, except the convolution cores are replaced by 5×5 and 3×3 . The aim of the third part which is consist of four convolution cores of 1×1 size is to reduce the feature depth, and an image feature description is obtained from the flatten layer. The fully-connected layer with sigmoid activation is used for reconstructing the gist feature extracted from the paired image. And the l_2 loss function is chosen for minimizing the distance between the output of our model and the gist feature descriptor because of its better performance in scene recognition [17].

4.2 Detecting Loops

Loop closure detection is to verify the similarity of two images captured at different times. If the similarity is high, it is considered to be a loop. Therefore, we estimate

whether there is a loop closure by calculating the distance between the features of the Flatten layer obtained by CAENN from different frames.

We define the feature vector generated of the i -th image as V_p^i , and use the Manhattan distance to calculate the similarity of the two frames, as the Eq. 4 shown below:

$$D(i,j) = \left\| V_p^i - V_p^j \right\|_1 \quad (4)$$

In practical applications, if the similarity of two frames of images is lower than a threshold, it can be considered as a loop closure.

4.3 Implementation Details of Network Training

For training sets, we select MIT's Places365 standard dataset [19], which includes approximately 1.8 million images from 365 scenes, each of which has approximately 5,000 images. The loop closure dataset used in this paper is generated by using the above proposed method. Inspired by Merrill et al. [20], the gist feature of the image can be extracted during the process of making up the training set. In order to construct and train the model, we chose the caffe deep learning framework which is widely used in the field of image processing [21]. The weight initialization of the convolutional layer and the fully connected layer is Xavier method, and the deviation is initialized to constant. The specific hyperparameter configuration is shown in the Table 1.

Table 1. Hyper-parameters for the model training.

| Parameter | Value | Parameter | Value |
|--------------|-------|--------------|--------|
| Batch size | 256 | Solver type | SGD |
| Input height | 256 | Input weight | 256 |
| Lr_policy | Fixed | Base_lr | 0.001 |
| Momentum | 0.9 | Weight_decay | 0.0005 |

5 Experimental Results

In this section, the proposed CAENN method is compared with the DBOW2 algorithm which is the current loop closure detection benchmark from ORB-SLAM2 [3], the deep learning network AlexNet model [22] trained in the same dataset, and the gist feature which is the learning goal of our method. The comparative dataset used in this paper is the New College and City Centre datasets that are currently popular open source dataset in loop closure detection and our own recorded datasets.

5.1 Evaluation Criteria

At present, the performance of the loop closure detection is mainly achieved by drawing a precision-recall curve. The precision rate is the probability that all loop

closures detected in the algorithm are actual loop closures. The recall rate is the probability of being correctly retrieved in all real loops. Adjusting with the different thresholds, it is easily to get the corresponding accuracy and recall rate, after that these precision rates and recall rates are drawn into a chart to get the precision-recall curve. There are also many ways to interpret the precision-recall curve. Most researchers use the following two indicators to explain the precision-recall curve, including the area under curve (AUC) value, located between [0, 1], where the AUC value is the higher the better; and the maximum recall rate when the precision is 100%, which is also the bigger the better. By reading these values, it is easily distinguishing the performance of difference algorithm.

5.2 Experiment on Open Source Dataset

The City Centre dataset and the New College dataset [23] are both provided by the Oxford University Mobile Robotics Group. There is an example of the two datasets in the Fig. 5. The GPS information and the ground truth of the loop closure detection is also provided. The ground truth is specified as a binary correspondence matrix whose $Entry(i, j)$ is 1 if image i and image j were determined to be taken at the same place, and 0 otherwise. It is very different between the images captured by left and right camera which is influenced by the orientation, so the images captured by the left camera were selected to evaluate the algorithm performance. The precision-recall curve is plotted as shown in Fig. 6. As we can see, our method performs better than the gist feature method that is the learning object for our model, which means the modification of the original image in the training set is efficient. And our method gets higher value on the precision rate when the recall rate is 100% than the other methods. For the AUC value, which is as shown in Fig. 2, our method performs as good as AlexNet method on city centre dataset, but lower than DBOW method on New College dataset. By comparing the two datasets, it is found that there are much more pedestrians and vehicles in the City Centre datasets. Therefore, we consider that our method has a better performance in the scene with more dynamic objects. In order to verify it, we recorded a dataset from our campus.



Fig. 5. Examples of City Centre dataset (a), Examples of New College dataset (b).

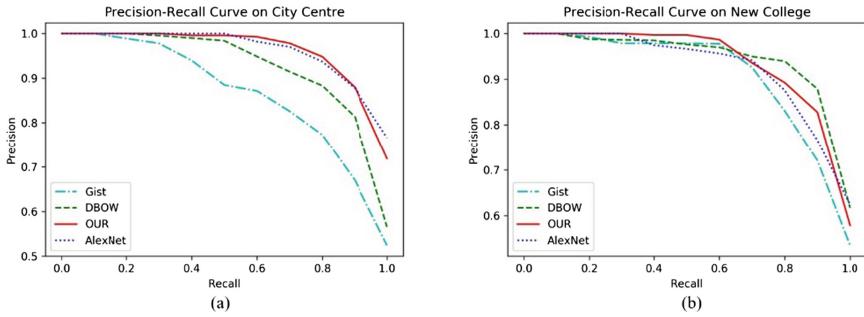


Fig. 6. The precision-recall curve on City Centre dataset (a), The precision-Recall curve on New College dataset (b).

Table 2. The area under curve (AUC) value.

| | Gist | DBOW | Ours | AlexNet |
|-------------|-------|-------|-------|---------|
| City Centre | 0.874 | 0.932 | 0.965 | 0.968 |
| New College | 0.911 | 0.952 | 0.945 | 0.927 |
| Campus | 0.873 | 0.773 | 0.948 | 0.884 |

5.3 Experiment on Our Campus

The campus dataset is recorded at the main street of our campus where there are always some people walking through it. And the sunny weather reduces the impact of the light on the recorded datasets. We hand-hold a mobile phone as a camera to record the dataset by walking the street twice, so the viewpoint of the image captured in the same place is a little different but the main difference is the position of pedestrians. Some examples are shown at the Fig. 7. The campus dataset is recorded as two 1080P, 60FPS videos, and then the videos are sampled and made pair of the image captured from the same place of the two videos into 100 image pairs. Figure 8 is shown the results of experimentation with this dataset. As expected, our method gets the better performance to the DBOW2 method even better than the AlexNet method. It is reasonable that AlexNet network does not specifically optimize the loop closure detection task than our model especially in the case of more dynamic objects.



Fig. 7. Examples of campus dataset.

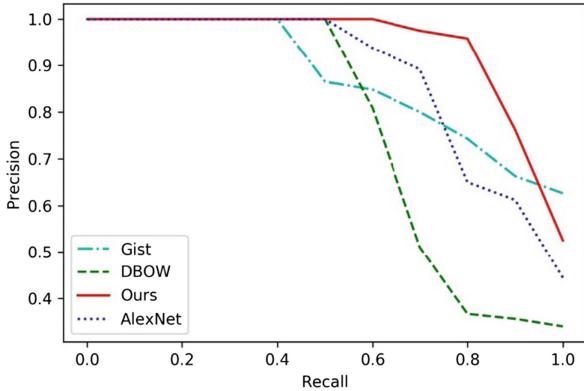


Fig. 8. The precision-recall curve on campus dataset.

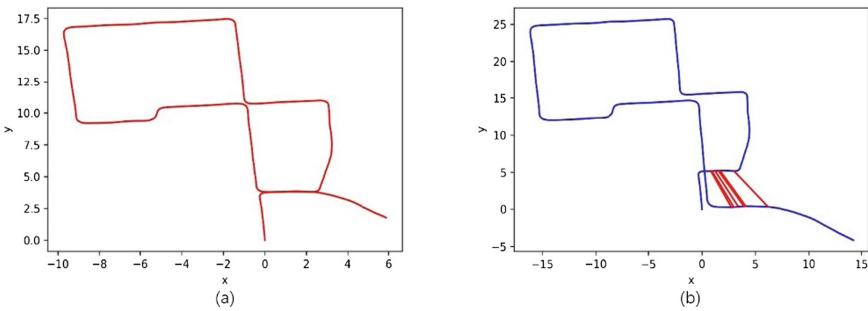


Fig. 9. The ground truth of part of the KITTI0 dataset (a). The Mono SLAM trajectory with detected loops (b).

5.4 Online Loop Closure

The Mono ORB-SLAM2 algorithm is used as the testing framework by replacing its own loop closure detection method (DBOW2) with our loop closure detection algorithm. It should be noted that we only do loop closure detection without map optimization. The online validation dataset what we choose is the KITTI dataset [24], which is the largest computer vision algorithm evaluation dataset in autopilot scenarios in the world at present. The first 1740 frames of KITTI0 dataset are intercepted which has a real loop closure scene to facilitate easily the observation of loop closure detection effect. The ground truth is shown at Fig. 9(a). As for the threshold of loop closure detection, we currently use a fixed value by experience. In the future, the similarity between the key frame and the adjacent key frame can be used as the reference of dynamic threshold. The loop closure detection by our approach is shown in the Fig. 9 (b). It can be seen that the map constructed by monocular SLAM has a large cumulative deviation due to the lack of real scale information. Our approach can not only accurately identify the loop closures in the real scene, but also has no false detection, which shows that our algorithm can be applied to the actual SLAM system.

5.5 Time and Space Consuming

The efficiency of the loop closure detection is also as important as the effective of it in the real time SLAM systems. The time consuming about the descriptor extraction for each image is counted because the similarity computation cost a little time which is much smaller than the description extraction in a not large map such as the experiment we have done and in a large map there are many ways to accelerate the speed of the similarity computation such as using hash method. The dataset used in this test is all the dataset from the above experiments. The result is the average time consuming and the running space that is an initialization state without any input, which is shown as Table 3. Note that the test is under i7-8750H CPU, and a GeForce GTX 1060 GPU.

Table 3. The average of time and space consuming.

| Model | Feature extract | | Running space | Storage space |
|---------|-----------------|---------|---------------|---------------|
| | CPU | GPU | | |
| Gist | 545 ms | N/A | N/A | N/A |
| DBOW | 10.1 ms | N/A | 234 MB | 145 MB |
| Ours | 75.35 ms | 1.66 ms | 44 MB | 14 MB |
| AlexNet | 695 ms | 10.9 ms | 236 MB | 233 MB |

From the Table 3, as you can see in the CPU mode, our method performs much quickly than AlexNet method and gist method. And in the GPU mode, our method costs only 1.66 ms, which is 5 times faster than the DBOW2 which is the fastest in CPU mode. With the development of the neural network engine on the embedded chip such as the A11 bionic from APPLE, it is easy to use deep learning method on a mobile device. Another thing should be known is that the storage of our model is only about 14 MB. The smallest running space means that there is much more memory for other algorithms in VSLAM such as odometry method and bundle adjustment method, what's more, it is possible to achieve a real VSLAM system with loop closure detection by using our model in a device with limited storage and computing resources.

6 Conclusion

This paper proposes a method based on convolution autoencoder neural network, which is effective and efficient in the VSLAM system. Our method can learn the scene recognition by using the gist feature as the learning goal. And more than this, with the training on the proposed loop closure dataset, our model gets the better performance compared to the gist feature, and performs better than DBOW2 method and the AlexNet method especially in the scene with dynamic objects. Another thing should be known is that our model only needs about 14 MB to storage and the running space is much lower than the other method compared in this paper. What's more, our model performs an average time about 1.6 ms for feature extraction of each image, which is much faster than the other methods among the paper.

Acknowledgments. This work was supported in part by the National Science Foundation of China (NSFC, No. 61673125, No. 61673055, No. 61673056, No. 61773056) and the National Key Research and Development Program of China under Grant 2017YFB1401203.

References

1. Labbe M, Michaud F (2014) Online global loop closure detection for large-scale multi-session graph-based SLAM. In: 2014 IEEE/RSJ international conference on intelligent robots and systems (IROS 2014). IEEE
2. Mur-Artal R, Montiel JMM, Tardos JD (2015) ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Trans Robot* 31(5):1147–1163
3. Mur-Artal R, Tardós JD (2017) Orb-slam2: an open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE Trans Robot* 33(5):1255–1262
4. Corke P et al (2013) Dealing with shadows: capturing intrinsic scene appearance for image-based outdoor localisation. In: 2013 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE
5. Csurka G et al (2004) Visual categorization with bags of keypoints. In: Workshop on statistical learning in computer vision. ECCV, vol 1, no 1–22
6. Cummins M, Newman P (2011) Appearance-only SLAM at large scale with FAB-MAP 2.0. *Int J Robot Res* 30(9):1100–1123
7. Gálvez-López D, Tardos JD (2012) Bags of binary words for fast place recognition in image sequences. *IEEE Trans Robot* 28(5):1188–1197
8. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110
9. Bay H, Tuytelaars T, Van Gool L (2006) Surf: speeded up robust features. In: European conference on computer vision. Springer, Heidelberg
10. Rublee E et al (2011) ORB: an efficient alternative to SIFT or SURF. In: 2011 IEEE international conference on computer vision (ICCV). IEEE
11. Shin D-W, Ho Y-S (2018) Loop closure detection in simultaneous localization and mapping using learning based local patch descriptor. *Electron Imaging* 2018(17):1–6
12. Hou Y, Zhang H, Zhou S (2015) Convolutional neural network-based image representation for visual loop closure detection. In: 2015 IEEE international conference on information and automation. IEEE
13. Xia Y et al (2017) An evaluation of deep learning in loop closure detection for visual SLAM. In: 2017 IEEE international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData). IEEE
14. Gao X, Zhang T (2015) Loop closure detection for visual slam systems using deep neural networks. In: 2015 34th Chinese control conference (CCC). IEEE
15. Arandjelovic R et al (2016) NetVLAD: CNN architecture for weakly supervised place recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition
16. Liu H et al (2018) An end-to-end siamese convolutional neural network for loop closure detection in visual slam system. In: 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE
17. Oliva A, Torralba A (2001) Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int J Comput Vis* 42(3):145–175

18. Singh G, Kosecka J (2010) Visual loop closing using gist descriptors in manhattan world. In: ICRA omnidirectional vision workshop
19. Zhou B et al (2018) Places: a 10 million image database for scene recognition. IEEE Trans Pattern Anal Mach Intell 40(6):1452–1464
20. Merrill N, Huang G (2018) Lightweight unsupervised deep loop closure. arXiv preprint [arXiv:1805.07703](https://arxiv.org/abs/1805.07703)
21. Jia Y et al (2014) Caffe: convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM international conference on Multimedia. ACM
22. Zhou, B et al (2014) Learning deep features for scene recognition using places database. In: Advances in neural information processing systems
23. Cummins M, Newman P (2008) FAB-MAP: probabilistic localization and mapping in the space of appearance. Int J Robot Res 27(6):647–665
24. Geiger A et al (2013) Vision meets robotics: the KITTI dataset. Int J Robot Res 32(11):1231–1237



Quasi-periodic J_2 Invariant Relative Orbits for Formation Flying

Qiang Luo¹, Yaru Zheng^{2(✉)}, and Ming Xu^{2(✉)}

¹ National University of Defense Technology, Changsha 410073, China

² School of Astronautics, Beijing University of Aeronautics and Astronautics,
Beijing 100191, China
xuming@buaa.edu.cn

Abstract. It is presented in this paper that an innovative concept quasi-periodic formation flying, which can hold bounded trajectories without any restriction initialization, compared with the traditional restriction on initial velocity by linear equation for relative dynamics. The improved periodic linear equation associated with J_2 perturbation is derived from the Taylor expansion of the J_2 relative dynamics and the analytic solutions of osculating orbital elements. And then it is constructed the controller preserving Hamiltonian structure just for the simplified representation of formation flying on the near-circular orbit. Finally it is investigated that the stability of the quasi-periodic J_2 invariant relative orbits generated from the equilibrium with its topology restructured from hyperbolic to elliptic by the Floquet multipliers of the controlled time-periodic Hamiltonian system.

Keywords: Quasi-periodic formation flying · Hamiltonian structure controller · Floquet multipliers

1 Introduction

There have been a lot of interests in formation flying in recent years, and many planned missions are beginning to take advantage of the benefits offered by the use of satellite formations [1, 2], such as the Gravity Recovery and Climate Experiment Mission (GRACE) developed by ESA and NASA, and the TerraSAR-X Add-On for Digital Elevation Measurement (TanDEM-X) developed by DLR (Germany). Now a large body of literature on relative dynamics and control has been led, and most of them have adopted Clohessy-Wiltshire (C-W) equation to design close periodic configurations for formation flying [3–5]. Here we will claim an innovative concept quasi-periodic formation flying, which can hold bounded trajectories without any restriction initialization, compared with the traditional restriction on initial velocity by C-W equation.

The main contributions in this paper are to derive the improved periodic C-W equation associated with J_2 perturbation (named as J_2 Type Periodic C-W Equation) derived from the Taylor expansion and the analytic solutions of osculating orbital elements, and then to construct the controller preserving Hamiltonian structure just for the simplified representation of formation flying on the near-circular orbit. Thus, the quasi-periodic formation flying is innovatively proposed which can hold bounded

trajectories without any restriction initialization. Finally, it is investigated that the stability of the quasi-periodic J_2 invariant relative orbits generated from the equilibrium with its topology restructured from hyperbolic to elliptic by the Floquet multipliers of the controlled time-periodic Hamiltonian system.

The significant outcome is to obtain quasi-periodic J_2 invariant relative orbits, which is the first time to be discovered. Quasi-periodic J_2 invariant relative orbits are quite different from J_2 invariant relative orbits that are proposed firstly by Schaub and Alfriend [6] and can be generated by the iterative algorithm from differential correction created by Xu [7]. The later requires two constraints on semi-major axis (a), eccentricity (e) and inclination (i). The bounded trajectories obtained from C-W equation are just close periodic orbit, but quasi-periodic orbits are much more universal than periodic orbits in the Cantor collection of bounded trajectories of formation flying are discovered in this paper.

2 Nonlinear Dynamics of Formation Flying

C-W equation has been modified by a number of scholars, relatively famous are: Carter [3] involved the drag resistance in C-W equation; Izzo [4] combined the long-term facts from J_2 perturbation with C-W equation and obtained the linear time-invariant model. Unlike what have been mentioned above, in this section, we derive the linear time-varying J_2 Type Periodic C-W Equation from Taylor expansion of the J_2 relative dynamics.

In inertial coordinates, the relative position vector is $\Delta\mathbf{r}_I = \mathbf{r}_{Follower} - \mathbf{r}_{Leader}$, and it can be denoted as $\Delta\mathbf{r} = [x \ y \ z]^T$ in the Leader orbit coordinates. In assumption that the transformation from inertial coordinates to the LVLH coordinate frame is

$$\mathbf{R}(\Omega, i, u) = \mathbf{R}_z(u)\mathbf{R}_x(i)\mathbf{R}_z(\Omega) \quad (1)$$

Thus, the relative position vector in the inertial coordinates can be expressed as

$$\Delta\mathbf{r}_I = \mathbf{R}^{-1} \cdot \Delta\mathbf{r} = \mathbf{F} \cdot \Delta\mathbf{r} \quad (2)$$

taking the second-order derivative of $\Delta\mathbf{r}_I$ with time comes

$$\Delta\ddot{\mathbf{r}}_I = \mathbf{F} \cdot \Delta\ddot{\mathbf{r}} + 2\dot{\mathbf{F}} \cdot \Delta\dot{\mathbf{r}} + \ddot{\mathbf{F}} \cdot \Delta\mathbf{r}. \quad (3)$$

Considering J_2 perturbation, the potential function of the Follower is

$$U|_F = -\frac{\mu}{r_F} + \frac{\mu J_2 R e^2}{2r_F^3} \left(3 \frac{z_F^2}{r_F^2} - 1 \right) \quad (4)$$

Where $r_F = \sqrt{(r+x)^2 + y^2 + z^2}$, $z_F = (r+x) \sin u \sin i + y \cos u \sin i + z \cos i$, and r, u, i are, respectively, the vector of the Leader, the argument of orbit and the orbital inclination.

By solving Eqs. (3) and (4) and eliminating φ_L , we can obtain

$$\Delta\ddot{\mathbf{r}} + 2\mathbf{F}^{-1}\dot{\mathbf{F}}\Delta\dot{\mathbf{r}} + \mathbf{F}^{-1}\ddot{\mathbf{F}}\Delta\mathbf{r} = -(\nabla \cdot \nabla_{\Delta\mathbf{r}})U_F|_{\Delta\mathbf{r}=0}\cdot\Delta\mathbf{r} \quad (5)$$

this equation is the general form of the J_2 type C-W equation.

The derivation process of the J_2 type C-W equation demonstrates that it is essentially the first-order approximation of the J_2 relative dynamics. Meanwhile, the general form is not restricted to orbital eccentricity, thus, is adapted to describe the relative motion of elliptical formation flying.

According to Eq. (5), the influences of J_2 perturbation on the relative dynamics are as following:

- (a) influence on kinematics, reflected in $\mathbf{F}^{-1}\dot{\mathbf{F}}$ and $\mathbf{F}^{-1}\ddot{\mathbf{F}}$, is that J_2 perturbation arouses the drift velocity between the orbit coordinates of the Leader and the Follower;
- (b) influence on dynamics, reflected in $-(\nabla_{\Delta\mathbf{r}} \cdot \nabla)U_F|_{\Delta\mathbf{r}=0}$, is that J_2 perturbation produces the differences between the potential functions of the Leader and the Follower.

3 Quasi-periodic Formation Flying and Quasi-periodic J_2 Invariant Relative Orbits

3.1 Structure-Preserving Stabilization for Hamiltonian System

Most astrodynamical problems have Hamiltonian structure, whose Hamiltonian function is expressed in the following form [8]:

$$H = \frac{1}{2}\mathbf{p}^T\mathbf{p} + \omega\mathbf{p}^T\mathbf{J}\mathbf{q} - V(\mathbf{q}) \quad (6)$$

and the Legendre transformation between Hamiltonian variables and physical variables is as following:

$$\begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 \\ -\omega\mathbf{J} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \dot{\mathbf{r}} \end{bmatrix} \quad (7)$$

where \mathbf{J} is the symplectic operator, and \mathbf{I} is the identical operator.

The equilibria of the Hamiltonian system can be solved from:

$$\begin{cases} \mathbf{r} = \mathbf{r}_0, \dot{\mathbf{r}} = \mathbf{0} \\ \left.\frac{\partial V}{\partial \mathbf{r}}\right|_{\mathbf{r}_0} = \mathbf{0} \end{cases} \quad (8)$$

Denote V_{rr} as the second derivative matrix of the pseudopotential function V to the position vector $\mathbf{r} = [x, y]$, and the element of V_{rr} has the general expression as

$V_{mn} = \frac{\partial V}{\partial m \partial n}$, $(m, n) = (x, y)$. And then a hyperbolic (saddle) equilibrium is resulted from $V_{xx} \cdot V_{yy} - V_{xy}^2 < 0$.

Then the characteristic equation for the linear Lagrangian dynamics is

$$\lambda^4 + B\lambda^2 + C = 0 \quad (9)$$

where $B = 4\omega^2 - V_{xx} - V_{yy}$, and $C = V_{xx} \cdot V_{yy} - V_{xy}^2$. For the hyperbolic system, the hyperbolic eigenvalues are $\pm\sigma$, and the stable\unstable manifolds are \mathbf{u}_\pm

$$\mathbf{T}_c = \{-\sigma^2 [G_1 \mathbf{u}_+ \mathbf{u}_+^T + G_2 \mathbf{u}_- \mathbf{u}_-^T] - \gamma^2 G_3 \mathbf{u} \bar{\mathbf{u}}^H\} \delta \mathbf{r} + 2\Delta \cdot \mathbf{J} \delta \dot{\mathbf{r}} \quad (10)$$

where the gains of the controller G_1, G_2, G_3

Some propositions and theorem have been achieved [8] as:

Proposition 1: The stable and unstable manifolds can be used to stabilize the system;

Proposition 2: The center manifolds can be used to stabilize the system;

Theorem of Poles Assignment: The two poles can be assigned at any different positions on the imaginary axis, and the G_1, G_2 , and G_3 required to stabilize the system are not unique.

3.2 Quasi-periodic Formation Flying Generated from Structure-Preserving Stabilization for Hamiltonian System

Most Formation-flying maintaining controllers are based on the classical the C-W equation, however, the C-W equation doesn't consider of the J_2 perturbation. Therefore when the controllers try to eliminate the drifts in the along-track direction inserted by J_2 perturbation, it will consume rather more ΔV . Knowing from the above section, the J_2 type C-W equation possesses a great accuracy. Therefore, taking this equation as the model to design the controller can definitely applied to the accurate J_2 relative dynamics or even the whole perturbation model.

This section will derive J_2 invariant relative quasi-periodic orbits based on the controllers preserving Hamiltonian structure, and accordingly come up with a fresh conception of formation-flying called quasi-periodic formation-flying.

Consider the controlled model

$$\Delta \ddot{\mathbf{r}} + \mathbf{A}(t) \Delta \dot{\mathbf{r}} + \mathbf{B}(t) \Delta \mathbf{r} = \mathbf{T}_c \quad (11)$$

where \mathbf{A} is skew symmetric, $\mathbf{A}(t)^T = -\mathbf{A}(t)$, but \mathbf{B} , unfortunately, has no type of symmetry, therefore this linear system no longer maintains Hamiltonian structure. The origin $(\Delta \mathbf{r} = 0, \Delta \dot{\mathbf{r}} = 0)$ is the equilibrium of any time. For any $t \in [0, T_d]$, if the system has the hyperbolic eigenvalues $\pm\sigma(t)$, then the equilibrium is hyperbolic; and the stable\unstable manifolds $\mathbf{u}_+(t)\backslash\mathbf{u}_-(t)$ satisfy with the following relationship:

$$\left[\sigma(t)^2 \mathbf{I} \mp \sigma(t) \mathbf{A}(t) + \mathbf{B}(t) \right] \mathbf{u}_\pm = \mathbf{0}. \quad (12)$$

The 3-D view of the quasi-periodic J_2 invariant relative orbit ($G = 0.5$) is shown in Fig. 2. Figure 3 illustrates the time history of the relative state: let controller gain be $G = 0.5$, relative state be $x_0 = y_0 = z_0 = 1 \text{ km}$ and relative velocity obtained from classical C-W equation be $\dot{y}_0 = -2\omega x_0$, $\dot{x}_0 = \omega y_0/2$, $\dot{z}_0 = 0$.

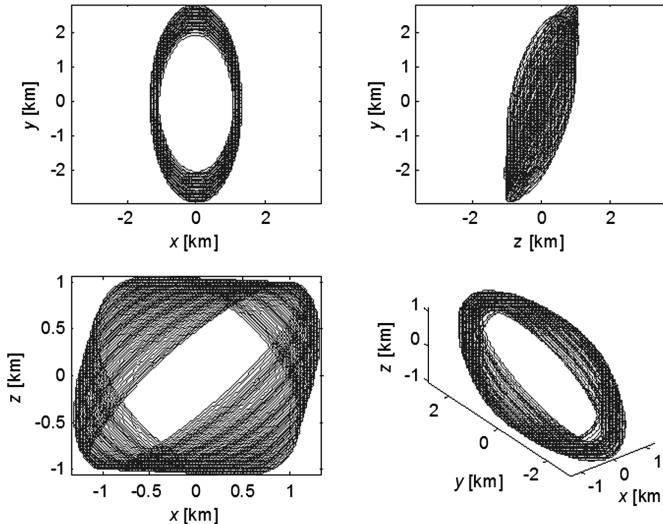


Fig. 1. 3-D view of Quasi-periodic J_2 invariant relative orbits

4 Stability and Spectra for Quasi-periodic J_2 Invariant Relative Orbits

4.1 Stability and Spectra

According to reference [8] and Sect. III.A, choosing the controller gain G that is large enough can change the topology of the equilibrium. Quite different from the autonomous Hamiltonian system, the controlled system is unstable even through its origin is always elliptic for all time.

Fortunately, the controlled Hamiltonian system is still time-periodic. Knowing from the stability theorem of the time-periodic Hamiltonian system, if all the Floquet multipliers of the system lie on the unit circle of complex domain, or the moduli of the multipliers equal to 1, then this Hamiltonian system is Lyapunov stable [9].

The Floquet multipliers in real and complex domains for the quasi-periodic J_2 invariant relative orbits is showed in Fig. 3: when the gain G is bigger than 0.27, all the

Floquet multipliers will lie on the unit circle, so the quasi-periodic J_2 invariant relative orbits are long-term structural stable.

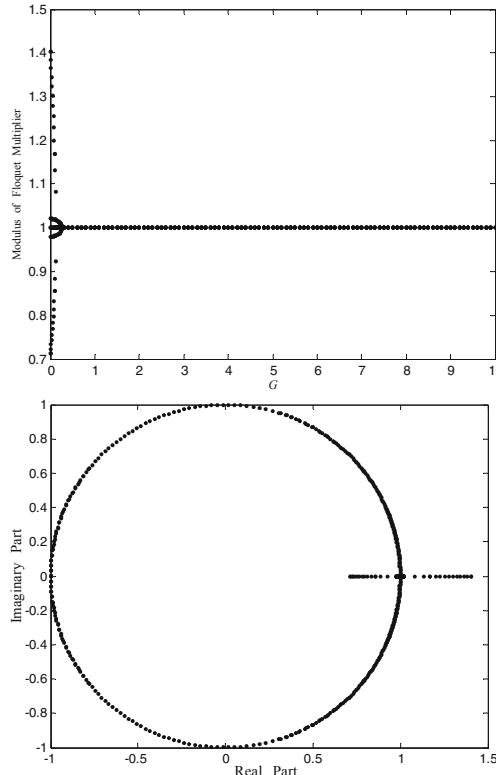


Fig. 2. Floquet multipliers of Quasi-periodic J_2 invariant relative orbits (upper: moduli of the multipliers vs. control gain G , below: Floquet multipliers and the unit circle in complex domain)

For the stable controlled Hamiltonian system ($G > 0.27$), its origin is elliptic during its period. So the Hamiltonian near the equilibrium can be reduced as

$$H = H_0 + \omega_1 \tau_1 + \omega_2 \tau_2 + \omega_3 \tau_3 + O(3) \quad (13)$$

where H_0 is the value of the Hamiltonian on the equilibrium, $\tau_i, i = 1, 2, 3$ are the action variables and $\omega_i, i = 1, 2, 3$ are the frequencies for the three resonant mode. The three resonant mode are the linear decomposition of the stable Lissajous orbits and its x, y, z axes cannot be completely decoupled. The fact reveals that the quasi-periodic J_2 invariant relative orbit is decomposed into three different librations with the spectra as $\omega_1 \ll \omega_2 \approx \omega_3 \approx \varpi$ displayed in Fig. 3.

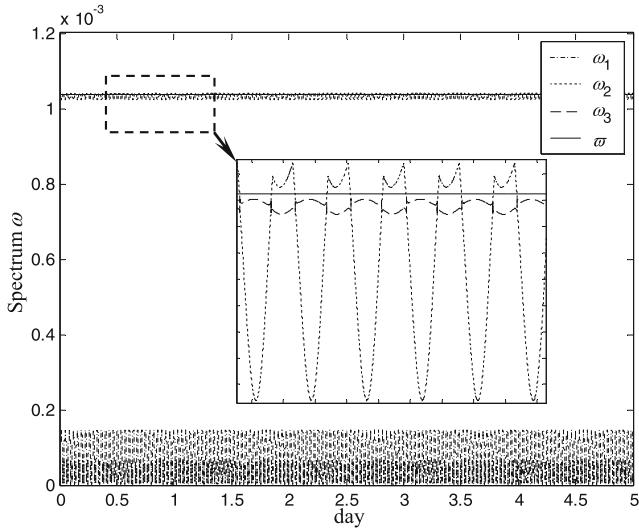


Fig. 3. Three librating spectra of Quasi-periodic J_2 invariant relative orbits

4.2 Physical Implementation of the Controller

The controller feedback is received based on current relative states and can be achieved by continuous-thrust high impulse low thrust propulsion. To generate the stable Lissajous orbits shown in Fig. 1 the acceleration needed is showed in Fig. 4. The biggest acceleration is in the 10^{-5}m/s^2 level and the consumption during 5 days is only $\Delta V = 2.4077\text{ m/s}$.

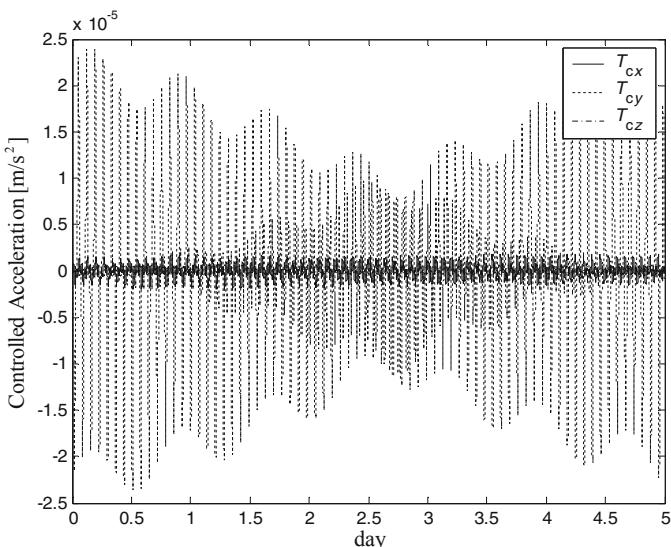


Fig. 4. History of physical implementation of the controlled acceleration

5 Conclusion

An innovative concept quasi-periodic formation flying is claimed in this paper, which can hold bounded trajectories without any restriction initialization compared with the traditional restriction on initial velocity by C-W equation. The improved periodic C-W equation associated with J_2 perturbation (referred as J_2 Type Periodic C-W Equation) is derived from the Taylor expansion of the J_2 relative dynamics and the analytic solutions of osculating orbital elements, to attain the time-periodic linear relative dynamics on the near-circular orbit. J_2 type periodic C-W equation is a perfect approximation to real J_2 relative dynamics, and is employed to design the controller to stabilize the motions of formation flying for its high fidelity to the relative dynamics and simplified representation in the form of periodic matrices.

The controller preserving Hamiltonian structure is constructed for J_2 type C-W equation to generate the quasi-periodic J_2 invariant relative orbits innovatively, which can is quite different from J_2 invariant relative orbits that is proposed firstly from relative elements by Schaub and Alfriend. The controller preserving Hamiltonian structure modifies the topology of the equilibrium from hyperbolic to elliptic, so some interesting inclusions can be achieved that all the trajectories (without any requirements on the initial states or orbital elements) near the equilibrium are bounded according to KAM theory, and the type of periodic orbit referred to as a Lyapunov orbit will emerge at some special initial values of the spacecraft's orbit when the oscillating spectra on radial, cross-track and along-track directions are pairwise-irrational.

References

1. Montenbruck O, Kirschner M, D'Amico S, Bettadpur S (2006) E/I-vector separation for safe switching of the GRACE formation. *Aerospace Sci Technol* 10(7):628–635
2. D'Amico S, Montenbruck O (2006) Proximity operations of formation flying spacecraft using an eccentricity/inclination vector separation. *J Guid Control Dyn* 29(3):554–563
3. Carter T, Humi M (2003) The Clohessy-Wiltshire equations can be modified to include quadratic drag. In: 13th AAS/AIAA space flight mechanics meeting, Ponce, Paper AAS 03-240
4. Izzo D, Sabatini M, Valente C (2003) A new linear model describing formation flying dynamics under J_2 effects. In: Proceedings of 17th AIDAA national congress, Rome, vol 1, pp 493–500
5. Schweighart S, Sedwick R (2002) High-fidelity linearized J_2 model for satellite formation flight. *J Guid Control Dyn* 25(6):1073–1080
6. Schaub H, Alfriend KT (2001) J_2 invariant relative orbits for spacecraft formations. *Celest Mech Astron* 79:77–95
7. Xu M, Xu S (2007) J_2 invariant relative orbits via differential correction algorithm. *Acta Mech Sin* 23(5):585–595
8. Xu M, Xu S (2008) Nonlinear dynamical analysis for displaced orbits above a planet. *Celest Mech Astron* 102(4):327–353
9. Scheeres DJ, Hsiao FY, Vinh NX (2003) Stabilizing motion relative to an unstable orbit: applications to spacecraft formation flight. *J Guid Control Dyn* 26(1):62–73



Stereo Visual Inertial Odometry for Unmanned Aerial Vehicle Autonomous Flight

Quanpan Liu, Zhengjie Wang^(✉), and Huan Wang

Beijing Institute of Technology, Beijing 100081, China
wangzhengjie@bit.edu.cn

Abstract. Visual–inertial SLAM system is very popular in the near decade for the navigation of unmanned aerial vehicle (UAV) system, because it is effective in the environments without the Global Position System (GPS). Due to size and weight constraints, only inexpensive and small sensors can be used. Therefore, there are still challenges in the computational efficiency and robustness of MAV autonomous flight algorithm. We present S-VIO: an optimization-based stereo visual-inertial odometry. Our approach starts with inertial measurement units (IMU) pre-integration, in which IMU measurements are accumulated between several frames using measurement pre-integration. After the initial state estimation converges, a highly precision stereo vision-inertial odometry is obtained by fusing IMU measurements and feature observations. Our approach is validated on the EuRoC MAV datasets. Experimental results prove that our S-VIO has higher accuracy and robustness than the most advanced visual-inertial fusion methods in some challenging situations.

Keywords: Navigation of UAV · IMU pre-integration · Stereo camera · SLAM · Visual-inertial odometry

1 Introduction

Accurate and robust navigation system is important for the autonomous flight of MAV, because correct attitude estimation is necessary for the stability of MAV in flight. Simultaneous Localization and Mapping (SLAM) is very popular in the near decade because it can estimate the motion of MAV and build the map of surrounding environment in real time [1]. Many types of sensors available for SLAM, such as laser lidar, IMU, monocular and stereo cameras. However, they have obvious respective shortcomings: laser lidar will have reflection problem when it encounters glass surface; measurements from cheap IMU typically have significant time-varying biases; and monocular camera are incapable of recovering the metric scale; in contrast to monocular camera, absolute scale problem is easier to deal with in stereo camera [2]. Therefore, stereo vision-inertial odometry (VIO) system, a multi-sensor fusion method, has received extensive attention and was widely used.

In scenarios such as danger investigation and emergency rescue, the VIO algorithm of MAV will face various challenges, such as motion-blurred images, intense changes in light intensity, low texture scenes and so on [3]. Therefore, the VIO algorithm must not only be accurate, but also robust.

To address such challenges, we present S-VIO, an optimization-based stereo VIO, which obtains a highly precision stereo vision-inertial odometry by minimizing IMU residuals and vision residuals in sliding windows. Compared to the monocular camera, our approach utilizes stereo camera in which metric scale can be recovered, has more robustness and accuracy in challengeable environment.

The structure of proposed S-VIO is shown in Fig. 1. The front-end processes raw measurement data from stereo camera and IMU, in which feature points are detected and matched, and IMU measurement data are pre-integrated. In visual-inertial alignment module, we estimate gyroscope bias, gravity vector, accelerometer bias and velocity in turn. These initial values will be used to start the back-end module of the stereo VIO system. In the back-end we obtain a highly precision stereo vision-inertial odometry by minimizing the sum of prior information, IMU residuals and visual residuals. To limit the complexity of S-VIO, the marginalization approach is used to remove useless states from the sliding window. New prior information is obtained on the basis of marginalized approach, and the constraint information of deleted state is retained.

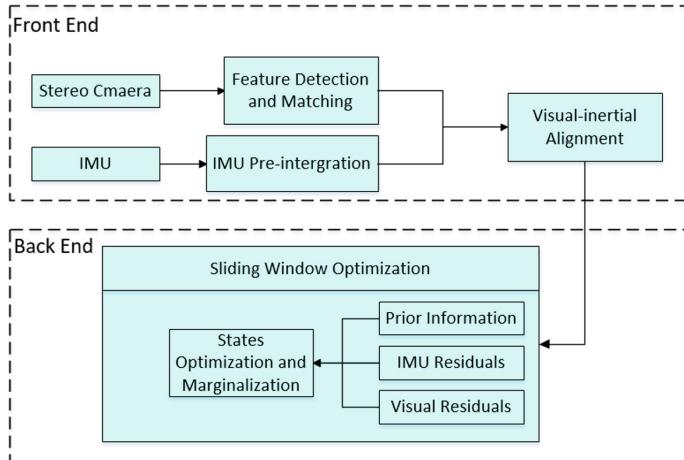


Fig. 1. Structure of our stereo visual-inertial odometry system

1.1 Related Work

There is a lot of research on visual-based odometry and SLAM. However, in this section, we skip the discussion of pure visual methods and focus only on the visual-inertial odometry.

Visual SLAM algorithm is the basis of VIO system, and its back-end can be divided into optimization-based and filtering-based approaches. ROVIO and MSCKF are two well-known tightly-coupled VIO based on filtering [4, 5]. ROVIO is an excellent monocular VIO algorithm, which uses the pixel strength error of patches on two adjacent frames to establish the observation equation [6]. MSCKF is also a well-known filter-based method. The MSCKF maintains several previous cameras pose in

the state vector, and uses visual measurements of the same feature across multiple camera views to form multi-constraint update [7, 8].

Optimization-based approaches obtain the optimal state estimate by minimizing IMU residuals and visual residuals. Compared with the filter-based method, the optimization-based method can repeat the linearization of state vector at different points, so it can obtain higher accuracy. OKVIS is an excellent VIO algorithm based on optimization. It optimizes state by minimizing visual and inertial errors in sliding window [9]. It has a perfect key frame selection strategy, as long as this strategy is satisfied, a new key frame will be generated. The oldest key frames are marginalized. VINS is improved on the basis of OKVIS, which introduces loop detection to eliminate accumulated error. The front-end of VINS uses L-K optical flow method to track features, which can effectively reduce the computational cost [10].

2 Mathematical Formula

2.1 Notations Definition

Figure 2 illustrates the notations definition in the stereo vision inertial odometry system. We define C and B as camera frames and IMU body frames respectively. W represents the geodetic coordinate system. We define $T_{WB} = (R_B^W, P_B^W)$ as the pose of frame B with respect to frame W. $T_{BC} = (R_C^B, P_C^B)$ is the extrinsic parameter between the left camera and IMU in stereo camera, R_C^B represents rotational matrix and P_C^B represents translation vector, which has been calibrated previously. $T_{t,t+1} \in SE(3)$ represents the motion of the camera frame C from time t to time $t + 1$, and $Z_{t,t+1}$ represents a pre-integrated IMU measurement between C_t and C_{t+1} . F_1 and F_2 are the feature points observed simultaneously by the camera frames C_t and C_{t+1} , and $P_{11}^{C_t}$ and $P_{11}^{C_{t+1}}$ are the projection of F_1 in the left camera frame in stereo camera at time t and $t + 1$.

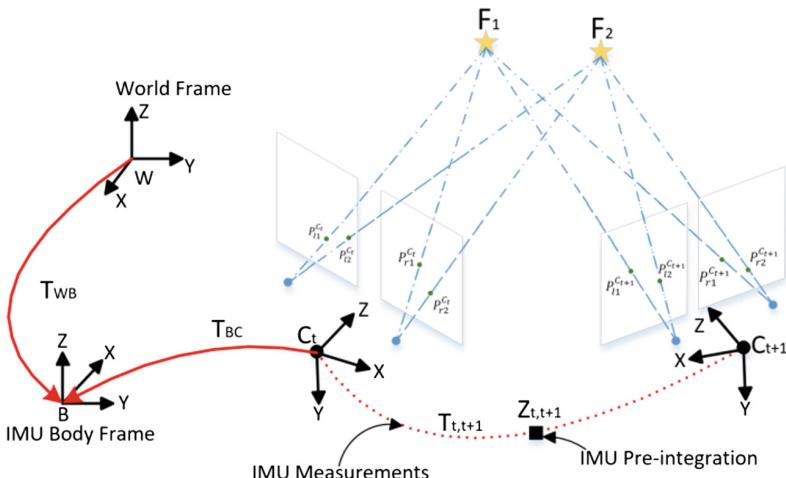


Fig. 2. An illustration of notations definition

2.2 IMU Pre-integration

IMU can provide triaxial angular velocity and triaxial acceleration relative to body frame B and the original measurements \tilde{w} and \tilde{a} contain bias and Gauss white noise:

$$\tilde{w} = w^B + b_w^B + n_w^B \quad (1)$$

$$\tilde{a} = R_B^W (a^W + g^W) + b_a^B + n_a^B \quad (2)$$

Where n_w^B , n_a^B are defined as Gauss white noise of gyroscopes and accelerometers, $n_w^B \sim N(0, \sigma_w^2)$, $n_a^B \sim N(0, \sigma_a^2)$. $g^W = [0, 0, g]^T$ is the gravity vector in frame W. R_B^W is the rotational part of the transformation matrix T_{WB} , which maps a point from the world frame W to a body frame B. b_w^B , b_a^B are the biases from gyroscope and accelerometer.

We define $P_{B_i}^W$, $V_{B_i}^W$, and $q_{B_i}^W$, as the IMU body state at time $t = i$, and a series of values of angular velocity and acceleration are known during $t \in [i, j]$, we can calculate the body state at time $t = j$:

$$P_{B_j}^W = P_{B_i}^W + V_{B_i}^W \Delta t + \iint_{t \in [i, j]} \left[R_{B_t}^W (\tilde{a}_t - b_{a_t}^B) - g^W \right] \delta t^2 \quad (3)$$

$$V_{B_j}^W = V_{B_i}^W + \int_{t \in [i, j]} \left[R_{B_t}^W (\tilde{a}_t - b_{a_t}^B) - g^W \right] \delta t \quad (4)$$

$$q_{B_j}^W = q_{B_i}^W \otimes \int_{t \in [i, j]} \frac{1}{2} \Omega (\tilde{w}_t - b_{w_t}^B) q_{B_t}^W \delta t \quad (5)$$

From Eqs. (3)-(5), we can know that the calculation of body state is based on body frame B_i . When the reference frame changes, we need to recalculate IMU measurements very time-consuming. In order to avoid unnecessary time consumption, we adopt IMU pre-integration technology to solve this problem [11, 12]. Equation (3)-(5) can be written as:

$$R_W^{B_i} P_{B_j}^W = R_W^{B_i} \left(P_{B_i}^W + V_{B_i}^W \Delta t - \frac{1}{2} g^W \Delta t^2 \right) + \alpha_{B_i}^{B_j} \quad (6)$$

$$R_W^{B_i} V_{B_j}^W = R_W^{B_i} \left(V_{B_i}^W \Delta t - g^W \Delta t \right) + \beta_{B_i}^{B_j} \quad (7)$$

$$q_W^{B_i} \otimes q_{B_j}^W = \gamma_{B_i}^{B_j} \quad (8)$$

Where

$$\alpha_{B_i}^{B_j} = \iint_{t \in [i, j]} \left[R_{B_t}^{B_i} (\tilde{a}_t - b_{a_t}^B) \right] \delta t^2 \quad (9)$$

$$\beta_{B_j}^{B_i} = \int_{t \in [i,j]} \left[R_{B_t}^{B_i} \left(\tilde{a}_t - b_{a_t}^B \right) \right] \delta t \quad (10)$$

$$\gamma_{B_j}^{B_i} = \int_{t \in [i,j]} \frac{1}{2} \Omega \left(\tilde{w}_t - b_{w_t}^B \right) q_{B_t}^{B_i} \delta t \quad (11)$$

From Eqs. (9)-(11), we can know that the pre-integration measurements $\alpha_{B_j}^{B_i}$ $\beta_{B_j}^{B_i}$ $\gamma_{B_j}^{B_i}$ and the body states of B_i are independent of each other, which means that there is no need to recalculate when the body states change. In the back-end nonlinear optimization, we use pre-integration measurement as a constraint factor between consecutive key frames. Since the pre-integration measurements $\alpha_{B_j}^{B_i}$ $\beta_{B_j}^{B_i}$ $\gamma_{B_j}^{B_i}$ is related to the bias of IMU, when the bias is updated, we will update the measurements by the first-order approximation method.

$$\alpha_{B_j}^{B_i} \approx \tilde{\alpha}_{B_j}^{B_i} + J_{b_a}^\alpha \delta b_a^B + J_{b_w}^\alpha \delta b_w^B \quad (12)$$

$$\beta_{B_j}^{B_i} \approx \tilde{\beta}_{B_j}^{B_i} + J_{b_a}^\beta \delta b_a^B + J_{b_w}^\beta \delta b_w^B \quad (13)$$

$$\gamma_{B_j}^{B_i} \approx \tilde{\gamma}_{B_j}^{B_i} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} J_{b_w}^\gamma \delta b_w^B \end{bmatrix} \quad (14)$$

Where $J_{b_a}^\alpha$ $J_{b_w}^\alpha$ $J_{b_a}^\beta$ $J_{b_w}^\beta$ $J_{b_w}^\gamma$ are the Jacobian matrices of pre-integrated measurements with respect to bias.

3 Visual Measurements

We use KLT optical flow algorithm to track the existing features and match stereo features [9]. Compared with the descriptor-based methods, the optical flow-based methods can save more computation. We use two strategies to eliminate outliers in image processing algorithms. RANSAC with fundamental matrix model and reverse check of KLT optical flow algorithm are applied to reject outliers in tracking. After outlier rejection, new feature points can be added through the FAST detector to meet the minimum threshold of feature points.

In this step, we have two key frame selection strategies. If the rotation distance and the translation between the last two adjacent frames calculated by IMU pre-integration is over the threshold value, we consider the current frame as a new key frame. Another criterion is the tracking standard of feature points. If the current frame tracking point is less than 70% of the last key frame, we also treat the frame as a new key frame.

We adopt pinhole camera model and map the three-dimensional feature point F_1 to the camera pixel coordinate system $P_{II}^{C_1}$ through the camera projection function π .

$$\pi(P_{II}^{C_1}) = \begin{bmatrix} f_u \frac{x_c}{z_c} + c_u \\ f_v \frac{y_c}{z_c} + c_v \end{bmatrix} \quad F_1 = [x_c \quad y_c \quad z_c]^T \quad (15)$$

Where $[f_u \ f_v]^T$ and $[c_u \ c_v]^T$ is the camera intrinsic parameters.

Our S-VIO combines the advantages of static stereo and monocular cameras. Absolute scale can be directly calculated from the fixed baseline known to the stereo camera. However, constant baselines limit the performance of static stereo. In our S-VIO system, stereo measurements are only used when the depth of feature is initialized, and then long-term tracking is performed with monocular measurements. By adopting vision Structure from Motion (SFM), we can triangulate all features observed between two consecutive keyframes. Finally, we adopt the perspective-n-point (PnP) method to recover the relative rotation and translation of all frames in the sliding window [13].

3.1 Visual-Inertial Alignment

The accuracy of initial values is very important for the final convergence of the non-linear visual-inertial navigation system. Therefore, in our visual-inertial alignment module, we estimate the gyroscope bias, gravity vector, accelerometer bias and velocity respectively.

3.1.1 Gyroscope Bias Correction

Rotation $R_{C_i}^{C_0}$ and $R_{C_j}^{C_0}$ between two adjacent keyframes C_j and C_i can be obtained from visual SFM. And integration $\gamma_{B_j}^{B_i}$ between them can be obtained by IMU pre-integration. We can estimate the gyroscope bias by solving the following least squares problem.

$$\min_{\delta b_{w_t}^B} \sum \left\| R_{C_j}^{C_0 -1} \otimes R_{C_i}^{C_0} \otimes \gamma_{B_j}^{B_i} \right\|^2 \quad (16)$$

Where

$$\gamma_{B_j}^{B_i} \approx \tilde{\gamma}_{B_j}^{B_i} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} J'_{b_w} \delta b_{w_t}^B \end{bmatrix} \quad (17)$$

In Eq. (17), J'_{b_w} denotes the first-order approximation of the changing gyroscope bias. We can estimate the initial correction of gyroscope deviation $b_{w_t}^B$ by Gauss-Newton method. Then, the new gyroscope bias is used to update the IMU pre-integration.

3.1.2 Velocity, Gravity Vector Correction

We define the variables to be estimated as:

$$X_I = \begin{bmatrix} V_{b_0}^{b_0} & V_{b_1}^{b_1} & \dots & V_{b_n}^{b_n} & g^C \end{bmatrix}$$

Where $V_{b_k}^{b_k}$ is velocity in the body frame, g^C is the gravity vector in the camera frame C.

Assuming that we have pre-calibrated the extrinsic parameter $T_{BC} = (R_C^B, P_C^B)$, we can convert the pose from the camera frame to the body frame.

$$q_{B_i}^C = q_{C_i}^C \otimes q_C^{B-1} \quad (18)$$

$$P_{B_i}^C = P_{C_i}^C - R_C^C P_C^B \quad (19)$$

Considering two adjacent keyframes B_i and B_j in sliding window, then Eqs. (6)-(7) can be written as:

$$\alpha_{B_j}^{B_i} = R_{C_i}^{B_i} \left(P_{B_j}^C - P_{B_i}^C \right) - R_{B_i}^C V_{B_i}^{B_i} \Delta t + \frac{1}{2} g^C \Delta t^2 \quad (20)$$

$$\beta_{B_j}^{B_i} = R_{C_i}^{B_i} \left(R_{B_j}^C V_{B_j}^{B_j} - R_{B_i}^C V_{B_i}^{B_i} + g^C \Delta t \right) \quad (21)$$

We combine (18)-(21) to get the following formula:

$$\tilde{Z}_{B_j}^{B_i} = \begin{bmatrix} \alpha_{B_j}^{B_i} - P_{B_i}^C + R_{C_i}^{B_i} R_{B_i}^C P_C^B - R_{C_i}^{B_i} \left(P_{B_j}^C - P_{B_i}^C \right) \\ \beta_{B_j}^{B_i} \end{bmatrix} = H_{B_j}^{B_i} X_I + n_{B_j}^{B_i} \quad (22)$$

Where

$$H_{B_j}^{B_i} = \begin{bmatrix} -I \Delta t & 0 & \frac{1}{2} R_C^{B_i} \Delta t^2 \\ -I & R_C^{B_i} R_{B_j}^C & R_C^{B_i} \Delta t \end{bmatrix} \quad (23)$$

In the above formula, $R_{B_i}^C, R_{B_j}^C, P_{B_i}^C, P_{B_j}^C$ are obtained from visual SFM. Δt is the time interval between two consecutive frames. We can get the linear least squares problem:

$$\min_{X_I} \sum \left\| \tilde{Z}_{B_j}^{B_i} - H_{B_j}^{B_i} X_I \right\|^2 \quad (24)$$

By solving this linear least square problem, we can obtain velocities for every keyframes in the window and the gravity vector.

3.1.3 Accelerometer Bias Correction

So far, we have not corrected accelerometer bias. We fix the gravity vector g^C . Then we estimate the bias of accelerometer by Eq. (6). We define the variables to be estimated as:

$$X_A = \begin{bmatrix} V_{b_0}^{b_0} & V_{b_1}^{b_1} & \dots & V_{b_n}^{b_n} & \delta b_{a_t}^B \end{bmatrix}$$

We add accelerometer bias to Eq. (20):

$$\left[\alpha_{B_j}^{B_i} - P_C^B + R_C^{B_i} R_{B_i}^C P_C^B - R_C^{B_i} (P_{B_j}^C - P_{B_i}^C) - \frac{1}{2} g^C \Delta t^2 \right] = \begin{bmatrix} V_{B_i}^{B_i} \\ V_{B_j}^{B_i} \\ \delta b_{a_t}^B \end{bmatrix} \begin{bmatrix} -I \Delta t \\ 0 \\ -J_{b_a}^\alpha \end{bmatrix}^T \quad (25)$$

The Jacobian $J_{b_a}^\alpha$ denotes a first-order approximation of the changing accelerometer bias. Similar to Sect. 3.1.1, after estimating the accelerometer bias, we update all IMU pre-integration with a new accelerometer bias.

4 Tightly-Coupled Stereo VIO

After convergence of the visual inertial alignment module, these initial values will be used to start the back-end module of the stereo VIO system. We maintain a sliding window method to limit the complexity of the algorithm. As shown in Fig. 3, the optimized S-VIO includes visual measurement, pre-integrated IMU measurement, continuous IMU body state and camera pose.

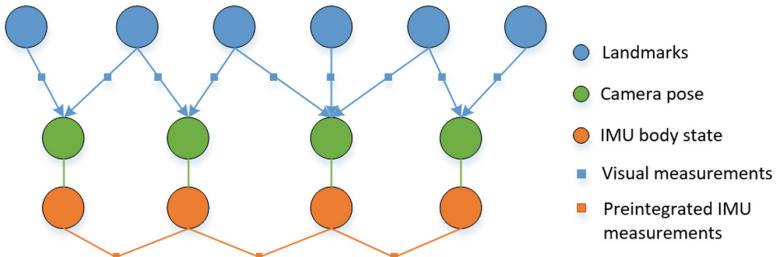


Fig. 3. The optimization-based stereo visual-inertial odometry

4.1 Formula

The state variables to be estimated by S-VIO are defined as:

$$\begin{aligned} X &= [X_0 \quad X_1 \quad \dots \quad X_n \quad X_C^B] \\ X_k &= [P_{B_k}^W \quad V_{B_k}^W \quad q_{B_k}^W \quad b_a^B \quad b_w^B] \quad k \in [0, n] \\ X_C^B &= [P_C^B \quad q_C^B] \end{aligned}$$

Where X_k contains the position, velocity and direction of the k^{th} IMU frame with reference to the world coordinate frame, as well as the bias of acceleration and angular velocity in the IMU coordinate frame. n is the number of keyframes in the sliding window. We obtain a highly precision stereo vision-inertial odometry by minimizing the sum of prior information, IMU residuals and visual residuals in the sliding window to obtain the maximum posteriori estimation [10].

$$\min_X \left\{ \|r_p - H_p X\|^2 + \sum_{k \in B} \left\| r_B(\tilde{Z}_{B_{k+1}}^{B_k}, X) \right\|_{P_{B_{k+1}}^{B_k}}^2 + \sum_{(l,j) \in C} \left\| r_C(\tilde{Z}_l^{C_j}, X) \right\|_{P_l^{C_j}}^2 \right\} \quad (26)$$

Where $\{r_p, H_p\}$, $r_B(\tilde{Z}_{B_{k+1}}^{B_k}, X)$ and $r_C(\tilde{Z}_l^{C_j}, X)$ are prior information, IMU residuals and visual residuals respectively.

4.2 IMU Measurement Residual

According to the mathematical model of IMU, given the IMU measurements between two adjacent keyframes B_i and B_j in the sliding window, the IMU measurement residual can be defined as:

$$r_B^{15*1}(\tilde{Z}_{B_j}^{B_i}, X) = \begin{bmatrix} \delta\alpha_{B_j}^{B_i} \\ \delta\theta_{B_j}^{B_i} \\ \delta\gamma_{B_j}^{B_i} \\ \delta b_a^B \\ \delta b_w^B \end{bmatrix} = \begin{bmatrix} R_W^{B_i}(P_{B_j}^W - P_{B_i}^W - V_{B_i}^W \Delta t + \frac{1}{2} g^W \Delta t^2) - \tilde{\alpha}_{B_j}^{B_i} \\ R_W^{B_i}(V_{B_j}^W - V_{B_i}^W \Delta t + g^W \Delta t) - \tilde{\beta}_{B_j}^{B_i} \\ 2 \left[q_{B_i}^{W-1} \otimes q_{B_j}^W \otimes (\tilde{\gamma}_{B_j}^{B_i})^{-1} \right]_{xyz} \\ b_a^{B_i} - b_a^{B_j} \\ b_w^{B_i} - b_w^{B_j} \end{bmatrix}_{15*1} \quad (27)$$

Where $[\cdot]_{xyz}$ represents the real part of the quaternion. $[\tilde{\alpha}_{B_j}^{B_i}, \tilde{\beta}_{B_j}^{B_i}, \tilde{\gamma}_{B_j}^{B_i}]^T$ are the IMU pre-integration between two adjacent key frames B_i and B_j .

4.3 Visual Measurement Residual

With the conventional pinhole-camera model, if the feature F_1 is observed for the first time in the i^{th} image, the re-projection error in the j^{th} image is defined as:

$$r_C(\tilde{Z}_l^{C_j}, X) = \pi^{-1} \left(\begin{bmatrix} \tilde{u}_{F_1}^{C_j} \\ \tilde{v}_{F_1}^{C_j} \end{bmatrix} \right) - R_B^C(R_W^{B_j}(R_{B_i}^W(R_C^B \pi^{-1} \left(\begin{bmatrix} \tilde{u}_{F_1}^{C_j} \\ \tilde{v}_{F_1}^{C_j} \end{bmatrix} \right) + P_C^B) + P_{B_i}^W - P_{B_j}^W) - P_C^B) \quad (28)$$

Where $[\tilde{u}_{F_1}^{C_j}, \tilde{v}_{F_1}^{C_j}]$ is the coordinates of the pixels projected from the feature point F_1 to the i^{th} frame image. $[\tilde{u}_{F_1}^{C_j}, \tilde{v}_{F_1}^{C_j}]$ is the projection of the same feature point on the j^{th} frame image. π^{-1} is a back-projection function that converts pixel coordinates into unit vectors through camera intrinsic parameters.

4.4 Marginalization

We adopt the marginalization strategy to restrict the complexity of the back-end and ensure the real-time and effective operation of the optimization-based S-VIO. As shown in Fig. 4, if the second latest frame is determined to be a key frame, the frame is retained in the sliding window and the oldest frame is removed at the same time. When

the oldest frame is marginalized, the measurements associated with the oldest frame and the feature points that can only be observed by the frame will also be removed. If the second latest frame is determined not to be a key frame, we will marginalize the frame and eliminate all visual measurements related to the frame [9].

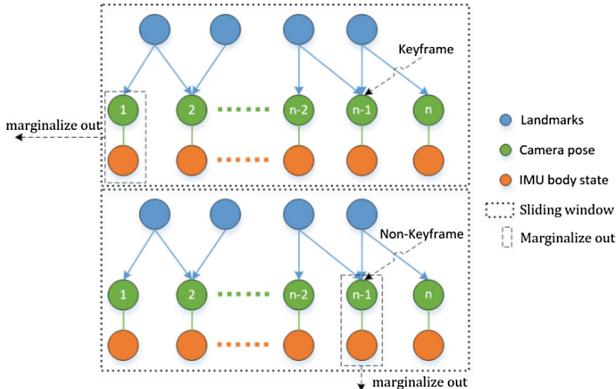


Fig. 4. An illustration of our marginalization strategy

We adopt Schur complement to remove the information related to the marginalized measurements. On this basis, new priori is constructed and added to the existing priori.

5 Experiment

We evaluated the accuracy and robustness of S-VIO on the EuRoC MAV Datasets [14]. The datasets contain stereo images, synchronized IMU measurements, and accurate motion and structure ground-truth. The S-VIO method is compared with the most advanced visual-inertial fusion methods VINS-Mono. For fairness, VINS-Mono

Table 1. The RMSE results of translation (m) are shown in the table below.

| Dataset | VINS-Mono | S-VIO |
|---------|-----------|----------|
| 1 | 0.266314 | 0.339084 |
| 2 | 0.497430 | 0.482999 |
| 3 | 0.550573 | 0.614529 |
| 4 | 0.678140 | 0.524223 |
| 5 | 0.229267 | 0.247859 |
| 6 | 0.436770 | 0.434792 |
| 7 | 0.375504 | 0.345927 |
| 8 | x | 0.179598 |
| 9 | 0.395870 | 0.370413 |
| 10 | 0.435668 | 0.426532 |

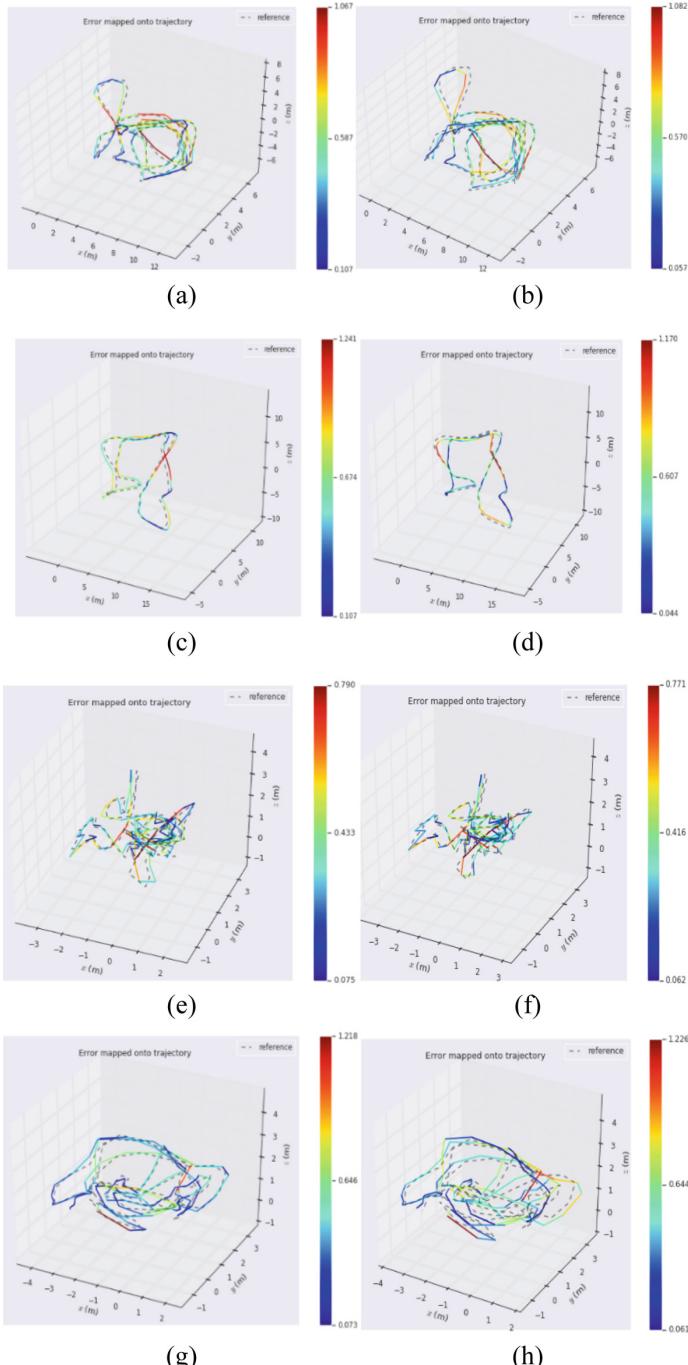


Fig. 5. The trajectory of keyframes on sequences. The four trajectory heat maps on the left are from VINS-Mono. The four trajectory heat maps on the right are from S-VIO. Colors represent the corresponding absolute pose errors.

does not use loop detection module. VINS-Mono is an excellent SLAM system based on feature point method, and its back-end adopts non-linear optimization method to obtain a highly precision monocular vision-inertial odometry.

In the experiment, we use the open source software package EVO to evaluate the accuracy and robustness of our algorithm. By comparing the estimated trajectory error with the actual ground trajectory error, we calculate the absolute attitude error (APE) as an index to evaluate the performance of the algorithm. Table 1 shows the root mean square error (RMSE) of all trajectory translations on the EuRoC MAV datasets.

As can be seen from Table 1, our S-VIO performs well and has high robustness when the camera moves in a challenging environment. By comparing their trajectories, we conclude that in challenging environments, our approach performs better than VINS-Mono. Figure 5 shows more intuitively the heat maps of trajectories estimated by S-VIO and VINS-Mono.

6 Conclusions

We present S-VIO: an optimization-based stereo visual-inertial odometry, which consists of two modules. The front-end module receives and processes raw measurement from stereo camera and IMU, and the back-end module obtains a highly precision stereo vision-inertial odometry by minimizing the sum of prior information, IMU residuals and visual residuals in the sliding window. We also compare our approach with the most advanced monocular VIO methods on the EuRoC datasets. Our S-VIO performs well in estimate the motion of the camera and is more robust in challenging environments. In the future, we plan to add line features to the front-end for VIO algorithm's outstanding performance in urban environment.

References

- Mur-Artal R, Montiel J, Tardós JD (2015) ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Trans Robot* 31:1147–1163
- Mur-Artal R, Tardós JD (2017) ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Trans Robot* 33:1255–1262
- Shen S, Mulgaonkar Y, Michael N, Kumar V (2014) Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft MAV. *IEEE international conference on robotics and automation*, pp 4974–4981
- Li M, Mourikis A (2013) High-precision, consistent EKF-based visual inertial odometry. *Int J Robot Res* 32:690–711
- Bloesch M, Omari S, Hutter M, Siegwart R (2015) Robust visual inertial odometry using a direct EKF-based approach. *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp 298–304
- Bloesch M, Burri M, Omari S, Hutter M, Siegwart R (2017) Iterated extended Kalman filter based visual-inertial odometry using direct photometric feedback. *Int J Robot Res* 36:1053–1072

7. Huang GP, Mourikis AI, Roumeliotis SI (2008) Analysis and improvement of the consistency of extended Kalman filter-based SLAM. In: Proceedings of the IEEE international conference on robotics and automation, pp 473–479
8. Mourikis AI, Roumeliotis SI (2007) A multi-state constraint Kalman filter for vision-aided inertial navigation. In: IEEE international conference on robotics and automation, pp 3565–3572
9. Leutenegger S, Lynen S, Bosse M, Siegwart R, Furgale P (2015) Keyframe-based visual–inertial odometry using nonlinear optimization. *Int J Robot Res* 34:314–334
10. Shen S, Michael N, Kumar V (2015) Tightly-coupled monocular visual-inertial fusion for autonomous flight of rotorcraft MAVs. In: IEEE international conference on robotics and automation
11. Lupton T, Sukkarieh S (2009) Efficient integration of inertial observations into visual SLAM without initialization. IEEE/RSJ international conference on intelligent robots and systems, pp 1547–1552
12. Forster C, Carlone L, Dellaert F, Scaramuzza D (2016) On-manifold preintegration for real-time visual–inertial odometry. *IEEE Trans Robot* 33:1–21
13. Yang Z, Shen S (2017) Monocular visual–inertial state estimation with online initialization and camera–imu extrinsic calibration. *IEEE Trans Autom Sci Eng* 14:39–51
14. Burri M et al (2016) The EuRoC micro aerial vehicle datasets. *Int J Robot Res* 35:1157–1163



A Traffic Network Congestion Diffusion Model Based on Gene Regulation Mechanism

Xiangwei Wu, Ning Huang^(✉), and Kan Xu

School of Reliability and System Engineering, Beihang University,
Beijing 100191, China
hn@buaa.edu.cn

Abstract. Under signal lights regulation, local congestion faults at intersections are coupled with each other and form regulation coupling relationship. Network congestion can spread from the local section to the whole network under this coupling relationship. And the current research on network congestion diffusion is modeled by considering a variety of physical connections and influencing factors, and the research on traffic lights regulation focuses on regulation strategies based on real-time data. These studies lack the support of the congestion diffusion mechanism under the regulation coupling relationship, which is insufficient to support the research on this fault propagation process and the coupling relationship between local faults. Because there is an urgent need for a new method to solve the above problem, based on the regulation mechanism of gene transcription system and the characteristics of traffic network, this paper creatively puts forward the congestion diffusion model of traffic network under the regulation coupling relationship, and this model can be used to evaluate the congestion of the whole network. The case study demonstrates the use of the model, and the results show that the proposed model can support the study of network congestion diffusion mechanism under the regulation coupling relationship.

Keywords: Regulation coupling relationship · Gene regulation mechanism · Congestion diffusion

1 Introduction

The congestion diffusion is a typical phenomenon that the congestion fault is constantly diffused in the network under the influence of the coupling relationship. In the traffic network, traffic lights control the behavior of vehicles on the road section. Under the effect of signal lights regulation, local congestion faults at intersections are coupled with each other and form regulation coupling relationship. Since regulation coupling relationship is a common congestion diffusion coupling relation, it is particularly important to model the network congestion diffusion under this coupling relation, analyze the process of network congestion diffusion, and understand the mechanism of network congestion diffusion.

The current research on network congestion diffusion is mainly based on cascading failure model. At first, Crucitti [1] established the congestion diffusion model considering

network transmission efficiency based on the classical capacity load model in 2003, but because there are many kinds of coupling relations in the traffic network, such as the regulation coupling relationship, it is not reasonable to study the congestion diffusion of the traffic network only considering the transmission efficiency of the network. After that many scholars have analyzed the influence of several types of network coupling relations on congestion diffusion in modeling. Zhan [2] considered the interaction between the network topology and the load. Parshani [3] considered the network failure under the connection and dependency of the network at the same time. Su [4] studied the failure of one node on the same dependent chain affecting the whole link. Rui [5] considered the traffic redistribution strategy under the coupling relationship between normal path and failure path. Wang [6] considered the coupling relationship between two types of nodes in the network. And some scholars analyzed the network fault propagation by considering the concepts of asymmetric coupling cluster [7] and dynamic coupling cluster [8]. In conclusion, the current model based on cascading failure mainly models the network fault propagation from the perspective of topology structure and flow change, and focuses on the functional faults of nodes or edges in the network, while network congestion is a performance failure, and we cannot simply use “functional” or “failure” to describe the state of nodes or edges. These models describe a variety of physical connection relationships and coupling relationships, but they cannot support the study of network congestion diffusion under regulation coupling relationship.

Meanwhile, based on cellular automata model, many scholars studied the influence of factors, such as autonomous driving [9], network topology [10], slow start [11], acceleration and deceleration [12], on traffic network congestion diffusion. However, the model based on cellular automata reflects the influence of various factors in the traffic network by setting corresponding rules, so as to be closer to the reality, and does not focus on the coupling relationship between local congestion faults in traffic networks. Similarly, it cannot support the study of network congestion diffusion under regulation coupling relationship.

As for the research on traffic signal light regulation, initially some researchers focused on isolated traffic lights [13]. However, the local optimal solution is not the global optimal solution, so this paper studies the congestion diffusion of the whole network under the regulation coupling. Aiming at the traffic light system of the whole urban road network, the current research mainly uses the real-time traffic flow information to propose signal control strategies, such as Younes [14]’s intelligent signal control algorithm and Liu [15] ‘ s distributed control method integrating V2X network. Some scholars use wireless sensor network [16], device to device communication [17], roadside infrastructure and onboard sensors [18] to obtain traffic flow information and propose their signal control strategies respectively. In conclusion, the current research on signal lights regulation mainly focus on collecting real-time traffic flow information and proposing signal control strategies. Then the control strategy is optimized according to the simulation results. However, these studies do not focus on the coupling relationship between local congestion faults in traffic networks, and lack the research on the congestion diffusion mechanism under the regulation coupling relationship.

In conclusion, since the existing researches do not support the study of congestion diffusion mechanism under the regulation coupling relationship, a new method is required to study the network congestion diffusion mechanism under the regulation

coupling relationship. Modern medical science uses the method of biomolecular network to study the complex coupling relationship between biomolecules. In order to study the regulation coupling relationship, we model the network congestion diffusion based on the regulation mechanism of gene transcription system. This model can help analyze the process and mechanism of network congestion diffusion, provide theoretical support for alleviating the network congestion diffusion, and provide theoretical basis for proposing better regulation strategies.

In this paper, a model of network congestion diffusion with regulation coupling relationship is proposed. In Sect. 2, regulation mechanism in gene transcription system is briefly introduced. In Sect. 3, the traffic congestion diffusion model based on regulation coupling relationship is proposed. In Sect. 4, taking Barcelona urban road network as an example, the model proposed in this paper is used to analyze the network congestion diffusion under the regulation coupling relationship. Finally, Sect. 5 provides our conclusion.

2 Regulation Mechanism in Gene Transcription Systems

This paper compares the change of load on the traffic network section to the change of biomolecular concentration and compares traffic signal lights to the transcription factors in the process of gene transcription. The relevant basis for the above analogies is as follows:

- Similar regulation causes. Both need to adapt to changes in the external environment. Transcription factors maintain normal changes in intracellular molecular concentration through their regulation actions when cellular environment is altered. Similarly, in the traffic network, the regulation strategy should be changed according to the change of network load.
- Similar regulation logic. In transcriptional regulation, the proportional relationship between polymeric and non-polymeric concentrations of transcription factors controls the rate of protein molecule synthesis. In traffic network regulation, it is necessary to control different signal timing parameters according to different loads on the road.
- Similar regulation results. In the process of gene expression, once the external environment changes, the synthesis rate of biomolecules can be changed through the regulation of transcription factors, so that the concentration of biomolecules in cells can be maintained at a normal level. Similarly, as regulatory factors, traffic signal lights redistribute the traffic flow.

Therefore, the modeling method of coupling relationship in biomolecular networks [19] is used for reference to analyze the regulation coupling relationship in this paper. During the transcription phase of gene expression, according to the changes of protein molecular concentration, the following dynamic equation can be obtained:

$$\frac{dY}{dt} = k_1 S \frac{K_d^p}{K_d^p + Y^p} - k_2 E_T \frac{Y}{K_m + Y} \quad (1)$$

where the first term on the right side of the equation represents the protein production rate regulated by transcription factors, and different parameter p corresponds to different polymer morphology. The second term on the right side of the equation represents the protein degradation process. And Y represents protein (transcription factor) molecular concentration, k_1 is the maximum transcription rate of mRNA, S is the external signal strength, K_d is the concentration of transcription factor polymer, p is Hill constant, k_2 is the degradation rate of protein, E_T is the concentration of degradation enzyme, and K_m is Machaelis coefficient. It can be seen from formula (1) that transcription factors, as key units of mRNA synthesis, control the whole gene expression process.

3 Traffic Network Congestion Diffusion Model Based on Regulation Coupling Relationship

The model of biomolecular network based on the regulation mechanism in gene transcription system provides a new idea for the research of complex diseases. However, there are some difference between the regulation coupling relationship in traffic network and the regulation mechanism of gene transcription, so it is necessary to modify the model. In order to describe the load change in traffic network under regulation coupling relationship, the dynamic function of the traffic load is proposed as follows:

$$\begin{aligned} \frac{dl_{ij}(t)}{dt} &= H_i(\alpha) \sum_{p \in \Gamma_i} f_i(t, p) v_{pi}(t) - H_j(\alpha) \sum_{q \in \Gamma_j} f_j(t, q) v_{jq}(t) \\ &= \frac{l_{p'i}^\alpha(t)}{\sum_{p \in \Gamma_i} l_{pi}^\alpha(t)} \sum_{p \in \Gamma_i} f_i(t, p) v_{pi}(t) - \frac{l_{qj}^\alpha(t)}{\sum_{q \in \Gamma_j} l_{qj}^\alpha(t)} \sum_{q \in \Gamma_j} f_j(t, q) v_{jq}(t) \end{aligned} \quad (2)$$

where the term on the left side of the equation represents the overall change rate of load on the road segment. The first term on the right indicates the increased load of this segment due to regulation coupling relationship, which can be analogous to the process of protein molecule synthesis in the transcriptional regulation model. The second term represents the load reduction due to the regulation coupling relationship of the section. $H_i(\alpha)$ represents the regulation coupling intensity of this section at this time, which can be understood as a coefficient in signal control, which is related to the section load and controls the numerical value of load transfer.

In addition, in formula (2) $f_i(t, p)$ is a demonstration function, which is used to judge which road section around the intersection i under this phase has the right of way, thus indicating whether the load can be transferred. $v_{pi}(t)$ represents the load transfer rate on the current section.

The reference of the model in this paper is mainly reflected in the regulation intensity $H_i(\alpha)$. The specific calculation formula is as follows:

$$H_i(\alpha) = \frac{l_{p'i}^\alpha(t)}{\sum_{p \in \Gamma_i} l_{pi}^\alpha(t)} \quad (3)$$

where $l_{p'i}(t)$ represents the load on the road section that can enter the intersection i (i.e. obtain the right of way), and $\sum_{p \in \Gamma_i} l_{pi}^\alpha(t)$ represents the weighted sum of the load on the road section around the intersection (the weight is the regulation factor α). Similar to the parameter p in formula (1), the regulation coefficient α of the model in this paper is proposed to control the regulation coupling intensity under different road segment loads, which affects the numerical value of the regulation intensity $H_i(\alpha)$ from the perspective of road segment loads. The physical meaning can be understood as follows: for a given right-of-way road section, the regulation intensity $H_i(\alpha)$ is proportional α to the power of the current load on the road section, that is, more green time will be allocated to the section with greater load. The definitions of the corresponding parameters in the model are shown in Table 1.

Table 1. Partial parameters of the transcription system model and the model in this paper

| Parameters and implications in a transcription system model | Parameters and meanings in the model in this paper |
|--|--|
| $K_d^p/K_d^p + Y^p$ is the direct proportionality coefficient of the current transcription rate of the cell, which is related to the concentration of the transcription factor polymer and the total concentration of the transcription factor | $H_i(\alpha) = l_{p'i}(t) / \sum_{p \in \Gamma_i} l_{pi}^\alpha(t)$ represents the regulation intensity under the current phase, which is related to the current section load and all the load at the junction. Γ_i represents the set of edges around node i |
| K_d represents the current polymeric concentration of transcription factors | $l_{p'i}(t)$ represents the load on the current segment to be transferred |
| Y represents the current transcription factor concentration | $\sum_{p \in \Gamma_i} l_{pi}^\alpha(t)$ represents the load-weighted sum of all segments entering the intersection |
| p represents the Hill coefficient, specifically the polymer-binding state of the transcription factor, which controls the direct proportionality coefficient in the model | α is the regulation coefficient in this model, which is used to control the regulation coupling intensity |

In traffic network, the traffic signal controls the load distribution on the road. When a certain section is congested, the traffic flow on the normal section cannot enter the congested section, and on the other hand, the traffic flow on the congested section will flow into the normal section in the next green light direction. In both cases, the adjacent section will be overloaded to form congestion, which eventually leads to the diffusion of congestion in the network.

A two-dimensional directed Lattice network is used as the overall network topology of the model, where nodes and edges represent the intersection and road in the actual traffic network respectively. In this paper, the evaluation index of network congestion in reference [20] is adopted, and the calculation method is as follows:

$$D = \sqrt{\frac{1}{|\{E\}|} \sum_{ij \in \{E\}} (l_{ij} - \bar{l}_{ij})^2} \quad (4)$$

where $\{E\}$ represents the set of all edges in the network, $|\{E\}|$ represents the total number of edges in the network, and \bar{l}_{ij} represents the average load of the network. This measurement method is more suitable for the model in this paper, mainly because the total traffic in this model remains unchanged, and the network initial parameter setting only contains the load information. When the index D is very large, it means that part of the network side bears a large load and congestion is more likely to occur. When index D is relatively small, the load distribution in the network is relatively even, indicating that the network is in normal operation.

4 Case Study and Discussions: Barcelona Road Network

Based on the model proposed in this paper, part of the road network in Barcelona [21] is taken as the research object for research and analysis. We abstract the topology of this part of road network, and obtain the network topology structure as shown in Fig. 3. It can be roughly regarded as a $12 * 10$ two-dimensional grid network, which can be modified to a certain extent at some intersections and sections.

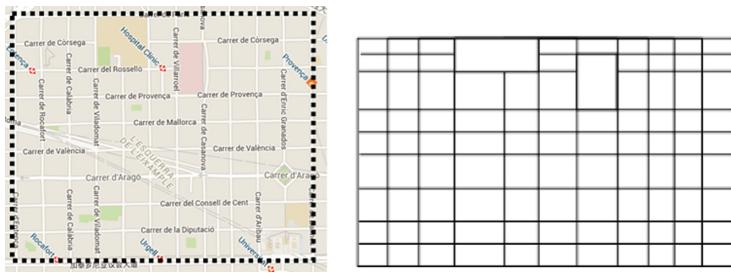


Fig. 1. Road traffic network and topological structure of Barcelona city

In the actual network, the intersection and road section on the boundary also have the load transfer with the external intersection or road section, but this part of the intersection and road section and the external load transfer process is ignored in this paper. In the network topology of Fig. 1, there are 119 nodes and 426 directed edges. After obtaining the topology of the case, the network congestion diffusion is studied according to the model proposed in this paper. Simulation was performed 10 times in total. The network load parameters p is set as $[0.05, 0.1, 0.15, \dots, 0.5]$, and the regulation coupling coefficient is set as $[0, 0.5, 1, 1.5, 2]$. It was assumed that the network node processing capacity $K = 0.001$. The simulation process is as follows:

Step 1: Input the number of network nodes and the total step of simulation time.

The model uses a two-dimensional Lattice network topology with 119 nodes and 426 directed edges to determine the network size. The total step length of simulation time is designed as 10000.

Step 2: Set the initial load on the edge of the network and the processing capacity on the nodes.

In this model, the directed edge of the network represents the road section in the traffic network, so the initial load assigned to each edge of the network represents the initial traffic flow on the road section, and the network load parameter $p = 0.05$; At the same time, the network node represents the junction, and the processing capacity of the node is also set in the initial state of simulation. The traffic capacity of the junction limits the transfer rate on the section, and the processing capacity of the network node $K = 0.001$. In addition, the regulation coefficient α is $[0, 0.5, 1, 1.5, 2]$.

Step 3: Set the initial phase of nodes in the network.

Different phases determine the right of way on the current section.

Step 4: Calculation of network load transfer.

Traverse every node in the network in the following order.

Step 4.1: Calculate the current phase of the node.

Firstly, according to the initial phase and the current time step of the node, the demonstration function $f_i(t, p)$ is used to determine the current phase, so as to judge which road section around the node gets the current right of way.

Step 4.2: Calculate the coupling intensity under the current phase.

The coupling intensity of the current phase is calculated by formula (3) according to the load $l_{p'i}(t)$ to be transferred and the load weighted and $\sum_{p \in \Gamma_i} l_{pi}^\alpha(t)$ on the surrounding road and the phase of the node

Step 4.3: Calculate the load transfer amount.

According to the regulation coupling intensity under this phase, the traffic capacity of the intersection (default $K = 0.001$) to calculate the load transfer amount on the corresponding road section;

Step 5: Calculate network congestion measurement value D.

In the current time step, after all nodes in the network are traversed, the current network congestion measure D is calculated according to formula (4), and the current time step is +1.

Step 6: Determine whether the simulation time is reached.

If the maximum simulation step T is not reached, Step 4– Step 5 are continued, and the network congestion measurement changes with time are recorded. Otherwise, the round simulation is ended to obtain the change of congestion measurement value with time.

Taking the change of network congestion measurement D with different values of alpha under $K = 0.001$ and $p = 0.05$ as an example, the simulation results are shown as follows (Fig. 2).

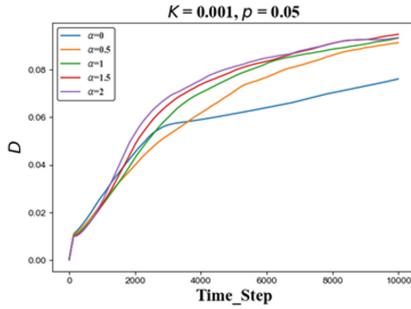


Fig. 2. Changes in a set of network congestion metrics

The regulation coupling coefficient that minimizes the degree of network congestion is calculated under different load parameters, and the result is shown in Fig. 3.

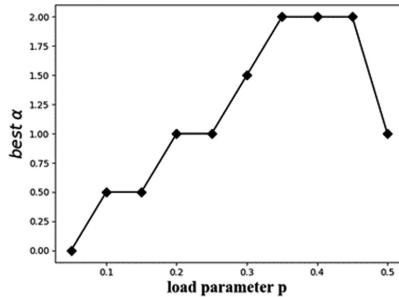


Fig. 3. Shows the regulation coupling coefficient that leads to the minimum degree of network congestion

In Fig. 3, under different load parameters, it is not that the greater the regulation coefficient is, the more favorable it is to alleviate network congestion. Based on the simulation results, the coupling coefficient α in the network can be adjusted according to different network processing capacity and initial load conditions, so as to minimize the degree of network congestion and alleviate the diffusion of network congestion.

Next, comparative verification was carried out. The control group was set as the model without considering regulation coupling relationship, that is, the special case that set the regulation coefficient $\alpha = 0$ (constant). At this time, the model in this paper is simplified to only consider the relationship between the capacity and load of the road section. By comparing the load and capacity of each road section in each time step, the congestion degree of the whole network is calculated, which has been studied in a large number of literatures such as literature [1]. Other initial parameters are guaranteed to be the same during the simulation. The result is shown in Fig. 4.

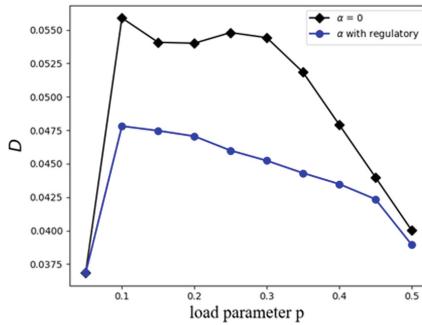


Fig. 4. The comparison of two types of models ($K = 0.001$)

In Fig. 4, compared with the model that does not consider regulation coupling relationship, the network congestion is alleviated to some extent after the coupling coefficient α is adjusted according to the simulation results of the model in this paper, and it also reflects the necessity of considering regulation coupling relationship. In this paper, the model can be used to study the relationship between network processing capacity, initial load conditions and congestion degree under different regulation coefficients, which shows that the model can reflect the congestion diffusion mechanism under the regulation coupling relationship. The conclusion of this paper can provide reference for the optimization of regulation strategy.

5 Conclusion

This paper proposes a new congestion diffusion model based on regulation coupling relationship. This paper compares the change of load on the traffic network section to the change of biomolecular concentration creatively, and then builds the congestion diffusion model of the traffic network by combining the regulation coupling characteristics in the traffic network. Finally, the traffic network in part of the city of Barcelona is taken as an example for analysis. The simulation results show that the model in this paper can help the traffic network control and alleviate the network congestion to a certain extent. This model can help analyze the process and mechanism of network congestion diffusion, provide theoretical support for alleviating the network congestion diffusion, and provide theoretical basis for proposing better regulation strategies.

Acknowledgments. This work is supported by the Beijing Natural Science Foundation (Grant No. L171002), the National Natural Science Foundation of China (Grant No. 61872018) and the National Natural Science Foundation of China (Grant No. 61773044).

References

- Crucitti P, Latora V, Marchiori M (2003) A model for cascading failures in complex networks
- Zhan X, Ukkusuri SV, Rao PSC (2017) Dynamics of functional failures and recovery in complex road networks. *Phys Rev E* 96(5):052301
- Parshani R, Buldyrev SV, Havlin S (2011) Critical effect of dependency groups on the function of networks. *Proc Natl Acad Sci* 108(3):1007–1010
- Su Z, Li L, Peng H, Kurths J, Xiao J, Yang Y (2014) Robustness of interrelated traffic networks to cascading failures. *Sci Rep* 4:5413
- Rui L, Wang X, Qiu X (2018) Analysis of cascade fault optimization based on regional fault and traffic reallocation in complex networks. *IEEE Access* 6:24818–24828
- Wang J, Jiang C, Qian J (2014) Robustness of Internet under targeted attack: a cascading failure perspective. Academic Press Ltd., Cambridge
- Li M, Liu RR, Jia CX, Wang BH (2013) Critical effects of overlapping of connectivity and dependence links on percolation of networks. *New J Phys* 15(9):093013
- Bai Y, Huang N, Wang L, Wu Z (2016) Robustness and vulnerability of networks with dynamical dependency groups. *Sci Rep* 6:37749
- Cruz-Piris L, Lopez-Carmona MA, Marsa-Maestre I (2009) Automated optimization of intersections using a genetic algorithm. *IEEE Access*, 1
- Olmos LE, Muñoz JD (2016) Traffic gridlock on a honeycomb city
- Sui QH, Ding ZJ, Jiang R, Huang W, Sun D (2012) Slow-to-start effect in two-dimensional traffic flow. *Comput Phys Commun* 183(3):547–551
- Zhao HT, Liu XR, Chen XX, Lu JC (2017) Cellular automata model for traffic flow at intersections in internet of vehicles. *Phys A: Stat Mech Appl* 494:40–51
- Pandit K, Ghosal D, Zhang HM, Chuah CN (2013) Adaptive traffic signal control with vehicular ad hoc networks. *IEEE Trans Veh Technol* 62(4):1459–1471
- Younes MB, Boukerche A (2016) Intelligent traffic light controlling algorithms using vehicular networks. *IEEE Trans Veh Technol* 65(8):5887–5899
- Liu W, Qin G, He Y, Jiang F (2017) Distributed cooperative reinforcement learning-based traffic signal control that integrates v2x networks' dynamic clustering. *IEEE Trans Veh Technol PP(99)*:1
- Younis O, Moayeri N (2017) Employing cyber-physical systems: dynamic traffic light control at road intersections. *IEEE Internet Things J* 4(6):2286–2296
- Tang C, Wei X, Hao M, Zhu C, Wang R, Chen W (2018) Traffic signal phase scheduling based on device-to-device communication. *IEEE Access* 6:47636–47645
- Liu J, Wan J, Jia D et al (2017) High-efficiency urban traffic management in context-aware computing and 5 g communication. *IEEE Commun Mag* 55(1):34–40
- Novák B, Tyson JJ (2008) Design principles of biochemical oscillators. *Nat Rev Mol Cell Biol* 9(12):981–991
- Zheng JF, Gao ZY, Zhao XM, Dai S, Fu BB (2010) Self-organized diffusion of congestion in complex networks. *Phys A* 389(2):342–348
- Joan B, Qian LY (2018) Barcelona: a full scale urban planning road. *Architect*, (1):31–41



Search-Based Motion Planning for MAVs

Decentralized Formation Control

Yu Zheng^{1(✉)}, Yunsong Ge¹, Longhong Wang¹, Xu Zeng²,
and Xuedong Zhao²

¹ Beihang University, Beijing 100191, China
yzheng@buaa.edu.cn

² Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Abstract. This paper presents a search-based motion planning method with motion primitives to solve the Decentralized Formation Control problems of Micro Aerial Vehicles (MAVs) in unknown environment. Given a team of N MAVs, our algorithm finds trajectories that navigate MAVs from their start positions to non-interchangeable goal positions (the labeled MAVs planning problem) in a collision-free manner. We use Lifelong Planning A^* (LPA^*) algorithm to leverage incremental search techniques for dynamic environment, and an additional graph pruning process is necessary to maintain the correctness and optimality of the planning results for searching with motion primitives. The proposed planner is shown to be efficient and its resolution complete and optimal. We demonstrate the feasibility of the generated formation control algorithm in simulations and real-world experiments.

Keywords: MAVs · Search-based · Motion primitives · Decentralized formation

1 Introduction

Micro Aerial Vehicle (MAV) teams have become attractive solutions for many challenging problems, such as collaborative search and rescue [1], delivery [2], and construction [3]. These settings require MAVs to have the fundamental capability of autonomously navigating amongst their neighbors. Many MAVs formation control algorithms have been presented in the past decades. Decentralized approaches can broadly be characterized as solutions for discrete or continuous workspaces. In the discrete domain, proposed algorithms have used sampling-based [4] planners to search the joint configuration space of the team. These approaches can guarantee safety, optimality, and completeness, however, their complexities grow exponentially with the number MAVs and quickly becomes computationally prohibitive. A decentralized formation flight control strategy via local information exchange and relative position control is presented [5], which consists of two parts, one is to synchronize the motion of UAVs, while the other is to achieve and maintain the formation geometry. [6] employs a unique extension-decomposition-aggregation scheme to transform the overall complex formation control problem into a group of sub-problems, which work via boundary interactions or disturbances. Alternatively, the problem has been formulated as a Mixed

Integer Program [7] and solved with optimization techniques. Other approaches, such rule-based methods [8] and decomposition into independent sub-problems [9], trade optimality for computational improvements

In this paper, we present a search-based motion planning method with motion primitives to solve the decentralized formation control problems of Micro Aerial Vehicles (MAVs) in unknown environment. The Lifelong Planning A^* (LPA^*) algorithm is used to leverage incremental search techniques for dynamic environment, and an additional graph pruning process is necessary to maintain the correctness and optimality of the planning results for searching with motion primitives. The proposed planner is shown to be efficient and its resolution complete and optimal. Figure 1 shows three MAVs formation control using our algorithm in real environment.



Fig. 1. Three MAVs formation control using our algorithm in real environment

2 Algorithms Description

2.1 Concurrent Assignment and Planning of Trajectories (CAPT)

In the unlabeled problem ($N_g = N$, MAV teams from their start positions to non-interchangeable goal positions), a team of N MAVs must navigate to N_g goals. However, it does not matter which MAV arrives at which goal, as long as all goals are visited. To Solve this problem, we need not only assign goals to MAVs but also generate safe trajectories to those goals. However, the coupling of assignment and trajectory planning paradoxically decreases the problem complexity.

Let $\varphi : A_N \rightarrow \mathbb{R}^2$ represent the goal assignment, where $\varphi^i = g^i$ indicates g^i is assigned to the MAV starting at s^i . CAPT finds the assignment φ that minimizes the sum of the distances squared traveled by all MAVs, explicitly:

$$\arg \min_{\varphi} \sum_{i=1}^N \|s^i - \varphi^i\|^2 \quad (1)$$

For first-order MAVs, the optimal trajectory between a start position, s_1 , and a goal position, g_1 is found by optimizing the cost functional:

$$\begin{aligned} \arg \min_{\phi} & \int_0^T \left\| \frac{d}{dt} \Phi^i(t) \right\|_2^2 dt \\ \Phi^i(0) &= s_1, \Phi^i(T) = g_1 \end{aligned} \quad (2)$$

The solution to Eq. (2) is

$$\Phi^i(t) = s_1 + \frac{t}{T}(s_1 - g_1) \quad (3)$$

Given a CAPT assignment and v_{\max} , the maximum time required for any MAV to reach its goal is given by:

$$T = \max_{i \in A_N} \frac{\|s^i - g^i\|_2}{v_{\max}} \quad (4)$$

2.2 Motion Primitives

The expression of the dynamical system is obtained:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad A = \begin{bmatrix} 0 & I_m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & I_m \\ 0 & \cdots & 0 & 0 \end{bmatrix}_{qm \times qm}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I_m \end{bmatrix}_{qm \times m} \quad (5)$$

and the control input is $u(t) \in U := [-u_{\max}, u_{\max}]^m \subset \mathbb{R}^m$. We consider a lattice discretization $U_M := \{u_1, \dots, u_M\} \subset U$, where each control vector $u_m \subset \mathbb{R}^3$ will define a motion of short duration τ for the system. Given an initial state $x_0 := [p_0^T, v_0^T, a_0^T, \dots]$, we generate a motion primitive of duration $\tau > 0$ and apply a constant control input $u(t) \equiv u_m \in U_M$ for $t \in [0, \tau]$ so that,

$$u(t) = p_D^{(n)}(t) = \sum_{k=0}^{K-n} d_{k+n} \frac{t^k}{k!} \equiv u_m \quad (6)$$

Because control input is constant, it applies that all coefficients that involve time need to zero, $d_{(n+1):K} = 0 \Rightarrow u_m = d_n$. And integrating the control input $u(t) = u_m$ with an initial condition x_0 results in

$$p_D(t) = u_m \frac{t^n}{n!} + \dots + a_0 \frac{t^2}{2} + v_0 t + p_0 \quad (7)$$

So the resulting trajectory of the linear time-invariant system in Eq. (2) is:

$$x(t) = \underbrace{e^{At}}_{F(t)} x_0 + \underbrace{\left[\int_0^t e^{A(t-\sigma)} B d\sigma \right]}_{G(t)} u_m \quad (8)$$

Because of the duration τ and control input are fixed, the cost of motion primitive according to Eq. (4) is $(\|u_m\|^2 + \rho)\tau$.

2.3 Incremental Trajectory Planning

In order to guarantee safety and optimality, it is necessary to update the planned trajectory when new information of moving MAVs is updated. Re-planning from scratch every time is not efficient since we may waste time searching places that are already explored in previous planning epochs. We use Lifelong Planning A^* (LPA^*) algorithm to leverage incremental search techniques for dynamic environment. For searching with motion primitives, an additional graph pruning process is necessary to maintain the correctness and optimality of the planning results. By combining LPA^* and graph pruning, we can efficiently solve the re-planning problem in a dynamic environment.

(a) LPA^* with motion primitives

It is computationally expensive to construct the graph in free space due to feasibility checking of edges. Therefore, if we are able to reuse the graph from the previous plan and only update costs of affected edges as the map changes, we are able to save a significant amount of time in the new planning query. LPA^* is popular incremental graph search algorithms that have been widely used in real-time re-planning. Therefore, LPA^* is a better choice for fast re-planning that plans the same optimal trajectory as A^* but expands much fewer states.

In LPA^* , once the graph has local inconsistency due to a map update, the algorithm will re-expand the affected states until all the locally inconsistent states become consistent. To update edge costs, we need to check the related edges in the graph according to the latest map. And this update is fast enough if the map is updated incrementally in a relatively small amount.

(b) Graph Pruning

When the MAV starts following the planned trajectory, the start state for the new planning moves to the successor of the previous start state along the planned trajectory. Now, since the start state for the graph has been changed, the start-to-state cost of all the graph MAVs need to be updated. Since our graph is directional (edges are irreversible), most of the existing graph that originated from the old start state becomes unreachable. Therefore, it is important to prune an existing graph and update the start-to-state cost of the remaining states according to the new start state. In this step, we do not need to explore new states, check for collision against map or calculate any heuristic values, thus it is very fast.

2.4 Collision Checking Between MAVs

Denote $c_i(t)$ as the i -th MAV configuration which is a non-linearly moving polyhedron (NMP) in \mathbb{R}^m , it is represented as the MAV geometry $c_{i,0} = \{p | A_i^T p \leq b_i\}$ that

centered at MAV's center of mass following a trajectory ${}^i\Phi(t)$. Thus, it can be represented as:

$$c_i(t) = \{p | A_i^T(p - {}^i\Phi(t)) \leq b\} \quad (9)$$

For MAV i and j , they are not colliding with each other if and only if

$${}^i\Phi(t) \cap [c_{i,0} \oplus c_j(t)] = \emptyset \quad (10)$$

where “ \oplus ” denotes the Minkowski addition. Constraint (10) can be verified by solving for roots of a polynomial equation. For a team of MAVs, we can verify whether the i -th MAV's trajectory is collision-free by checking (10) against all the other robots.

2.5 Decentralized Planning

In the decentralized case, each MAV re-plans at their own clock rate and there is no priority. It is practical to assume that the MAV is able to share information about its current trajectory with other MAVs. For accurate collision checking, we also assume there is a global time frame and a local time frame for each MAV representing its trajectory start time. Use τ and t to represent the time in global time frame and local time frame respectively. The conversion between this two frames is simply $t = \tau - \tau_s$ where τ_s is the start time in global time frame of the trajectory. Thus, we formulate the collision checking for Z MAVs in the decentralized manner as:

$${}^i\Phi(\tau - \tau_i^s) \cap \bigcup_{j=0}^{Z-1} [c_{i,0} \oplus c_j(\tau - \tau_j^s)] = \emptyset \quad (11)$$

or in local time frame as:

$${}^i\Phi(t) \cap \bigcup_{j=0}^{Z-1} [c_{i,0} \oplus c_j(\tau + \tau_i^s - \tau_j^s)] = \emptyset \quad (12)$$

Here $\tau_i^s - \tau_j^s$ is a constant for the given trajectory pair ${}^i\Phi$ and ${}^j\Phi$. Since we plan for MAV i with the presence of MAV j , $\tau_i^s - \tau_j^s \geq 0$ should always be true.

Denote jT as the duration of trajectory ${}^j\Phi$. Ideally, we use the whole trajectory ${}^j\Phi(t)$ from $t = 0$ to jT for collision checking when plan for MAV i . However, we know that MAV j is also constantly re-planning, such that the future trajectory of j can be meaningless to be considered in collision checking. We are able to improve the efficiency of collision checking in (12) by setting a cut-off time T_c . Namely, we ignore the part of trajectory ${}^j\Phi(t)$ of other MAVs j for the domain $t > {}^jT_c$.

Consequently, as T_c gets smaller, the computational time for inter-MAV collision checking is also smaller. The smallest T_c for a complete solution is determined by the system's dynamic constraints. For example, for a second order system that is constrained by maximum velocity v_{\max} and acceleration a_{\max} , the smallest value of T_c is

the minimum time it takes to stop the robot from the maximum velocity. Thus, we set the value of T_c for trajectory Φ_j as

$${}^j T_c = \min\{\bar{v}_{\max}/\bar{a}_{\max}, {}^j T\} \quad (13)$$

To make the algorithm complete, we need to ignore MAV j for $t > {}^j T_c$ when planning for MAV i instead of treating it as a static obstacle. Figure 2 and Fig. 3 shows the results of two planning tasks using the decentralized planning with T_c .

3 Analysis of Simulation Results

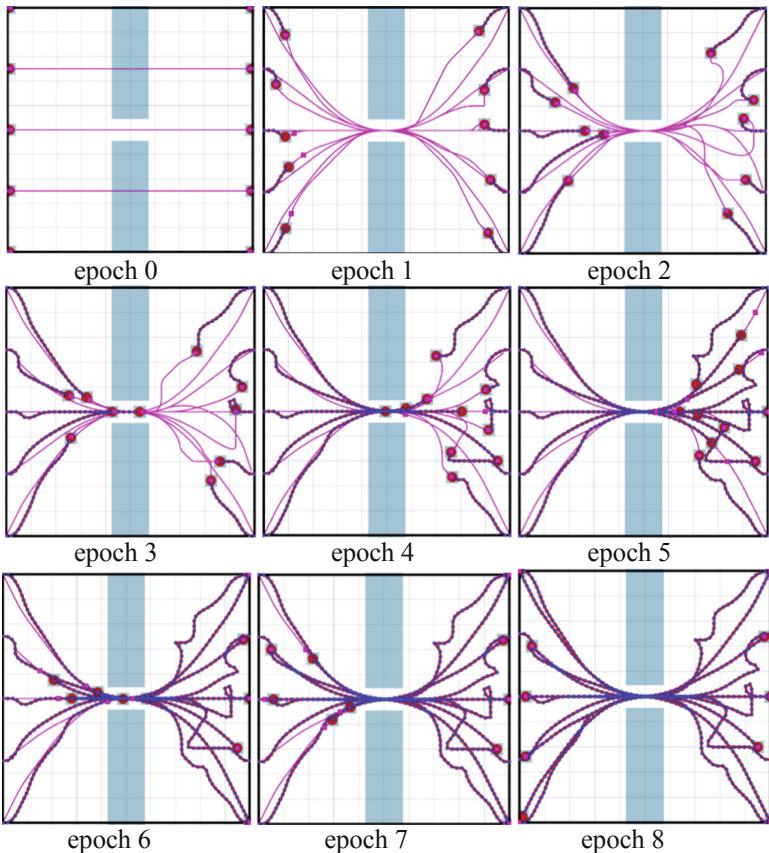


Fig. 2. Decentralized formation control in unknown narrow gap environment

Figure 2 shows the decentralized formation control in unknown narrow gap environment, and epoch 0 to epoch 8 demonstration the entire process of crossing the narrow gap. Each MAV has its own ID address, and the other MAVs are regarded as dynamic

obstacles. Using the principle that MAVs with smaller ID pass first can effectively improve the execution efficiency of the formation control task.

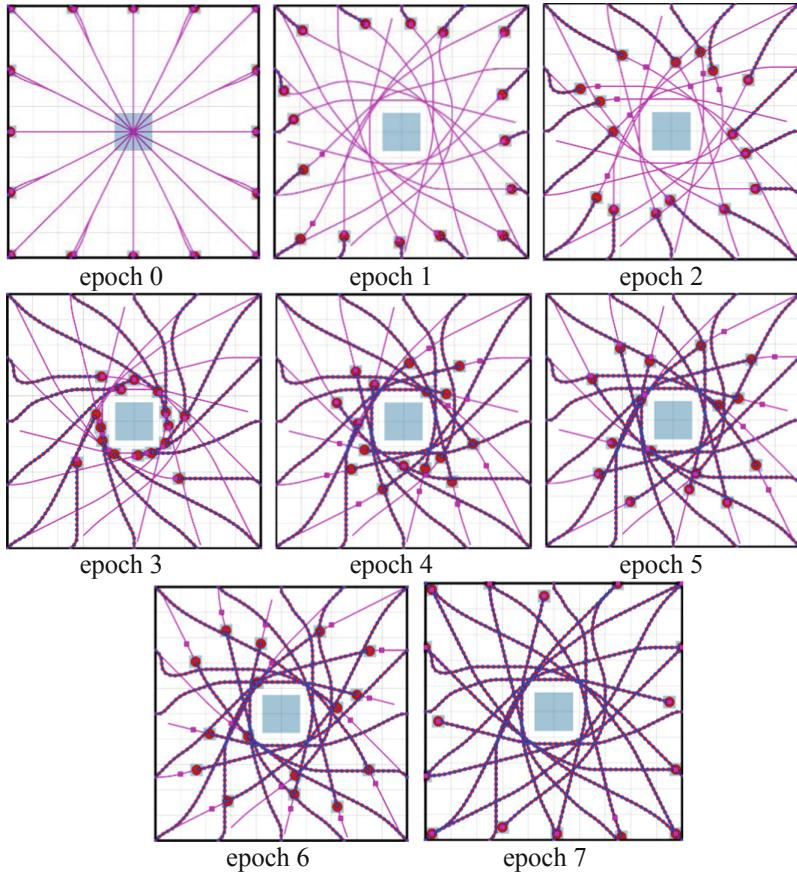


Fig. 3. Decentralized formation control in unknown obstacle environment

Figure 3 shows the decentralized formation control in unknown obstacle environment, and epoch 0 to epoch 7 demonstrate the entire process of obstacle modeling. By flying in an unknown obstacle scenario with a sufficient number of MAV teams, the shape of the unknown obstacle can be modeled effectively, and then extended to the construction of the complex scene with unknown obstacles, which can be quickly and effectively get the geometry information of unknown scene.

4 Analysis of Experimental Results

We develop Formation Control Software (FCS) for formation algorithms flight tests, and the status (position, velocity, attitude, etc.) of each MAV was obtained during the flight. The FCS has a strong security policy and can be re-planned in real time (Figs. 4, 5 and 6).



Fig. 4. Formation control software



Fig. 5. MAVs formation control tests in real environment

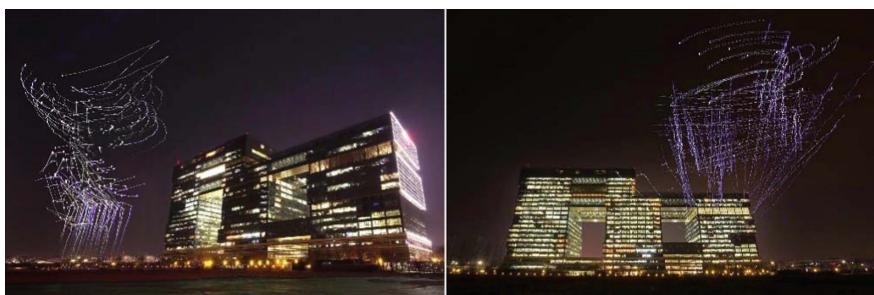


Fig. 6. Delayed flight path display of MAVs formation control tests using SLR camera

5 Conclusion

We use simulations and real-world experiments demonstrate the feasibility of the search-based motion planning method with motion primitives to solve the formation control problems of MAVs in unknown environment. And we extend to the construction of the complex scene with unknown obstacles, which can be quickly and effectively get the geometry information of unknown scene. We use Lifelong Planning A* (*LPA**) algorithm to leverage incremental search techniques for dynamic environment, and an additional graph pruning process is necessary to maintain the correctness and optimality of the planning results for searching with motion primitives. The proposed planner is shown to be efficient and its resolution complete and optimal.

References

1. Michael N et al (2012) Collaborative mapping of an earthquake-damaged building via ground and aerial robots. *J Field Robot* 29(5):832–841
2. Reisinger D (2016) Watch amazon's prime air complete its first drone delivery. *Fortune*. <http://fortune.com/2016/12/14/amazon-prime-air-delivery/>
3. Stewart J (2013) Google tests drone deliveries in project wing trials. *BBC News*. <http://www.bbc.com/news/technology-28964260>
4. Solovey K, Salzman O, Halperin O (2016) Finding a needle in an exponential haystack: discrete RRT for exploration of implicit roadmaps in multi-robot motion planning. *Int J Robot Res (IJRR)* 35:501–513
5. Yuan L (2012) Decentralized formation flight control of multiple fixed-wing UAVs with a virtual leader. In: Proceedings of the 31st Chinese control conference, Hefei, China
6. Yang A et al (2014) Stability analysis and implementation of a decentralized formation control strategy for unmanned vehicles. *IEEE Trans Control Syst Technol* 22(2):706–720
7. Yu J, LaValle SM (2013) Planning optimal paths for multiple robots on graphs. In: IEEE international conference on robotics and automation (ICRA), Karlsruhe, Germany
8. de Wilde B, ter Mors AW, Witteveen C (2013) Push and rotate: cooperative multi-agent path planning. In: International conference on autonomous agents and multi-agent systems (AAMAS), Saint Paul, MN, USA
9. Standley TS, Korf RE (2011) Complete algorithms for cooperative pathfinding problems. In: International joint conference on artificial intelligence (IJCAI), Barcelona, Spain



Researching the Monitoring Scheme Built for the Characteristic Parameters of Incineration Equipment

Fengguang Huang^{1,2,3}(✉), Lin Tian¹, Daoqing Hu¹, Wei Wang^{1,2,3},
Bosen Li¹, Qi Lu^{1,2,3}, Xiaolin Xing^{1,2,3}, Xi Chen^{1,2,3},
Xiaoman Zhai^{1,2,3}, and Yansha Luo^{1,2,3}

¹ 101 Institute of the Ministry of Civil Affairs, Beijing 100070, China
huangfengguang@163.com

² Key Laboratory of Pollution Control, Ministry of Civil Affairs,
Beijing 100070, China

³ Institute of the Ministry of Civil Affairs Environmental Monitoring Center
Station, Beijing 100070, China

Abstract. By studying the special characteristics of incinerators, we have developed an on-line monitoring scheme based on PHP and MYSQL technology for incineration equipment. This scheme is capable of immediately monitoring and analyzing the parameters of flue gas, fan, fuel injector and other related parameters tested by incineration detection equipment installed in the incinerator. Moreover, the scheme can remotely send control instructions for the on-line incinerator equipment. The implementation and application of this scheme will increase the efficiency of monitoring on-line incineration equipment and analyzing relevant data.

Keywords: Incineration equipment · Real-time monitoring · Remote control

1 Introduction

Incineration equipment is commonly used in funeral services, but on the other hand, it is the main source of air pollution. With the government having increasing requirements for environmental protection, researchers have been paying more and more attention to the environmental pollution caused by incineration equipment, and placing a focus on the research of the flue gas characteristics, and the running and operating characteristics of key equipment components. Some research findings were obtained through using portable and fixed detection equipment, as these characters were tested, collected and analyzed.

With the continuous development of the Internet of Things (IoT) in the traditional industries in China, the applications of IoT are becoming more and more popular. Researchers have attempted to introduce IoT technology into the monitoring of incineration equipment over the past years. It is found that by using IoT technology to upload the parameters of incineration equipment, which are collected by the detection equipment, to the network server where the parameters are stored, downloaded and

remotely analyzed, the collection and analysis efficiency of the incineration equipment characteristics will be improved. However, at present, there is still a lack of analysis of the collected parameters based on the online platform and the on-line control of the incineration equipment based on real-time data.

Therefore, aiming at an immediate monitoring and real-time analysis of flue gas parameters, fan parameters and relevant parameters of the injection machine tested by incineration detection equipment, we have developed an on-line monitoring scheme, which addresses the characteristics of various parameters detected by the incinerator's detection equipment. Through this scheme, the control instruction can be sent for the incinerator's on-line equipment, after which the remote control of the equipment.

2 Designing the Scheme and the Interface

2.1 The Scheme Design

By studying the working characteristics of incineration detection equipment, we have constructed a B/S system architecture scheme [1] by using PHP and MYSQL technology, and archived the purpose of the remote, immediate monitoring and real-time analysis of the parameters detected by the detection equipment through the WEB or INTERNET browser. The user access interface is intuitionistic, concise, and has strong manoeuvrability [2]. The structure of the monitoring scheme and the access framework are shown in Fig. 1.

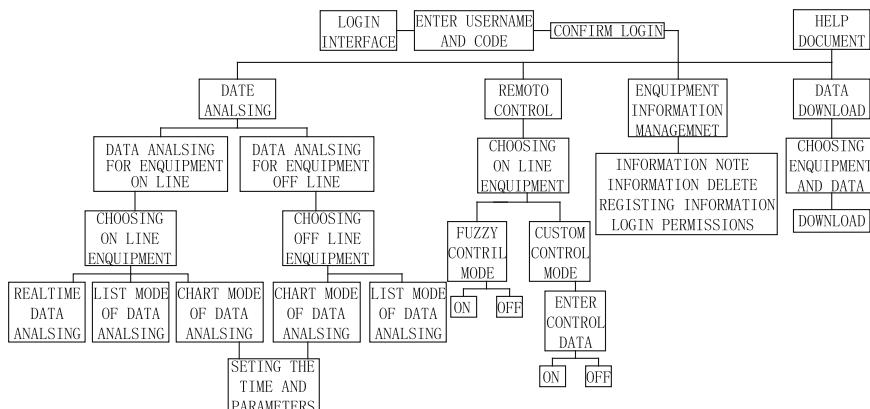


Fig. 1. A basic layout of the project's system structure and accesses

2.2 The Interface Design

Login Interface and Home Page: On the login page, users need to enter the user name and login password for identity confirmation, and then click “Login” to access the home page of the monitoring system (shown in Fig. 2). The home page provides

operational functions such as data analysis, remote control, device list, and information base. The device list enables the equipment management functions, and the information base enables the data downloading and the user guide.



Fig. 2. The login interface and the home page

Data Analysis: Data analysis can help users to analyze the incineration equipment's data. The data analysis management covers on-line monitoring and off-line analysis. You may click on the "Data analysis" interface, and select on-line monitoring or off-line analysis to check the equipment. Then, the real-time flue gas data such as SO₂, CO, O₂, NO, and parameters of the fan, fuel injection and other equipment can be viewed. The real-time data analysis interfaces are shown in Fig. 3. The system data is collected with a frequency of 5 s per data group. In case of connection delay, the server will mark the received data by the time that it received the most recent data. In case the field sensor does not function, it will pass down the "abnormal" data to the server. The operators may trace this "abnormal" data for the location and the type of the sensor, before repairing the sensor.

Remote Control: Users can remotely control the incineration equipment by using this function. Users can click the interface of remote control, select the device needs to be controlled in a list of on-line devices, get into the page, and send the control instruction to the incineration equipment by starting the remote fuzzy control or the remote custom control mode. The real-time parameters collected in the field are displayed on the right side of the chart, which is convenient for users to view the feedback information of the remote control. The main interface of the remote control is shown in Fig. 4.

Other Functions: Besides these functions, the scheme covers the device list, information base, etc. Using the list of devices, users can edit and complete the equipment information, such as device number and the address of online and offline devices. On the information base interface, users can add or remove the device from the list, and can add the user name, the real name, the contact information and notes to the identity database.

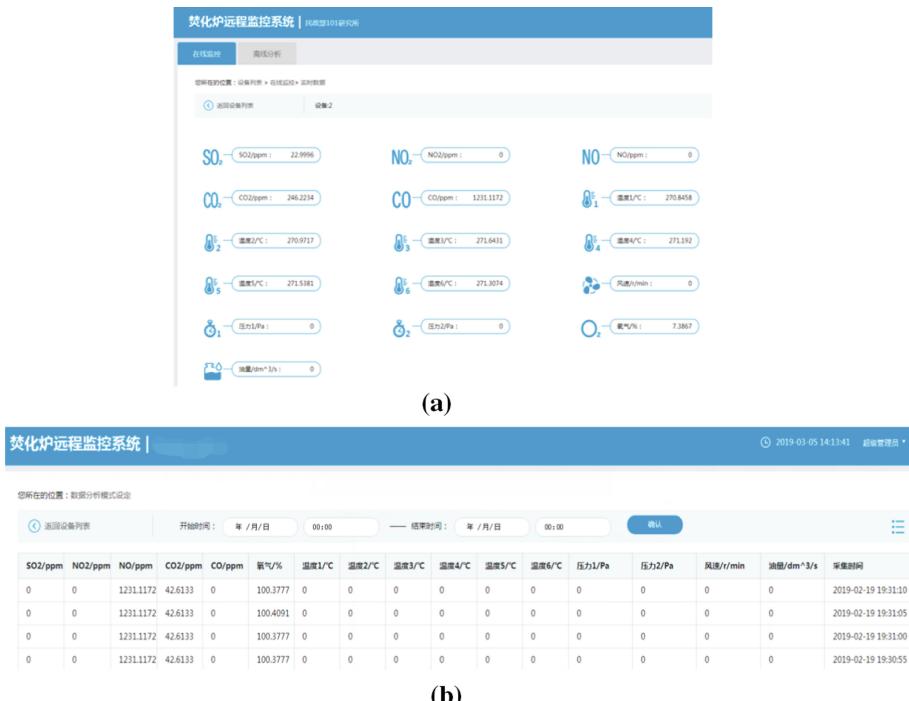


Fig. 3. (a) Data analysis (On-line). (b) Data analysis (Off-line)

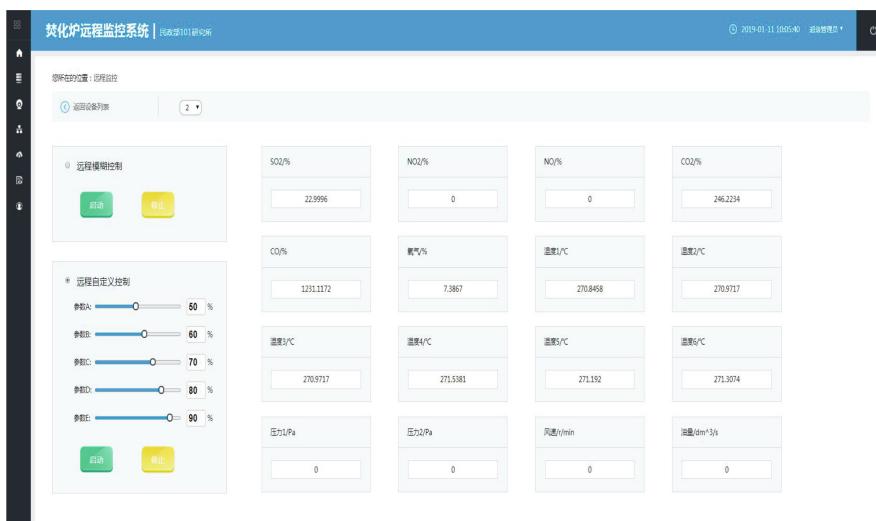


Fig. 4. Remote control interface

3 The Implementation of Key Technologies

3.1 Database

By studying the working characteristics of incineration detection equipment, we have constructed a B/S system architecture scheme by using PHP and MYSQL technology, before achieving the purpose of a real-time analysis [3]. The database design has considered the data information, such as the monitoring interface data and the logical operation data. The monitoring interface data refers to the monitoring interface which is based on PHP and MYSQL technology, and defines the main function types and structure layout of the detection scheme. Logical operation data is mainly used for the communication between the interface layer and the database layer, and defining logic level, operation process, order and other logical characteristics [4]. The main procedures for database implementation are as follows.

```
* @param Request $request [description]
* @return [type]      [description]
public function test(Request $request) {
    $data = $request->param();
    // $data1['data'] = $data['data'];
    $data1['number'] = $data['you'];
    $data1['time'] = time();
    // Db::name('test')->insert($data1);
    $str = substr($res, 0, -2); // Remove the last two bits of the character
    $str1 = str_split($str, 8); // Divide characters into a unit per 8 bits
    foreach ($str1 as $key => $value) {
        $arr[] = implode("", array_reverse(str_split($value, 2))); // Character inversion
        foreach ($arr as $key => $value) {
            $v = hexdec($value);
            $x = ($v & ((1 << 23) - 1)) + (1 << 23) * ($v >> 31 | 1);
            $exp = ($v >> 23 & 0xFF) - 127;
            $arr1[] = round($x * pow(2, $exp - 23), 4);
        }
    }
    // Transfer data to groups and into a database
}
```

3.2 Remote Control

Remote control function enables the remote control of the on-line devices [5]. Fuzzy control mode and custom control mode can be used for starting or stopping the incineration equipment. Custom control parameters can be entered into custom control mode, after which the incinerator equipment may be turned on or off. The main procedures for functional implementation are shown as follows.

```
/** * Remote Fuzzy control
 * @return [type] [description]
 */
public function client() {
    $num = input('num');//Device number
    $start = input('start');//Determine by parameter to start or stop 1 for startup 2 for Stop
    if($start == 1) { $para='55aa1100'; } else
        Gateway::$registerAddress = '127.0.0.1:1238';
    $client=Gateway::getClientIdByUid($num);
    $client1 = $client[count($client)-1];
    Gateway::sendToClient($client1,$para);
    return 1; } /** * Remote Custom Control
 * @return [type] [description] */
public function client1() { $num = input('num');//Device number $start = input (' start');//by parameter
to determine whether to start or stop 1 as a percentage of startup 2 for stop//Receive send parameters
$sa = sprintf("%03d", input('a'));
$b = sprintf("%03d",input('b'));
$c = sprintf("%03d",input('c'));
$d = sprintf("%03d",input('d'));
$e = sprintf("%03d",input('e'));
$foreach ($arr as $key => $value) {
    $para='55aa'.'.$num1.$value;
    Gateway::$registerAddress = '127.0.0.1:1238';
    $client=Gateway::getClientIdByUid($num);
    $client1 = $client[count($client)-1];
    Gateway::sendToClient($client1,$para);
    usleep(200000);//200 milliseconds execution
    $num1++; } } else
```

4 An Analysis of the Experimental Results

With the continuous development of funeral services in China, the real-time monitoring and the remote access system of flue gas and equipment operation data have gradually become the focus of the industrial research. In this paper, a monitoring scheme of online incineration equipment is developed. It uses PHP and MYSQL technology, and the immediate monitoring and expansion analysis of the relevant characteristics of flue gas parameters, fan parameters, fuel injector parameters and other incineration equipment characteristics, which are examined by the incinerator's detection equipment, are collected, and the remote control of the equipment is enabled. The paper will provide a reference support for the research of the real-time monitoring system of the incineration equipment.

References

1. Eve S, Hu W, Yang Z (2014) Design and implementation of sales management system based on PHP and Mysql. Comput Knowl Technol, (23)
2. Yu Y (2011) PHP typical module and project actual combat. Tsinghua University Press, Beijing
3. Han Y (2013) Research on student achievement management system based on PHP + MYSQL. Comput Knowl Technol, (32)
4. Showeming (2009) Website development based on PHP + MySQL. Logist Eng Manag, (06)
5. Ball D (2004) Implementation and application of MySQL database operation class. Comput programming skills and maintenance 2004 Phase 12



Finite-Time Consensus Control for the Second-Order Multi-agent Systems with Communication Delays

Yan Cui and Qinghua Li^(✉)

College of Physics and Information Engineering, Shanxi Normal University,
Linfen 041000, Shanxi, China
574383675@qq.com

Abstract. This paper is discussed the finite-time consensus control problem with communication delays for the second-order multi-agent systems. In order to achieve the consensus of systems with communication delays in finite time, a novel control algorithm with consensus control gain matrix is designed, which is based on the relative position and velocity error sates. By utilizing the stability theory of functional differential equations and the finite-time Lyapunov stability theorem, finite-time consensus of the second-order systems is realized. Furthermore, the allowable upper bound of the communication delays for the multi-agent systems is obtained. Finally, simulation examples verify the correctness and validity of the theoretical results.

Keywords: Finite-time consensus · Second-order multi-agent systems · Communication delays

1 Introduction

The consensus of network is a fundamental distributed control and optimization problem. It is to develop distributed control protocols that enable a group of agents to reach a consensus on certain quantities of interest [1–3]. It also has broad applications, such as unmanned air vehicles formation flying [4, 5], cooperative control of the satellites system [6, 7] and the distributed sensor network [8–10].

The finite time consensus theory of multi-agent has investigated in many literatures [11–13]. In [11], it discussed the finite-time consensus problem for the multi-agent systems with external disturbances. In [12], it studied the finite-time consensus problem of multi-agent systems with continuous time varying interaction topology. In [13], the leader-following consensus problem for first-order integrator dynamics with noise perturbation and communication delays was studied.

Recently, consensus problems for the second-order multi-agent systems with communication delays have been researched in some literatures [14–18]. In [14], a new kind of distributed control protocol was designed, which based on the adaptive intermittent control ideas for the nonlinear second-order multi-agent systems with mixed time-delays. In [15], the consensus problem of a non-leading second-order multi-agent system with non-uniform delay was studied. In [16], based on the relationship between

the root of the characteristic equation, the consensus problem of a series of general second-order multi-agent systems with communication delays were discussed. In [17], the network topology with the directed spanning tree could be reach consensus when the time delay was less than the threshold. [18] proposed a distributed algorithm to drive all agents to achieve a desired formation for systems with nonuniform time-delays. In [19], the finite time adaptive problem of for the first-order neural networks with time-delays and unknown parameters was solved.

Actually, actual networks with communication delays is generally present. In addition, in the second-order multi-agent systems, there are few literatures to research systems with communication delays in finite time. Therefore, it is of great significance to research the finite time consensus of the second-order multi-agent systems with delays. Then we design a control algorithm with consensus control gain matrix, which is based on the relative position and velocity error sates in this paper.

The rest of this article is organized as follows. Section 2 gives the preliminaries and problem. The main results are discussed in Sect. 3. In Sect. 4, examples are provided to show the effectiveness of the results. Conclusions are stated in Sect. 5.

2 Preliminaries and Problem Formulation

The multi-agent systems consisting of a leader and N followers are considered, where the leader is 0, the followers are 1, 2, ..., N and the network topology composed of N followers is noted as an undirected graph $\bar{G} = (\bar{V}, \bar{E}, \bar{A})$. The leader and N followers are described by an undirected graph $G = (V, E, A)$, among them $V = \{0, 1, 2, \dots, N\}$, $A = a_{ij} \in R^{(N+1) \times (N+1)}$, $i, j = 0, 1, \dots, N$, $\bar{V} = \{0, 1, 2, \dots, N\}$, $\bar{A} = a_{ij} \in R^{(N+1) \times (N+1)}$, $i, j = 1, 2, \dots, N$. $D \in R^{n \times n}$ is a diagonal matrix, which represents the diagonal weight of the communication between the follower and the leader, the diagonal element is d_i , that is $D = \text{diag}\{d_1, d_2, \dots, d_N\}$, $d_i = a_{i0}$. $d_i > 0$ if and only if leader 0 is the neighbor state of follower i .

In this article, all followers of the multi-agent systems are studied, which move on the same plane, so the dynamic system satisfies

$$\begin{cases} \dot{x}_i(t) = Ax_i(t) + Bv_i(t) \\ \dot{v}_i(t) = Cv_i(t) + Du_i(t) \end{cases}, t \geq 0, i = 1, 2, \dots, N \quad (1)$$

where $x_i(t) \in R$ is the position, $v_i(t) \in R$ is the velocity and $u_i(t) \in R$ is the control input of the i -th agent, $A \in R^{n \times n}$ is the position matrix, $B \in R^{n \times n}$ and $C \in R^{n \times n}$ are the velocity matrixes. $D \in R^{n \times n}$ is the control input matrix.

The dynamic equation of the leader is as follows

$$\begin{cases} \dot{x}_0(t) = Ax_0(t) + Bv_0(t) \\ \dot{v}_0(t) = Cv_0(t) \end{cases} \quad (2)$$

where $x_0(t) \in R$ is the position of the leader, $v_0(t) \in R$ is the velocity of the leader.

Converted (1) (2) to matrix form as follows

$$\begin{bmatrix} \dot{x}(t) \\ \dot{v}(t) \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ D \end{bmatrix} u, \quad \begin{bmatrix} \dot{x}_0(t) \\ \dot{v}_0(t) \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} \quad (3)$$

Let $\xi(t) = [x_1(t), x_2(t), \dots, x_n(t), v_1(t), v_2(t), \dots, v_n(t)]^T$
Thus the system model (1) becomes

$$\dot{\xi}_i(t) = M\xi_i(t) + Nu_i(t) \quad (4)$$

The leader's model (2) becomes

$$\dot{\xi}_0(t) = M\xi_0(t) \quad (5)$$

$$\text{where } M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}, N = \begin{bmatrix} 0 \\ D \end{bmatrix}$$

In order to analyze consensus of the second-order multi-agent systems with communication delays in finite time, which is based on the position and the velocity error states, a consensus protocol with a control gain matrix is designed, so the position and the velocity of the system can reach consensus in finite time. Then the distributed control protocol as follow

$$\begin{aligned} u_i(t) = K & [\sum_{j \in N_i} a_{ij}(x_j(t - \tau) - x_i(t - \tau) + v_j(t - \tau) - v_i(t - \tau)) \\ & + d_i(x_0(t - \tau) - x_i(t - \tau) + v_0(t - \tau) - v_i(t - \tau))] \end{aligned} \quad (6)$$

where $K \in R^{n \times n}$ is a consensus gain matrix, the design method will be given later.

3 Main Results

Denoted the position and the velocity error as follows

$$\begin{cases} e_{xi}(t) = x_i(t) - x_0(t) \\ e_{vi}(t) = v_i(t) - v_0(t) \end{cases} \quad (7)$$

Converted (7) as a matrix $e_i = [e_{xi} \ e_{vi}]^T$

Inserting the control protocol (6) into (4) (5) and rewriting the error state form

$$\dot{e}_i(t) = Me_i(t) + NK[\sum_{j \in N_i} a_{ij}(e_j(t - \tau) - e_i(t - \tau)) - d_i e_i(t - \tau)] \quad (8)$$

Note $e(t) = [e_1^T(t), e_2^T(t), \dots, e_n^T(t)]^T$

Rewriting (8) in matrix, we get

$$\dot{e}(t) = (I_n \otimes M)e(t) - (\mathbf{H} \otimes MK)e(t - \tau) \quad (9)$$

where $H = L + D$. L is the laplace matrix. D is the communication weight matrix.

Definition 1 [20]. The leader-following consensus of system (4) (5) is said to be achieved in finite-time if, for each follower $i \in \{1, 2, \dots, N\}$, existing non-negative real number $t^* > 0$, $\varepsilon > 0$, $\eta > 0$, there is a local state feedback u_i of $\{\xi_j : j \in N\}$, the error state of closed-loop system satisfies

$$\|e_i(t)\| \leq \eta e^{-\varepsilon(t-t^*)}, \forall t > t^* \quad (10)$$

for any initial condition $\xi_i(s) = \xi_i(0)$, $s \in (-\infty, 0]$, $i = 0, 1, \dots, N$

Several lemmas are given for the finite time consensus analysis.

Lemma 1 [20].

- (a) The matrix H is a non-negative symmetric matrix and has nonnegative eigenvalues.
- (b) The matrix H is a positive definite if and only if the graph G is strongly connected.

Lemma 2 [21]. Let $V(t) > 0$ for $t \in R$, $\tau \in [0, \infty)$ and $t_0 \in R$. Suppose $\dot{V}(t) \leq -aV(t) + b(\sup_{t-\tau < s < t} V(s))$ for any $t > t_0$. If $a > b > 0$, then $V(t) \leq \sup_{-\tau \leq s \leq 0} V(t_0 + s)e^{-\gamma(t-t_0)}$, $\forall t > t_0$, where $0 < \gamma < a$ is determined by $\gamma - a + be^{\gamma\tau} = 0$.

Assumption 1. The pair (M, N) is stabilizable.

Assumption 2. The network topology G is strongly connected.

Theorem 1. Considering the system (4) (5) under the control protocol (6). Suppose Assumptions 1 and 2 are satisfied and communication delay is satisfied $\tau \leq \tau^*$, the state of all the followers can achieve the consensus tracking control in finite time.

$$\tau^* = \frac{\sqrt{\beta(\kappa - \beta\lambda_n^2)\lambda_{\min}(Q)}}{\sqrt{2\lambda_{\max}(Q) \times \|(\mathbf{Q}NN^TQ)^2\|} \times \sqrt{(\|M^TM\| + \|\mathbf{H}^T\mathbf{H} \otimes (NN^TQ)^T(NN^TQ)\|)}} \quad (11)$$

Proof. Under Assumption 2, the network topology G is strongly connected. H is a positive definite matrix. Let $\kappa = \min\{\lambda_i(H), i = 1, 2, \dots, N\}$, the eigenvalue of H is $\lambda_i(i = 1, 2, \dots, N)$. Since (M, N) is stabilizable, there exists a solution $Q > 0$ to the following Riccati inequality

$$QM + M^TQ - 2\kappa QNN^TQ + \kappa I_n = 0 \quad (12)$$

The matrix K in (6) is constructed as

$$K = N^T Q \quad (13)$$

According to Leibniz-Newton formula

$$e(t - \tau) = e(t) - \int_{t-\tau}^t \dot{e}(s)ds \quad (14)$$

Rewritten the error dynamic system (9) to

$$\dot{e}(t) = (I_n \otimes M - H \otimes NK)e(t) + (H \otimes NK) \int_{t-\tau}^t \dot{e}(s)ds \quad (15)$$

Let $Q > 0$ be a solution of Riccati (12). Consider the Lyapunov function candidate

$$V(t) = e^T(t)(I_n \otimes Q)e(t) \quad (16)$$

The time derivative of this Lyapunov candidate along the trajectory of system (9) is

$$\begin{aligned} \dot{V}(t) &= e^T(t)[(I_n \otimes Q)(I_n \otimes M - H \otimes NK) + (I_n \otimes M^T - H \otimes K^T N^T) \\ &\quad (I_n \otimes Q)]e(t) + 2 \int_{t-\tau}^t e^T(t)(I_n \otimes P)(H \otimes NK)\dot{e}(s)ds \\ &\leq e^T(t)[(I_n \otimes (QM + M^T Q))e(t) - e^T(t)[H \otimes (QNK + K^T N^T Q)]e(t) \\ &\quad + \beta e^T(t)(H^T H \otimes I_n)e(t) + \frac{1}{\beta}[(I_n \otimes QNK) \int_{t-\tau}^t \dot{e}(s)ds]^T \\ &\quad \times [(I_n \otimes QNK) \int_{t-\tau}^t \dot{e}(s)ds] \end{aligned} \quad (17)$$

where $0 < \beta < \frac{\kappa}{\lambda_{\max}(H^T H)}$

Since H is symmetric, there exists an orthogonal matrix $\phi \in R^{N \times N}$, such that

$$\phi^T H \phi = \Gamma \equiv \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \quad (18)$$

where $\lambda_1, \dots, \lambda_N$ are the eigenvalues of matrix H . $\phi_i^T H = \lambda_i \phi_i^T$. According to Lemmas 1, 2 and Assumption 2 ensure that $\lambda_i > 0$.

Setting $\hat{e}(t) \triangleq [\hat{e}_1^T(t), \hat{e}_2^T(t), \dots, \hat{e}_n^T(t)]^T = (\phi \otimes I_n)e(t)$, (17) becomes

$$\begin{aligned} \dot{V}(t) &\leq \hat{e}^T(t)[(I_n \otimes (QM + M^T Q)) - \Gamma \otimes (2QNN^T Q)]\hat{e}(t) + \beta \lambda_n^2 \hat{e}^T(t)\hat{e}(t) \\ &\quad + \frac{1}{\beta}[(I_n \otimes QNN^T Q) \int_{t-\tau}^t \dot{e}(s)ds]^T \times [(I_n \otimes QNN^T Q) \int_{t-\tau}^t \dot{e}(s)ds] \\ &\leq \hat{e}^T(t)[(I_n \otimes (QM + M^T Q)) - \Gamma \otimes (2QNN^T Q)]\hat{e}(t) + \beta \lambda_n^2 \hat{e}^T(t)\hat{e}(t) \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{\beta} \| (QNN^T Q)^2 \| \int_{t-\tau}^t \int_{t-\tau}^t \dot{e}(s) \dot{e}(u) ds du \\
& = \sum_{i=1}^n \hat{e}_i^T(t) [QM + M^T Q - 2\kappa QNN^T Q] \hat{e}_i(t) \\
& + \beta \lambda_n^2 \hat{e}^T(t) \hat{e}(t) + \sigma \sup_{t-2\tau \leq s \leq t} e^T(s) e(s) \\
& \leq -(\kappa - \beta \lambda_n^2) \sum_{i=1}^n \hat{e}_i^T(t) \hat{e}_i^T(t) + \sigma \sup_{t-2\tau \leq s \leq t} e^T(s) e(s) \\
& = -(\kappa - \beta \lambda_n^2) e_i^T(t) e_i^T(t) + \sigma \sup_{t-2\tau \leq s \leq t} e^T(s) e(s) \\
& \leq -aV(t) + b \sup_{t-2\tau \leq s \leq t} V(s)
\end{aligned} \tag{19}$$

Thus $\dot{V}(t) \leq -aV(t) + b \sup_{t-2\tau \leq s \leq t} V(s)$, where

$$\begin{aligned}
\sigma &= \frac{2h^2}{\beta} \| (QNN^T Q)^2 \| \times (\|M^T M\| + \|\mathbf{H}^T \mathbf{H} \otimes (NN^T Q)^T (NN^T Q)\|) a = \frac{\kappa - \beta \lambda_n^2}{\lambda_{\max}(Q)} b \\
&= \frac{\sigma}{\lambda_{\min}(Q)} \\
h &< \frac{\sqrt{\beta(\kappa - \beta \lambda_n^2) \lambda_{\min}(Q)}}{\sqrt{2\lambda_{\max}(Q) \times \| (QNN^T Q)^2 \|} \times \sqrt{(\|M^T M\| + \|\mathbf{H}^T \mathbf{H} \otimes (NN^T Q)^T (NN^T Q)\|)}}
\end{aligned} \tag{20}$$

According to the Lemma 2, for any $t \geq t_0$, there is a positive real number γ

$$V(t) \leq \sup_{-2\tau \leq s \leq 0} V(t_0 + s) e^{-\gamma(t-t_0)} \tag{21}$$

Thus further obtaining the following formula

$$\|e(t)\| \leq \left(\frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)} \right)^{\frac{1}{2}} \sup_{-2\tau \leq s \leq 0} \|e(t_0 + s)\|^2 e^{-\frac{\gamma(t-t_0)}{2}} \tag{22}$$

According to the Definition 1, it is known that the agent state of all followers can achieve consensus in finite time.

Remark. The upper bound of the delay is more conservative, which ensures that all followers can track the dynamic leader in finite time.

4 Examples and Simulation

We assume that the network topology of the multi-agent system contains 4 nodes. The leader $x_0(t)$ is a global reachable node. Multi-agent topology is fixed and undirected, see Fig. 1. All agents move on the same plan. L , D and H as follows

$$L = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad H = \begin{bmatrix} 2 & 0 & -1 & 0 \\ 0 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

Then calculate the eigenvalue of H $\lambda_{\min} = 0.3820$, $\lambda_{\max} = 3.6180$

$$\text{Assume } A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

$$N = \begin{bmatrix} 0 \\ D \end{bmatrix}$$

Solving (12), get the solution of the Riccati equation as follow

$$Q = \begin{bmatrix} 0.3397 & -0.1840 & 0.1215 & 0.3040 \\ -0.1840 & 0.8744 & -0.3226 & 0.1350 \\ 0.1215 & -0.3226 & 0.3845 & 0.0075 \\ 0.3040 & 0.1350 & 0.0075 & 0.4315 \end{bmatrix} \quad K = N^T Q = [0 \quad 0 \quad 0 \quad 0.4315]$$

The initial state of four followers and one leader are $x_0 = (1, 0)^T$, $x_1 = (8, 0)^T$, $x_2 = (7, 0)^T$, $x_3 = (0, 5)^T$, $x_4 = (0, 3)^T$; $v_0 = (1, 0)^T$, $v_1 = (2, 0)^T$, $v_2 = (3, 4)^T$, $v_3 = (0, 6)^T$, $v_4 = (7, 8)^T$, let $\beta = 0.025$. According to the Theorem 1, the maximum communication delay that the follower system can tolerate under the consensus control protocol $h = 0.0030$.

Figures 2 and 5 describe the simulation curve of the trajectories of the position and the velocity with a leader with a delay of $\tau = 0.001$, where a plane composed of x-axis and y-axis, and the time axis is orthogonal to the x-y plane. Figures 3 and 4 respectively show the position errors of the x and y-coordinate between the follower and the leader. Figures 6 and 7 are the velocity errors between the follower and leader.

From all figures we can see, as long as the communication delay between the agents does not exceed the maximum allowable delay upper bound, the followers can track the dynamic leader in finite time.

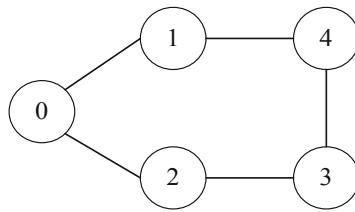


Fig. 1. Network topology of multi-agent system with a leader

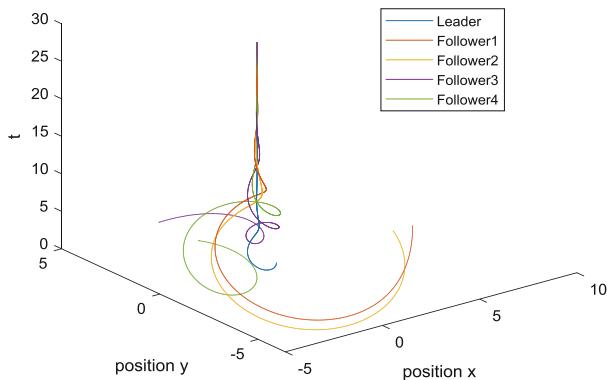


Fig. 2. Position state trajectories of five agents in the x-y plane

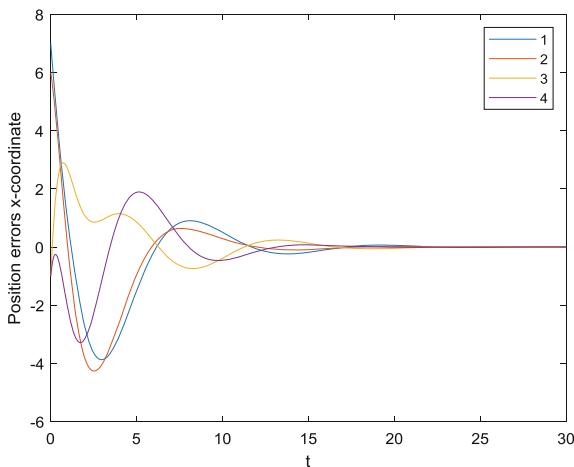


Fig. 3. Position errors between the leader and the following agents x-coordinate

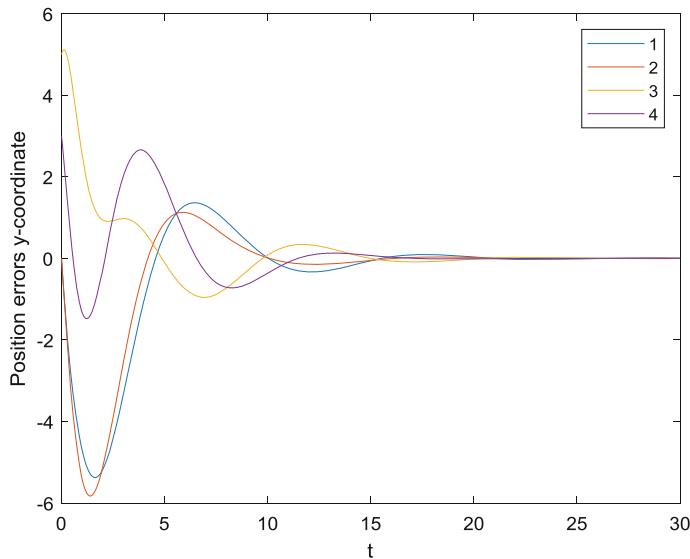


Fig. 4. Position errors between the leader and the following agents y-coordinate

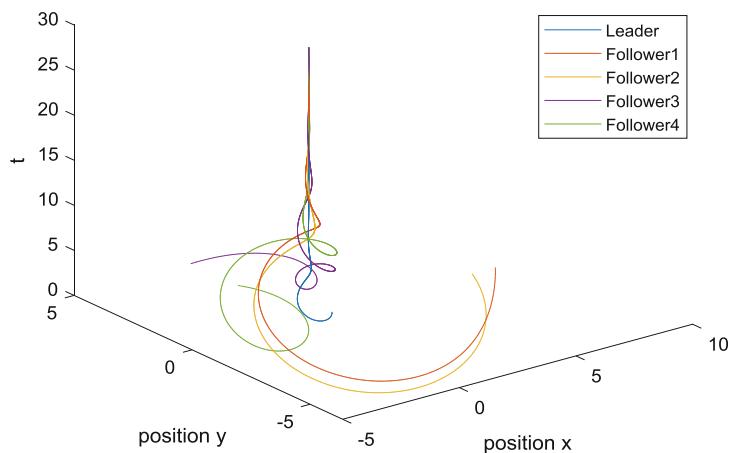


Fig. 5. Velocity state trajectories of five agents in the x - y plane

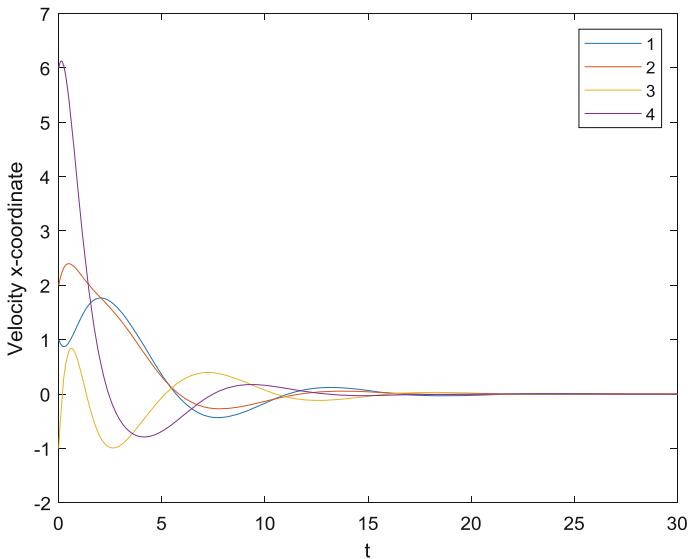


Fig. 6. Velocity errors between the leader and following agents x-coordinate

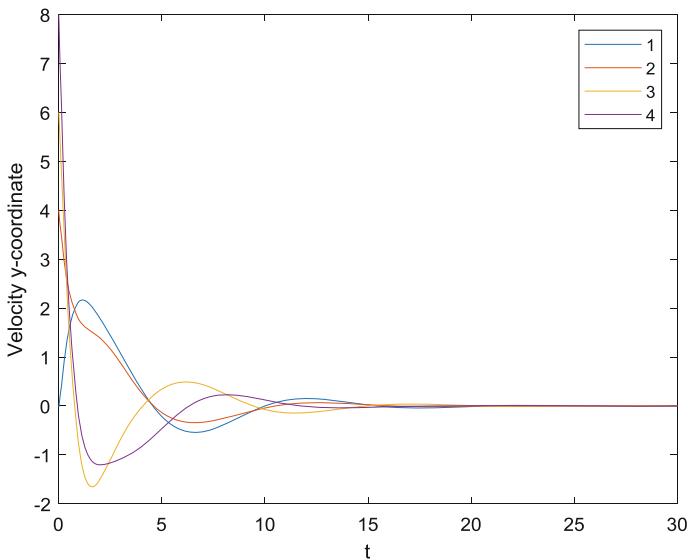


Fig. 7. Velocity errors between the leader and following agents y-coordinate

5 Conclusions

The leader-following finite-time consensus problem of the second-order multi-agent systems with communication delays is investigated by introducing a control gain matrix and designing the corresponding control algorithm. The upper bound of delay is

obtained for the consensus problem. In the future work, the finite-time consensus problem of discrete multi-agent systems with disturbances and uncertainties will be further studied and expand the related results to the switching topology.

Acknowledgments. This work was supported by the NSFC (61503231, 61473015).

References

1. Li Z, Ren W, Liu X (2013) Distributed consensus of linear multi-agent systems with adaptive dynamic protocols. *Automatica* 49(7):1986–1995
2. Li H, Liao X, Chen G (2013) Leader-following finite-time consensus in second-order multi-agent networks with nonlinear dynamics. *Int J Control Autom Syst* 11(2):422–426
3. Sun F, Guan ZH (2013) Finite-time consensus for leader-following second-order multi-agent system. *Int J Syst Sci* 44(4):727–738
4. Ghapani S, Mei J, Ren W (2016) Fully distributed flocking with a moving leader for Lagrange networks with parametric uncertainties. *Automatica* 67(5):67–76
5. Hui L, Li J (2009) Terminal sliding mode control for spacecraft formation flying. *IEEE Trans Aerosp Electron Syst* 45(3):835–846
6. Haghhighi R, Pang CK (2017) Robust concurrent attitude-position control of a swarm of underactuated nanosatellites. *IEEE Trans Control Syst Technol* 8:1–12
7. Varma S, Kumar KD (2012) Multiple satellite formation flying using differential aerodynamic drag. *J. Spacecraft Rockets* 49(2):325–336
8. Bazerque JA, Giannakis GB (2010) Distributed spectrum sensing for cognitive radio networks by exploiting sparsity. *IEEE Trans Sig Process* 58(3):1847–1862
9. Mohammadi A, Asif A (2015) Consensus-based distributed dynamic sensor selection in decentralised sensor networks using the posterior Cramér-Rao lower bound. *Sig Process* 108 (108):558–575
10. Khan UA, Kar S, Moura JMF (2010) DILAND: an algorithm for distributed sensor localization with noisy distance measurements. *IEEE Trans Sig Process* 58(3):1940–1947
11. Li S, Du H, Lin X (2011) Finite-time consensus algorithm for multi-agent systems with double-integrator dynamics. *J Tianjin Univ Technol* 47(8):1706–1712
12. Tong P, Chen S, Wang L (2018) Finite-time consensus of multi-agent systems with continuous time-varying interaction topology. *Neurocomputing* 284(6):187–193
13. Li W, Chen Z, Liu Z (2013) Leader-following formation control for second-order multiagent systems with time-varying delay and nonlinear dynamics. *Nonlinear Dyn* 72(4):803–812
14. Li H, Zhu Y, Wang J (2017) Consensus of nonlinear second-order multi-agent systems with mixed time-delays and intermittent communications. *Neurocomputing* 251(3):115–126
15. Tang ZJ, Huang TZ, Shao JL (2012) Consensus of second-order multi-agent systems with nonuniform time-varying delays. *Neurocomputing* 97(1):410–414
16. Hou W, Fu M, Zhang H (2017) Consensus conditions for general second-order multi-agent systems with communication delay. *Automatica* 75(6):293–298
17. Yu W, Chen G, Cao M (2010) Some necessary and sufficient conditions for second-order consensus in multi-agent dynamical systems. *Automatica* 46(6):1089–1095
18. Huang Y, Li Y (2019) Distributed rotating formation control of second-order leader-following multi-agent systems with nonuniform delays. *J Franklin Inst* 356(5):3090–3101

19. Li S, Peng X, Tang Y (2018) Finite-time synchronization of time-delayed neural networks with unknown parameters via adaptive control. *Neurocomputing* 45(8):65–74
20. Liu X, Xu B (2012) Distributed finite-time tracking control for multi-agent systems with communication delays. *Commun Theor Phys* 48(6):155–161
21. Halanay A (1996) Differential equations: stability oscillations time lags. Academic Press, New York



Social Network Emergency Incident Portrait Based on Attention Mechanism

Jiannan Chen, Junping Du^(✉), Lei Shi, Zhe Xue, 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 social networks, more and more people use social networks to publish and disseminate national security emergencies. In order to effectively control the spread and development of Chinese social network national security emergencies, we need to make an effective portrait of the emergencies. However, Chinese social network information has two research difficulties, such as text irregularity and few data sets in related fields, which may result in inaccurate event portrait results. In order to solve the above problems, we propose an algorithm based on the attention mechanism of Chinese part-of-speech tagging results (BLTAC) to perform emergency event portrait of Chinese social networks, which can efficiently perform emergency portraits. The BLTAC algorithm can be used to extract the Chinese social network emergency text entity name, and use the extracted entity name to describe the emergency event to perform event portrait. The experimental results show that the F1-score of our algorithm for the entity names recognition in each category on the Weibo dataset is improved compared with the other methods.

Keywords: Entity name recognition · Social network event portrait · Attention mechanism

1 Introduction

Online social networking is an online application service with powerful information publishing, acquisition and sharing capabilities. Social networks have become the most important platform for people to express their opinions and attitudes toward public affairs or hotspot emergencies. Social networks play a more and more significant role in the occurrence and development of emergencies. Recently, the portraits of emergencies have gradually become a hot topic in academic research. With the maturity of Knowledge Graph technology, there are more perspective to study event portraits. Entity name recognition is an important technology of the Knowledge Graph. There are two difficulties in the extraction of entity names in Chinese social network, especially hotspot emergencies: First, Chinese social network information is nonstandard and there is a lot of noise. Second, Chinese social network information, especially national security related corpus is small. Therefore, to solve the above two difficulties, it is

necessary to obtain relevant entity names from Chinese social network national security event information.

Entity name extraction technology is a task of natural language processing [1, 2]. In recent years, entity name recognition tasks have often been used as a sequence labeling task. With the continuous development of neural network technology, the technology of entity name recognition has developed rapidly. Collobert et al. [3] uses neural networks to identify entity names without supervision. Yao et al. [4] uses a convolutional neural network (CNN) model. Then Chen et al. [5] and Huang et al. [6] further uses the Long and Short Time Memory (LSTM) model. Bonadiman et al. [7] introduced the idea of deep neural network (DNN). He et al. [8] focuses on the vectorization of sentence-levels. Peng et al. [9] applies the Chinese word segmentation technique to the entity name recognition process. Xu et al. [10] uses cross-media ideas for joint training. However, the above methods do not fully take into account the characteristics of Chinese social network information, especially the contextual association of Chinese social network segmentation results and the characteristics of word boundaries. To solve the above problem, based on the BLSTM and CRF sequence labeling model, this paper introduces the attention mechanism of Chinese part-of-speech tagging results. An entity name recognition method based on attention mechanism is proposed to improve the accuracy of portraits of social security national security emergencies. In the next chapter we will use BLTAC (the attention mechanism of Chinese part-of-speech tagging results based on Bi-directional long and short memory network and conditional random field) to represent the proposed method. On the Weibo National Security Incident data set, the proposed algorithm increases recognition performance for events, time, location, and organization which ultimately improves the accuracy of portraits of national security events in Chinese social networks. The contributions of this article are as follows.

- (1) For the first time, the method of entity name recognition is used in the task of the social network national security emergency portrait to perform event portraits.
- (2) Based on the BLSTM-CRF network, the Chinese text part-of-speech tagging result is introduced to promote the accuracy of the attention-enhancing entity name recognition.

2 BLTAC Algorithm Description

This paper proposes an entity name recognition algorithm that introduces the attention mechanism of Chinese part-of-speech tagging results. As shown in Fig. 1, the entire system framework consists of three parts, the BLSTM layer, the attention mechanism layer, and the CRF layer. Specifically, BLSTM is used to obtain globalized bidirectional information from sentences. Adding the attention value of the part-of-speech tagging on the basis of BLSTM is beneficial to the extraction of the specific attribute entity name. Finally, the output of the attention layer is sequenced by the CRF layer. The following will explain in detail the specific role and implementation of each part.

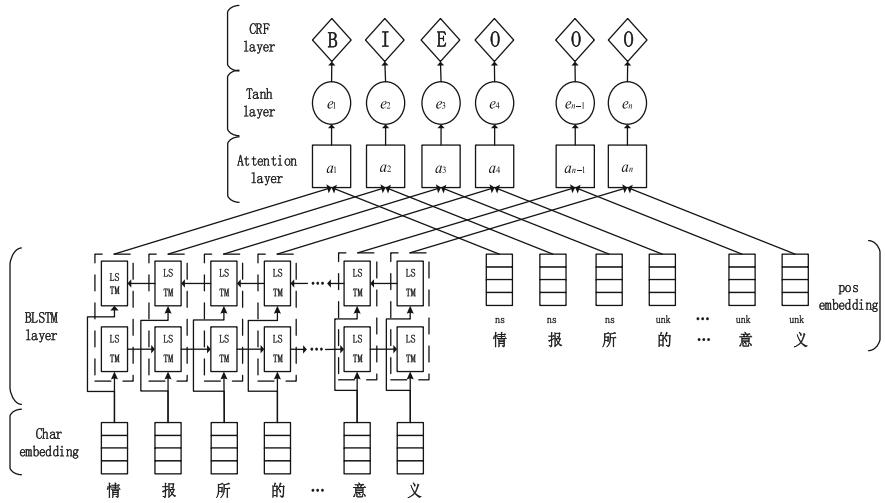


Fig. 1. Schematic diagram of BLTAC

2.1 BLSTM Layer

The role of BLSTM layer is to deal with the long-range dependency problem in social network text data. RNN is now widely used in many NLP-related tasks. Based on the model, the researchers proposed the LSTM model. The difference is that the LSTM model incorporates a special design in the hidden layer. The memory unit, which to some extent weakens the effects of gradient disappearance or gradient explosion, and can find the interdependence of data over a long period of time. On the basis of LSTM, this paper uses the technology of BLSTM. LSTM network structure is composed of “forgotten gate”, “input gate”, “candidate gate”, and “output gate”. h_{t-1} is the output result of the hidden layer of the previous LSTM unit, then C_{t-1} is the state result of the last LSTM unit, x_t is the word vector input result of this article, f_t is the output result of the “forgetting gate”, σ denotes the sigmoid activation function, i_t and \tilde{C}_t is output, tanh is the tanh activation function, o_t is the output of the “output gate”, h_t is the output of the hidden layer of the current cell. The entire LSTM unit involves the following formula.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

$$h_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \otimes \tanh(f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t) \quad (4)$$

The output to the long and short memory network has two hidden output results, forward \vec{h}_t and backward \bar{h}_t . The output of the BLSTM layer is shown in Eq. (5).

$$\mathbf{h}_t = [\vec{h}_t, \bar{h}_t] \quad (5)$$

2.2 Attention Layer

Attention mechanism is a selection mechanism used to allocate limited information processing capabilities. The purpose of using attention mechanism is to selectively focus on certain important information, and correspondingly ignore other information received at the same time. The main role of the attention network is to improve the effect of boundary detection and range detection on the specific entity name in the social security emergency content entity name. For example, when the result of the Chinese part-of-speech tagging is the location in the text data, the recognition weight value of the word entity name in the text data is relatively large. The structure of attention layer is shown in Fig. 2.

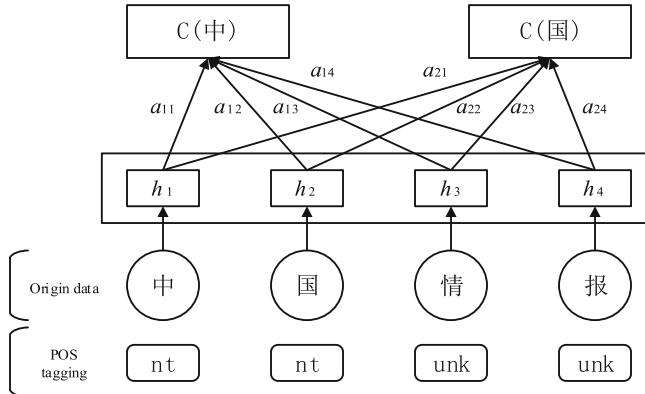


Fig. 2. Schematic diagram of attention layer

According to the calculation process shown in Fig. 2, the output of the attention layer of each original data can be obtained by using the output of the hidden layer in the BLSTM layer. Meanwhile the probability distribution value can be calculated by Eq. (6).

$$C_i = \sum_{j=1}^{T_x} a_{ij} h_j \quad (6)$$

h_j is the output of the BLSTM hidden layer, a_{ij} is the probability distribution of the part-of-speech tagging result, C_i is the result vector of the final result of the attention

layer. In the above, by introducing the Chinese part of speech in the social security emergency information of the social network, different weight values can be assigned to the words of different part of speech, thereby improving the accuracy of extracting the physical entity name.

2.3 CRF Layer

The special conditions on the linear chain are often used to solve the labeling problem. Specifically, X can be used as an input sequence which is the output result of the hidden layer of the BLSTM. P_{n*k} can be used as the BLSTM output probability matrix. P_{ij} is the probability that the i -th word is marked as the j th label. Then for the label sequence Y , the definition is given in Eq (7)

$$S(X, Y) = \sum_{i=0}^n A_{yi,yi+1} + \sum_{i=0}^n P_{i,yi} \quad (7)$$

A is the state transition matrix, A_{ij} represents the probability of moving from the i th tag to the j th tag. Using the maximum likelihood method to find the maximum value of $S(X, Y)$, so that the best output tag sequence can be obtained. Finally, the predicted label is converted to a single label vector to generate the desired label result.

3 Experiment

We use the collected Sina Weibo national security emergency data for evaluation and selected a total of 10 categories of national security emergencies, covering military, political, economic and other aspects. Through the data set, the proposed BLTAC algorithm is compared with the other four comparison algorithms on the data set. The evaluation metrics include precision, recall, and F1-Score.

3.1 Experiment Settings

3.1.1 Data Set

The Weibo data set includes the training Weibo set and the test Weibo set. The training data includes 7784 Weibo, and the test data includes 1950 Weibo. The detailed distribution of Weibo data in Table 1. The proportion of the training Weibo data and the test Weibo data label is basically the similar, which ensures the consistency of the data.

Table 1. Data corpus structure

| | train_data | test_data |
|-----------------------------|------------|-----------|
| Number of Weibo | 7784 | 1950 |
| Number of characters | 428019 | 105210 |
| Number of marked characters | 73332 | 19403 |
| Mark ratio | 17.13% | 18.44% |

In the experiment, all the labels of the text data are divided into LOC, ORG, PER, TIME and O. B_XXX is the beginning of the entity name, E_XXX is the end of the entity name, and M_XXX is the intermediate data of the entity name. In addition, “LOC” indicates location information, “ORG” indicates organization information, “PER” indicates person name information, and “TIME” indicates time information.

3.1.2 Network Parameter Setting

For the BLTAC network framework, the length of the word embedding vector is 300 dimensions, the length of the embedding vector of the Chinese part-of-speech tag is set to 600 dimensions, and the output dimension of the hidden layer for the BLSTM network is 300 dimensions. The batch_size is 32, the size of epoch is 32, the learning rate is 0.001, the dropout is set to 0.5 during training, and the optimizer uses Adam.

3.2 Experimental Comparison of Different Methods

We use the P (precision), R (recall) and F1-Score indicators to evaluate the performance of entity name extraction on the Weibo dataset. HMM, CRF, BLSTM, BLSTM +CRF are used to conduct comparative experiments. The experimental results are shown in Table 2. The last line of each table is the test result of the BLTAC algorithm proposed by us. The black marked data in the table is the highest evaluation method for each column.

Table 2. Entity name extraction performance comparison

| | B-XXX | | | M-XXX | | | E-XXX | | |
|-----------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | precision | recall | F1 | precision | recall | F1 | precision | recall | F1 |
| MM | 0.6221 | 0.6382 | 0.6215 | 0.6374 | 0.6763 | 0.6508 | 0.569 | 0.5826 | 0.5677 |
| CRF | 0.7832 | 0.5826 | 0.6663 | 0.7576 | 0.6553 | 0.7 | 0.7222 | 0.5367 | 0.6141 |
| BLSTM | 0.7418 | 0.587 | 0.6543 | 0.7301 | 0.6699 | 0.6956 | 0.7016 | 0.498 | 0.5774 |
| BLSTM+CRF | 0.8231 | 0.7384 | 0.7776 | 0.8428 | 0.74 | 0.7879 | 0.8116 | 0.6714 | 0.7344 |
| BLTAC | 0.8353 | 0.7631 | 0.7967 | 0.8455 | 0.7733 | 0.8071 | 0.7988 | 0.7285 | 0.7613 |

It can be clearly seen from the table that the performance index of the BLTAC method proposed in our paper is superior to that of other comparative experiments in most of the entity names. BLSTM can capture the forward and backward features of the sequence, so that the sequence information can be more fully obtained. The introduction of the CRF model on the basis of BLSTM is to sequence the output of the hidden layer of BLSTM. On the basis of the above, the attention mechanism is introduced in the use of the features of part-of-speech tagging in Chinese texts. It has a more important influence on the words of the corresponding part of speech, so that the boundary of the entity name is identified and the effect of entity name recognition is improved. Therefore, the BLTAC method is superior to the BLSTM+CRF method.

In addition, the overall improvement effect of each group of the BLTAC method is about 2%–3%, indicating that the increase is larger. From the perspective of F1-Score, the test scores of the test groups B_LOC, M_LOC, E_LOC, B_PER, and E_PER are more obvious, because the part-of-speech tagging for the location and the name of the person is better. However, for the precision of B_XXX, the effect of the method BLSTM_CRF is better than that proposed in this paper. The possible reason is that the number of E_XXX in the data set is small and the effect of training is poor. Secondly, the length of the entity name in Chinese data is longer. The introduction of part-of-speech tagging does not significantly improve the accuracy of the recognition of the end of the entity name.

3.3 The Effect of Parameters on Entity Name Extraction

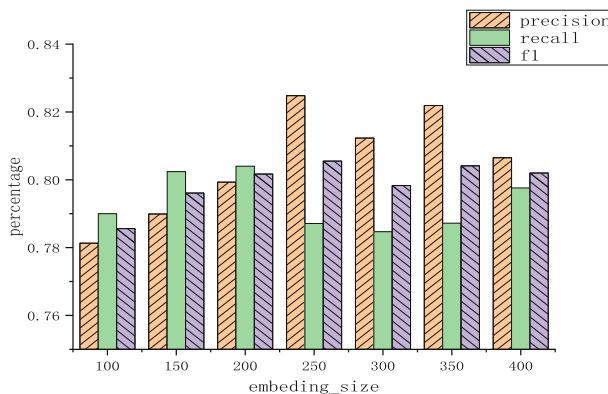


Fig. 3. Effect of model parameter embedding_size on BLTAC

The main training parameters of the BLTAC method are the length of the Weibo text word vector, i.e., embedding_size. The parameter analysis was performed on the Weibo data, and the average precision, recall rate and F1-Score were compared. The experimental results are shown in Fig. 3, embedding_size has a significant impact on the model. The parameter embedding_size tests 7 sets of experiments of 100, 150, 200, 250, 300, 350, 400. Overall, for F1-Score, the value of F1-Score increases as embedding_size increases, until embedding_size = 250 reaches a peak. Then, As the embedding_size increases, it drops slightly and remains steady.

3.4 Social Network National Security Emergency Portrait

The social network national security emergencies can be portraited using the trained BLTAC model. As shown in Fig. 4, From the Tianjin bombing portrait, we can clearly see when and where the incident occurred and related people and institutions.

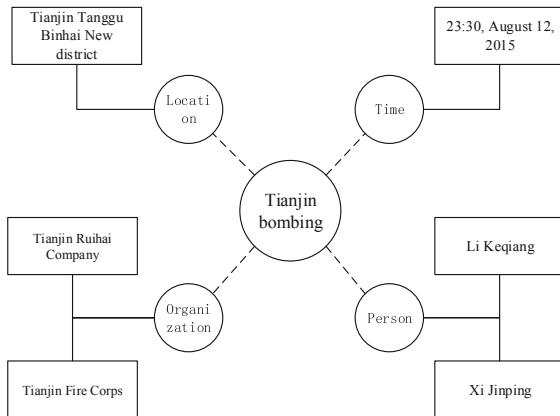


Fig. 4. Tianjin bombing incident portrait results

4 Analysis of Experimental Results

This paper proposes a social network security emergencies portraits method which uses the attention mechanism of Chinese text part-of-speech tagging based on BLSTM and CRF. By extracting the important entity names (locations, organizations, people, and time) of the social network national security emergencies, the corresponding keyword images are made, so people can understand the state of the incident and the corresponding people related to the incident quickly and in detail. We combine the datasets of Chinese datasets and social network national security datasets, and then extract the features of texts and the features of textual part-of-speech annotations. The features are first processed by the Bi-directional long and short memory network, then processed by the corresponding attention layer, and finally, the sequence is marked by the conditional random field layer. In this paper, we use multiple sets of experiments to verify that the BLTAC method achieve higher training accuracy, higher recall rate and higher F1-Score than other comparisons.

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

References

1. Ni J, Florian R (November 2016) Improving multilingual named entity recognition with Wikipedia entity type mapping. In: Conference on empirical methods in natural language processing, pp 1275–1284
2. Peng N, Dredze M (August 2016) Improving named entity recognition for Chinese social media with word segmentation representation learning. In: Meeting of the association for computational linguistics, pp 149–155

3. Collobert R, Weston J, Bottou L et al (2011) Natural language processing (Almost) from scratch. *J Mach Learn Res* 12(1):2493–2537
4. Yao C, Qu Y, Jin B, Guo L, Li C, Cui W, Feng L (2016) A convolutional neural network model for online medical guidance. *IEEE Access* 4:4094–4103
5. Chen J, Weihua LI, Chen JI, Xuze J, Yanbu G (2018) Bi-directional long short-term memory neural networks for Chinese word segmentation. *J Chin Inf Process* 32(2):29–37
6. Hunag Z, Xu W, Yu K (2015) Bidirectional LSTMCRF models for sequence tagging. *Comput Sci* 13(2):245–255
7. Bonadiman D, Severyn A, Moschitti A (2015) Deep neural networks for named entity recognition in Italian. *CLiC it* 51
8. He H, Sun X (2016) F-score driven max margin neural network for named entity recognition in Chinese social media. arXiv preprint [arXiv:1611.04234](https://arxiv.org/abs/1611.04234)
9. Peng N, Dredze M (2016) Improving named entity recognition for Chinese social media with word segmentation representation learning. In: Proceedings of 54th annual meeting of the association for computational linguistics, pp 149–155
10. Xu J, He H, Sun X, Ren X, Li S (2018) Cross-domain and semisupervised named entity recognition in chinese social media: a unified model. *IEEE/ACM Trans Audio Speech Lang Process* 26(11):2142–2152



Network Attack Detection Method Based on Convolutional Neural Network

Nan Ding¹, Yang Liu^{1(✉)}, Yimin Fan¹, and Dangyang Jie²

¹ School of Automation Science and Electrical Engineering, Beihang University,
Beijing 100083, China
ylbuaa@163.com

² Beijing Institute of Aerospace Systems Engineering, Beijing 100076, China

Abstract. The existing traditional network attack detection algorithms have low detection accuracy and are difficult to cope with external malicious and capricious network attacks in the communication and control systems. In order to solve this problem, this paper designs a network attack detection method based on convolutional neural network technology, which improves the recognition accuracy of attack detection for network attack signals. Specifically, a convolutional neural network structure with overlapping pooling layers is designed, and the improved data preprocessing method and parameter tuning make the recognition accuracy of the obtained convolution network on the KDD99 data set greatly improved. Finally, the experimental results show that the proposed network tuning strategy is effective, and the neural network structure designed in this paper is superior to other existing intelligent detection algorithms in terms of recognition accuracy.

Keywords: Network environment · Attack detection system · Deep learning · Convolutional neural network

1 Introduction

The application and popularity of Internet technology brings us convenience as well as many hidden dangers of network security. Especially in various control and communication networks in military and aerospace, information and system security is particularly important. Therefore, in order to intercept and respond to external attacks before the system is seriously harmed, many researchers have begun to research and design attack detection algorithms, to provide real-time protection against network attacks for the systems.

Attack detection is to collect information from key points in a system, and analyze the information in detail to find out whether there is an attack phenomenon or a violation of system security in the system or network [1]. In the traditional attack detection system design, the depiction of attack features can only be some fixed sequence, but in reality, the means of intruders using system vulnerabilities may change. In short, the intruder's attack modes are constantly evolving, and traditional attack detection systems are unable to identify these various modes of attack, which limits the system's adaptability.

With the development of artificial intelligence technology, the field of network attack detection began to reduce the error rate by using different emerging machine learning or deep learning techniques. Chung et al. constructed a set of attack detection models by combining various machine learning algorithms such as support vector machines, Bayesian classification and decision trees [2]. Meanwhile, Pan et al. proposed a k-means and Support vector machine combined with hybrid machine learning technology to detect attacks [3]. In [4], k-means algorithm was used to calculate the similarity between data, and adjusted the parameters to detect DDos attacks and worm attacks simultaneously. Besides, the LSSVM model [5] and hidden Markov model [6] were also proposed for network attack detection. The traditional machine learning method has achieved remarkable results in the field of network attack detection. However, the performance of machine learning technology is dependent on feature engineering, because it needs to artificially construct sample features. To solve this problem, researchers have introduced deep learning techniques. In [7], the probabilistic neural network was designed as attack detection techniques, while the deep belief network was applied to attack detection and achieved better results than other traditional machine learning methods [8].

The main work and innovations of this paper have the following three points: (1) The convolutional neural network technology is introduced into the field of network attack detection. By using its high-level abstraction and ability to extract features, the KDD 99 data set in the field of network attack detection is classified and identified. (2) The overlapping pooling layer is used instead of the common pooling layer to reduce the loss of feature information between different layers in the convolutional network. (3) With the improved data preprocessing method, unlike other existing similar work, our paper expands the one-dimensional input data into a two-dimensional matrix and cancels the one-hot processing. After the above improvement strategy, the detection rate of our convolutional neural network on this data set is increased to 99.84%. Compared with other intelligent detection algorithms, the detection rate has been significantly improved.

The content of this paper is arranged as follows: Sect. 1 is the research content and related research work. We elaborate the theoretical basis of convolutional neural network in Sect. 2. The network model structure, data preprocessing, network training details and parameters are introduced in Sect. 3. Section 4 shows the relevant experimental results and analysis. Finally the summary of this paper is given in Sect. 5.

2 Related Preliminaries

Convolutional neural networks (CNN) are commonly referred to as neural networks that use convolution operations at least in one layer of the network instead of general matrix multiplication. The basic operating layers and related principles in convolutional neural networks are briefly introduced below.

2.1 Convolutional Layer

Convolution is a mathematical integral transformation method for two time-varying functions, which has been widely used in many aspects. Convolution is the result of a

summation of two variables after they are multiplied in a certain range. If the convolved variables are sequences $x(n)$ and $h(n)$, the result of the convolution is as shown in Eq. (1):

$$y(n) = \sum_{i=-\infty}^{+\infty} x(i)h(n-i) \quad (1)$$

Each convolution in a neural network is a feature extraction method that allows the original data features to be enhanced and the noise to be relatively weakened. The convolution kernel is convolved with the original input or the upper layer output feature, and added to the offset to obtain the feature map of the current layer, as shown in Eq. (2):

$$x_i^l = \sum_{i \in M} y_i^{l-1} \otimes k_{ij}^l + b_j^l \quad (2)$$

Where x_i^l is the i-th feature map obtained by the l-th layer, y_i^{l-1} is the i-th feature map of the $l - 1$ th layer output, k_{ij}^l is the j-th convolution kernels of the convolution operation between l-th layer and the i-th feature map of its input, \otimes is the multiplication of the corresponding elements of two matrices, b_j^l is the offset term added to the j-th feature map obtained by the convolution of l-th layer, M is a collection of all the labels of feature maps output by the previous layer.

Convolutional networks are often used in tasks such as image classification or detection, at this time, the original feature map of the input network is a two-dimensional matrix of image pixel values, and the shape conforms to (batch size, length, width, channels). Batch size is the number of pictures trained in batches, length and width refer to the length and width of the picture, channels are the number of picture channels, and the number of picture channels in common RGB format is 3. The shape of the feature map that is continuously transmitted and changed in the convolutional layer conforms to the above format. Convolutional layer extraction features are available in two modes: ‘Same Padding’ and ‘Valid Padding’. Due to the network attack detection task studied in this paper, the input data used is network traffic and communication data, which is a one-dimensional vector, so it needs to be processed before inputting the network. This part will be introduced in Sect. 3.1.

2.2 Pooling Layer

The input of the pooling layer comes from the output of the convolutional layer. The feature map of the pooling layer does not change in the number of channels, but change in the length and width. The purpose of adding the pooling layer is to perform secondary extraction on the features. The convolutional layer convolves each neighborhood in the feature map to obtain the neighborhood features of the data, while the pooling layer integrates all the neighborhood features extracted from the convolutional layer and obtains new features, that is, using the overall statistical characteristics of the

adjacent output of a certain location instead of the output of the network at that location. Therefore, the pooling layer often does not exist independently, but always follows the convolution layer. The result of the pooling is to reduce the features and the network parameters, so that the dimension reduction can be achieved. The calculation process is shown in Eq. (3):

$$x_j^l = f(\beta_j^l \text{down}(x_j^{l-1}) + b_j^l) \quad (3)$$

where $\text{down}(x_j^{l-1})$ represents subsampling the j-th feature map of the $l - 1$ th layer. A typical example of operation is to sum all of the pixels in different $n \times n$ sub-blocks of the input image. In this way, the output image is reduced by n times in both dimensions. Each output's feature map corresponds to its own multiplicative offset β and an additive bias b . $f(*)$ is the activation function. There are two ways to pool operations: maximum pooling and average pooling.

2.3 Fully Connected Layer

The fully connected layer is calculated in the same way as the operational layer in the classical BP neural network. Each neuron is connected to the neurons of the previous feature map. The output of each neuron can be expressed by Eq. (4):

$$h_{w,b}(x) = f(w^T x + b) \quad (4)$$

where A is the neuron input, B is the neuron output, C is the connection weight matrix, and D is the bias.

In convolutional neural networks, fully connected layers often appear in the last layers to weight and add previously extracted features. The previous convolution and pooling operations are equivalent to feature engineering, and the subsequent fully connected layers are equivalent to feature weighting. Finally, the fully connected layer transforms the two-dimensional feature map into a one-dimensional vector as the output of this layer, and the length of the vector is the number of categories of the classification.

2.4 Softmax

Under the common multi-category classification task, softmax is the last layer of the network and acts as a classifier. Softmax is usually connected to the fully connected layer, and the one-dimensional vector of the fully connected layer is numerically processed so that the sum of the elements in the vector is 1. Each element value in the vector obtained after processing can be regarded as the confidence that the input data belongs to a certain category. The softmax function is calculated as Eq. (5):

$$S_i(x) = \frac{e^{V_i}}{\sum_i^C e^{V_i}} \quad (5)$$

where V_i is the output of the previous operational layer, i is the category index, C is the total number of categories. The output of the softmax layer will be used as the output of the entire network for the calculation of the loss function, and the network parameters will be iterated and updated through the learning optimization algorithm based on its gradient to achieve the purpose of network training.

3 Proposed Attack Detection Method

The overall framework of the network attack detection model based on convolutional neural network proposed in this paper is shown in Fig. 1. It consists of the following three parts:

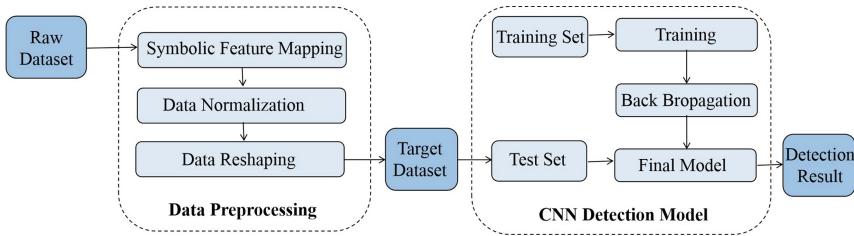


Fig. 1. Overall framework of network attack detection model based on convolutional neural network

(1) Data reading and preprocessing.

A standardized data set is obtained by digitizing and normalizing the symbolic feature attributes in the KDD 99 data set. Since the input data is a one-dimensional vector of length 41, in some recent similar work [9, 10], most of the literature adopts a one-dimensional convolution method: the convolution kernel and the pooled kernel are changed from matrix to vector. An improved data processing method is used in our paper: two-dimensional convolution and pooling are still used, but the one-dimensional input data is reshaped into a two-dimensional matrix before being inputted into the network. The network detection accuracy of these two methods are compared in Sect. 4.

(2) The specific structure of CNN.

The pre-processed data set is sent to the built model for training, and the obtained optimal features are identified by the Softmax classifier. In this paper, the experimental comparison of the accuracy of attack detection is carried out for a variety of feasible model structures, and finally the optimal one is selected.

(3) In the convolutional network model training, the Adam algorithm based on gradient descent is used to adjust the network parameters.

3.1 Data Set Introduction and Data Preprocessing

In 1999, the MIT Lincoln Laboratory simulated the US Air Force LAN, collecting 90,000 network connections and system audit data to form the KDD 99 data set, a classic data set in the field of attack detection. The data set has a total of 5 million pieces of data, providing 10% of the test subset and training subset, including 494,021 training samples and 31,029 test samples. Each piece of data is a network link record, which is marked as normal or abnormal. The abnormal types are divided into four major categories: DoS, Probe, R2L, and U2R. The four categories are divided into several subclasses, for a total of 39 types of attacks. The detailed type distribution and labels in the data set are shown in Table 1.

Table 1. KDD99 data set sample category distribution.

| Label | Category | Training set (10%) | Test set (Corrected) |
|-------|----------|--------------------|----------------------|
| 0 | Normal | 97278 | 60593 |
| 1–6 | PROBE | 4107 | 4166 |
| 7–16 | DOS | 391458 | 229853 |
| 17–24 | U2R | 52 | 228 |
| 25–39 | R2L | 1126 | 16189 |

Each data record in the data set contains 41 features, combined with category labels, for a total of 42 items. Among the 41 features, 38 numerical features and 3 symbolic features are included. For these features, separate preprocessing is required.

(1) The digitization of symbolic features

For the three symbolic features, we use a method of replacing with an integer type label. For example, for the feature “protocol_type”, three possible symbolic values: ‘tcp’, ‘udp’, ‘icmp’ are contained. Therefore, they can be converted into integer numbers of 0, 1, and 2, respectively. By performing the same operation on all three symbolic features, the symbolic features can be mapped to corresponding integer features.

Here, some recent similar work [11] performed one-hot processing on these transformed integer features. For example, for a feature ‘protocol_type’ containing three values, it is converted into a three-dimensional vector of [1,0,0], [0,1,0], [0,0,1], respectively. The purpose of this is because the different values of the symbolic features are independent, and the numerical relationship between the values will be introduced by simply labeling them with integer. After such one-hot processing, the impact of the numerical relationship of the features can be removed. But at the same time, the processed data will be expanded from the original 41-dimensional to 122-dimensional, and the values in a large number of dimensions are all 0, which introduces sparsity for the data and brings difficult to the improvement of network detection accuracy. Therefore, the one-hot processing is canceled in our paper. It is considered that in this data set, the negative impact of the mapped integer numerical relationship on the network is lower than the impact of a large number of sparse features.

(2) Normalization of numerical features

For the 38-dimensional numerical features, the size of the numerical features varies greatly due to the different dimensions. Therefore, in order to eliminate the influence of the dimensional difference, the values need to be normalized.

3.2 Network Structure

In order to best learn the characteristics of the data set and have higher detection accuracy, four kinds of convolutional neural network structures are established as shown in Table 2, and model training and performance comparison between them are carried out in Sect. 4.

Table 2. Convolutional neural networks with four different structures.

| Layers | CNN-1 | CNN-2 | CNN-3 | CNN-4 |
|---------------|-----------------------------------|-----------------------------------|-------------------------|-------------------------|
| Convolution-1 | 3×3 conv, stride 1, same | | | |
| Pooling-1 | 2×2 , stride 2 | 2×2 , stride 2 | 2×2 , stride 1 | |
| Convolution-2 | / | 2×2 conv, stride 1, same | | |
| Pooling-2 | / | 2×2 , stride 2 | 2×2 , stride 2 | 2×2 , stride 1 |
| FC-1 | 1000D | / | 1000D | |
| FC-2 | 40D, softmax | | | |

The different parameters for each of the operating layers in each structure are listed in Table 2. CNN-1 and CNN-2 are three-layer networks, but the number of convolutional layers and the number of fully connected layers are different. CNN-3 and CNN-4 are four-layer networks, but CNN-4 uses overlapping pooling layers. The size of the convolution kernel in the four network structures is taken as the common 3×3 and 2×2 , and the ‘Same Padding’ is used to make the shape of the feature map unchanged after convolution, also to ensure that the features of the edge portion of the input feature map can also be fully extracted by the convolution kernel. After each convolutional layer, a corresponding maximum pooling layer is added. After experiment and comparison, the network performance is better than the average pooling when taking the maximum pooling operation here. The experimental results can refer to Sect. 4.

Pooling is a very common operation in convolutional neural networks. The pooling layer often follows the convolutional layer, and the pooling is used to reduce the feature map dimensions of the convolutional layer output, while improving network performance (it is not prone to over-fitting). In general non-overlapping pooling, unlike convolution operations, pooling acts on areas of the image that do not coincide. The stride in the convolutional layer is usually smaller than the size of the convolution kernel, so that there is overlap between the two convolution windows. In the general pooling layer, the stride size is the same as the pooled sum, and the areas of adjacent pooling are only adjacent and do not overlap. This minimizes the dimension of the feature map, but it also causes some of the original information to be lost. Under the usual convolutional network image classification task, the data volume of a single image is much larger than the data volume of a single data in KDD 99. Therefore, the

image data set should be processed by data dimension reduction. And because of the local correlation of picture information, a lot of redundant information can be avoided by non-overlapping pooling, and network performance will not be affected by the loss of this part of redundant information. However, considering our task, which does not require too much data dimensionality reduction, there is no strong correlation between each feature inside each data. Loss of some features caused by non-overlapping pooling will have an impact on network performance. Therefore, the overlapping pooling layers in CNN-4 are used in our paper to further improve the detection accuracy of the network.

Since the amount of data per data is small, in order to prevent over-fitting of the model, in the model built in this paper, the dropout operation is added after the full connection layer 1, that is, each neuron in the fully connected layer in each training batch, both will be stopped to output with a certain probability p , and the weight will not be updated in this batch training. The probability that a neuron keeps outputting is $1-p$, which is the tunable parameter `keep_prob` in the dropout operation. In this paper, after the experiment, when the `keep_prob` takes 0.8, the network has the best test set accuracy.

4 Experimental Results

The environment configuration of the experiment in this paper is shown in Table 3.

The model used in this paper has a learning rate of 0.0001 and a batch size of 1000, and the optimized learning algorithm used is Adam.

The training set performance at different stages of training using a one-dimensional convolution and two-dimensional convolution network are compared in Table 4. The experimental results of the one-dimensional convolution use the data given in [10]. Under the premise of using the same network structure, the one-dimensional input data is reshaped into a 7×7 two-dimensional matrix, and the one-dimensional convolution is changed into a two-dimensional convolution. After the model is used for training, the second raw of experimental data in Table 4 is obtained and used for comparison. It can be seen that in different training iterations, the recognition accuracy of our two-dimensional convolution model is higher than that of one-dimensional convolution, and the accuracy is also improved faster than one-dimensional convolution.

Comparison of model performance using maximum pooling and average pooling are showed in Table 5. The network structure used is the structure of CNN-2 in Table 2, which contains two pooling layers, so four experiments were carried out: Max+Max means that both pooling layers are the maximum pooling; Max+Avg represents the first pooling layer is the maximum pooling, and the second pooling layer is the average pooling. The rest is the same. It can be seen that the performance of the four configurations is similar, and the Max+Max and Max+Avg configurations have the highest accuracy, but Max+Max has a smaller loss function. Therefore, the maximum pooling operation is selected for the pooling layer in the proposed structure of our paper.

Table 3. Experimental environment

| Experimental environment | Environmental configuration |
|--------------------------|-----------------------------|
| Operating system | Windows7 |
| Device | 1 GPU |
| Memory | 12 GB GPU memory, 80 GB SSD |
| Programming language | Python 3.6 |
| Deep learning framework | Tensorflow 1.6 |

Table 4. Comparison of experimental results of one-dimensional convolution and two-dimensional convolution

| Model/Epoch | 1 | 3 | 5 | 7 | 9 |
|--------------|--------|--------|--------|--------|--------|
| 1D Conv [10] | 56% | 70% | 80% | 90% | 94% |
| 2D Conv | 57.33% | 77.15% | 98.61% | 98.67% | 98.72% |

Table 5. Comparison of experimental results between average pooling and maximum pooling

| Model | Max+Max | Max+Avg | Avg+Max | Avg+Avg |
|---------------|---------------|---------------|--------------|---------|
| Accuracy | 98.05% | 98.05% | 98.03% | 97.97% |
| Loss function | 58.33 | 66.13 | 53.15 | 74.60 |

After the above experiments, two-dimensional convolution and maximum pooling are chosen in our paper, and four convolutional neural networks with different structures are built in Table 2. The four network structures are described in Sect. 3.2. The recognition accuracy of the four network structures after they are fully trained on the test set are showed in Table 6. It can be seen that for CNN-1 and CNN-2 with three-layer network structure, CNN-1 with two layers of fully connected layer performs better than CNN-2 with two layers of convolution. For a four-layer network structure CNN-3 and CNN-4 (two-layer convolutional layer and two-layer fully-connected layer), CNN-4 uses overlapping pooling, which has an improvement of recognition accuracy of about 0.3% compared to CNN-3 using non-overlapping pooling, achieving optimal performance in the four models.

The performance of the CNN-4 model and other machine learning and deep learning based models are compared in Table 7. It can be seen that the recognition accuracy of the convolutional network model designed in this paper is significantly better than that of other models, and the recognition ability of the network attack detection model is improved again.

Table 6. Comparison of experimental results of convolutional networks with four different structures

| Model | CNN-1 | CNN-2 | CNN-3 | CNN-4 |
|----------|--------|--------|--------|---------------|
| Accuracy | 98.74% | 98.05% | 99.53% | 99.84% |

Table 7. Comparison with experimental results of other attack detection algorithms

| Model | Accuracy |
|---|---------------|
| Model based on rough set and artificial immune calculation [12] | 97.86% |
| Fusion-based support vector machine model [13] | 98.20% |
| Deep belief networks model [14] | 98.59% |
| CNN-4 in this paper | 99.84% |

5 Conclusion and Discussions

In this paper, convolutional neural networks are used in network attack detection. It proves that the excellent performance and application range of convolutional neural networks are not limited to images. It can still perform various identification tasks well in other fields. Experiments show that the convolutional neural network can be used in the field of attack detection, and the network signal characteristics can be well extracted. In view of the structural characteristics of convolutional neural networks and some existing work, a convolutional neural network for network attack detection is built and three network optimization strategies are used to improves its classification performance on the KDD 99 data set. Compared with other artificial intelligence algorithms, the results show that the convolutional neural network is the current optimal detection model. In the future work, the model structure and parameters will continue to be optimized, the generalization performance of the network will be further improved, and the rapidity of the model for signal detection will be optimized, thereby enhancing its practical application value.

References

- Williams PD et al (2001) CDIS: towards a computer immune system for detecting network intrusions. In: 4th international symposium on recent advances in intrusion detection, RAID 2001, Proceedings DBLP, Davis, CA, USA, 10–12 October 2001 (2001)
- Chung S, Kim K (2015) CISC: a heuristic approach to enhance the performance of intrusion system using machine learning algorithms. In: Proceedings of the Korea Institutes of Information Security and Cryptology Conference, Korea (2015)
- Pan XL, Luo Y, Xu YT (2015) K-nearest neighbor bases structural twin support vector machine. Elsevier Science Publishers, Amsterdam
- Shin D, Choi K, Chune S et al (2016) Malicious traffic detection using K-means. J Korean Inst Commun Inf Sci 41(2):277–284
- Zhao F (2015) Detection method of LSSVM network intrusion based on hybrid kernel function. Mod Electron Tech 38(21):96–99
- Jha S, Tan K, Maxion RA (2001) Markov chains, classifiers and intrusion detection. In: Proceedings of the 14th IEEE computer security foundations workshop. IEEE, Piscataway
- Raman MRG, Somu N, Kirthivasan K et al (2017) A hypergraph and arithmetic residue-based probabilistic neural network for classification in intrusion detection systems. Neural Netw 92(23):89–96
- Gao N, Gao L, Gao Q et al (2015) An intrusion detection model based on deep belief networks. In: Second international conference on advanced cloud & big data, pp 247–252

9. Fan J, Lingzhi L (2017) Intrusion detection algorithm based on convolutional neural network. *Trans Beijing Inst Technol* 37(12):1271–1275
10. Yuefeng L, Cheng W, Yabin Z et al (2018) Deep convolutional neural network model for network intrusion detection system. *J Inner Mongolia Univ Sci Technol* 124(01):62–67
11. Ming W, Jian L (2017) Network Intrusion detection model based on convolutional neural network. *J Inf Secur Res* 3(11):990–994
12. Ling Z, Zhongying B, Shoushan L et al (2013) Integrated intrusion detection model based on rough set and artificial immune. *J Commun* 9:168–178
13. Desheng Z, Fei Z (2013) Research on intrusion detection system based on the SVM and the fusion technology. *Bull Sci Technol* 29(5):167–172
14. Chunlin L, Huangyue J, Hong W et al (2014) Detection of network intrusion based on deep learning. *Inf Secur Commun Priv* 10:68–71



Study on Path Planning for Family Service Robot Based on Improved Genetic Algorithm

Xuedong Jing, Ya-nan Chen^(✉), and Yuquan Xue

School of Mechanical Engineering, Shanghai Institute of Technology,
Shanghai 201418, China
810135367@qq.com

Abstract. An approach to obstacle avoidance and navigation for mobile robot based on structured environment is proposed. In this approach, the IGA (Improved Genetic Algorithm) is integrated with graph theory, and is applied to find the appropriate path; meanwhile, by application of the niche algorithm, the diversity of population after the initial population generated can be maintained. The simulated annealing algorithm is employed to IGA so as to optimize the result and to avoid the large increase of individuals. The simulation results indicate that this method is successfully applied in large scale with less iteration and the convergence rate is improved.

Keywords: Path planning · Graph theory · Genetic algorithm · Simulated annealing algorithm · Niche algorithm

1 Introduction

Robot technology has been in-depth development, not only in the industrial field, but also in the service. It has a wide range of applications. At the same time, automatic path planning has great significance to the autonomous movement and route optimization of home service robot, which according to the known geographic information data to find a safe and shortest path from one point to another in complex environment. The navigation technology has been applied to many methods, such as simulated annealing algorithm, neural network algorithm, artificial potential field algorithm [1], particle swarm optimization algorithm, ant colony algorithm and graph search method. Compared with other search methods, genetic algorithm adopts multi-point search, which has stronger global search ability and higher search efficiency, so it is more likely to find the global optimal solution. The strong robustness and flexibility make it widely used in path planning of mobile robot [2].

The problem of traditional genetic algorithm is that the evolution speed is difficult to control and the premature convergence often occurs. At the same time, there are too many empirical parameters to meet the real-time needs, which is not conducive to automatic processing. Especially in a complex environment, if a better initial solution can't be given, it is usually difficult to find a feasible path through hybrid mutation and other operations. Genetic algorithm usually produces redundant search, which affects the convergence rate of the algorithm makes the algorithm need a long time to calculate. In this paper, combine the niche technology, simulated annealing algorithm with

genetic algorithm to ensure that individuals can be dispersed in the global scope, and to avoid the function falling into local optimization. What's more, the new individual is adjusted to form a new solution, and the objective function is calculated to determine whether or not to receive the new solution, so as to obtain a better adaptive solution, to ensure the quasi-certainty of the evolutionary direction, and to prevent the degradation [3]. Thus the optimal path is obtained. The feasibility and effectiveness of the method are verified by simulation experiments.

2 Establishment of Environmental Model

In this paper, the location information of obstacles is obtained by laser radar Rplidar-A2, and using the improved ICP algorithm to match the data of different positions to get the entire map [4].

The environment used in this paper is the laboratory, just as Fig. 1 shown, through the laser radar Rplidar-A2 continuously scanning the environment, calibrating the position of mobile robot and obstacles. Finally get the environment map, which is shown in Fig. 2, left. Then modeled and simplified the environment by MATLAB, as is shown in Fig. 2, right.



Fig. 1. The panorama of laboratory.

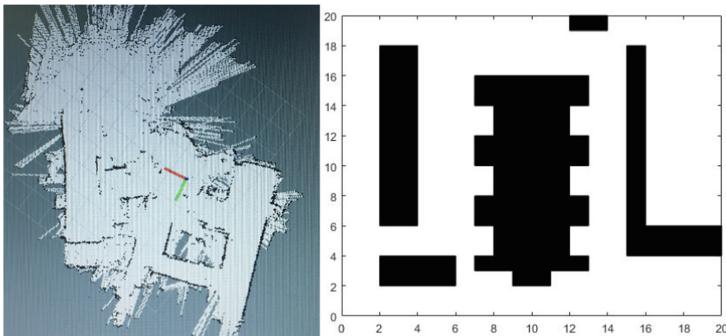


Fig. 2. Laboratory model map

3 Improved Genetic Algorithm

The flowchart of improved genetic algorithm is illustrated in Fig. 3.

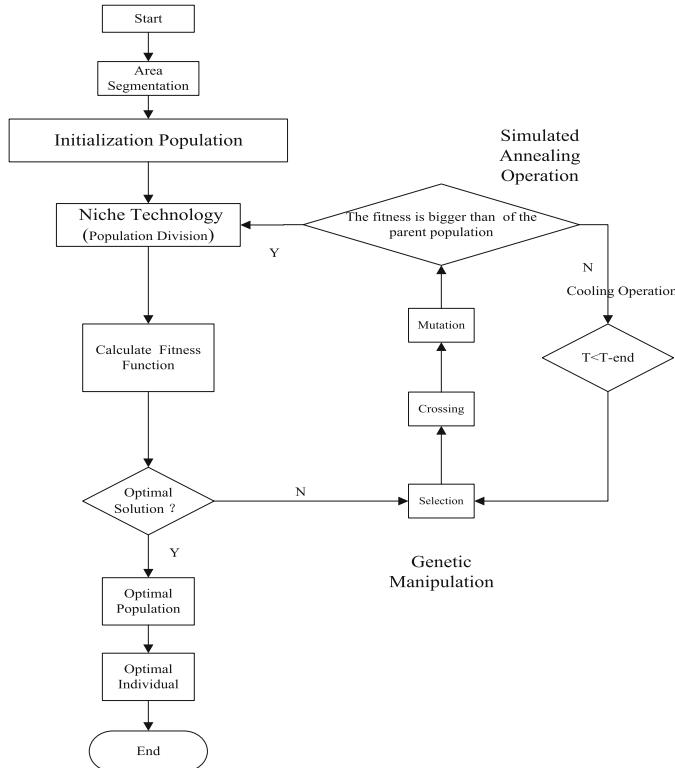


Fig. 3. Flowchart of improved genetic algorithm

3.1 Area Segmentation

In this paper divides the environmental map into 13 areas: *start*, *A1*, *A2*, *A3*, *B1*, *B2*, *C1*, *C2*, *D1*, *D2*, *E1*, *E2*, *End*, as is shown in Fig. 4, the areas are connected into a directed graph by directed line segments according to the graph theory, as shown in Fig. 5. According to the graph theory, the question of finding a path from the “*Start point*” to the “*End point*” can be transformed into a simple directed graph G , and the relation between regions can be represented by edge e , the feasible area of map partition can be represented by *start*, *A1*, *A2* ... *D1*, *D2*, *End*.

The graph relation corresponding to the path problem of any two points in this graph is $G = (V(G), E(G), \varphi_G)$ [5], of which the element of $V(G)$ represents the point of graph G , the element of $E(G)$ represents the edge of graph G , φ_G is the correlation function, it makes every element of $E(G)$ corresponding to an unordered element pair in $V(G)$, of which:

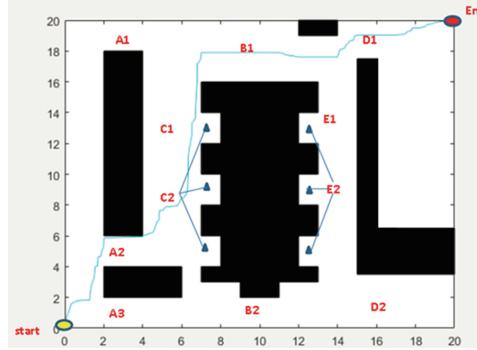


Fig. 4. The area segmentation of environment

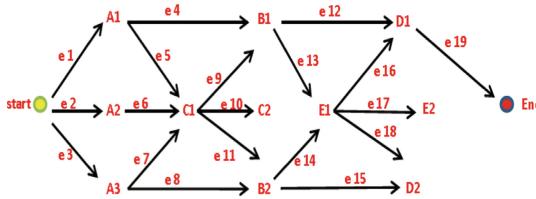


Fig. 5. Path directed graph G

$$V(G) = \{start, A1, A2, A3, B1, B2, C1, C2, D1, D2, E1, E2, End\},$$

$$E(G) = \{e1, e2, e3, e4, e5, e6, e7, e8, e9, e10, e11, e12, e13, e14, e15, e16, e17, e18, e19\},$$

$$\begin{aligned} \varphi_G(e1) &= start * A1; \varphi_G(e2) = start * A2; \varphi_G(e3) = start * A3 \\ \dots \varphi_G(e17) &= E1 * E2; \varphi_G(e18) = E1 * D2; \varphi_G(e19) = D1 * End \end{aligned}$$

Using b_{ij} to represent the correlation times of v_i and e_j , the directed correlation matrix $B(G)$ is:

$$B(G) = \left[\begin{array}{ccccccccc} Start & 1 & 1 & 1 & & & & & \\ A1 & -1 & & & 1 & 1 & & & \\ A2 & & -1 & & & & 1 & & \\ A3 & & & -1 & & & & 1 & 1 \\ B1 & & & & 1 & 1 & & & \\ B2 & & & & & -1 & -1 & & \\ C1 & & & & & & 1 & 1 & \\ C2 & & & & & & & -1 & \\ D1 & & & & & & & & -1 \\ D2 & & & & & & & & & -1 \\ E1 & & & & & & & & & 1 \\ E2 & & & & & & & & & & -1 \\ End & & & & & & & & & & 1 \end{array} \begin{array}{c} e1 \\ e2 \\ e3 \\ e4 \\ e5 \\ e6 \\ e7 \\ e8 \\ e9 \\ e10 \\ e11 \\ e12 \\ e13 \\ e14 \\ e15 \\ e16 \\ e17 \\ e18 \\ e19 \end{array} \right] \quad (1)$$

The directed correlation matrix $B(G)$ represents the association between regions: 1 means positive correlation; -1 means reverse correlation; 0 means no correlation. The point set of graph $G = (V, E)$ is V ; a_{ij} is the number of edges between point v_i and v_j ; $M(G)$ is the adjacency matrix of graph G:

$$M(G) = \begin{bmatrix} start & A1 & A2 & A3 & B1 & B2 & C1 & C2 & D1 & D2 & E1 & E2 & End \\ start & 1 & 1 & 1 & & & & & & & & & \\ A1 & 1 & & & 1 & & 1 & & & & & & \\ A2 & 1 & 1 & & & & & 1 & & & & & \\ A3 & 1 & & & & & 1 & 1 & & & & & \\ B1 & & 1 & & & & & 1 & & 1 & & 1 & \\ B2 & & & 1 & & & & & 1 & & 1 & & \\ C1 & & 1 & 1 & 1 & 1 & & 1 & & & & & \\ C2 & & & & & & 1 & & & & & & \\ D1 & & & & & 1 & & & & & 1 & & 1 \\ D2 & & & & & & 1 & & & & & 1 & \\ E1 & & & & & 1 & 1 & & 1 & 1 & & 1 & \\ E2 & & & & & & & & & & 1 & & \\ End & & & & & & & & & 1 & & & \end{bmatrix} \quad (2)$$

Each edge is given a weight, the edge represents the path, and the weight of the edge represents the length of the path. Path planning of any two-point can be guided and found by $B(G)$ and $M(G)$.

3.2 Encoding the Path

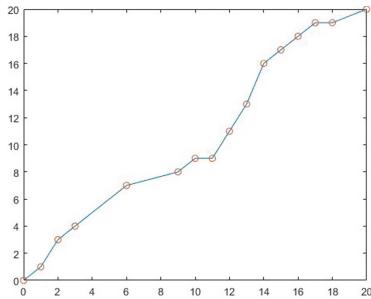
In this paper, the decimal system is used for individual paths, and the horizontal and vertical coordinates are simply coded independently. The structure is as follows:

$$\begin{aligned} X_i(X_{i1} \rightarrow X_{i2} \rightarrow \dots \rightarrow X_{ik} \rightarrow \dots \rightarrow X_{ij}) \\ Y_i(Y_{i1} \rightarrow Y_{i2} \rightarrow \dots \rightarrow Y_{ik} \rightarrow \dots \rightarrow Y_{ij}) \end{aligned} \quad (3)$$

Too much initial population will lead to large amount of calculation, and too little population will lead to the incorrect results. In this paper, 100 groups of path points will be distributed in the map environment at the beginning of the selection.

In formula (3), X_i represents all transverse coding of path i ; Y_i represents all vertical coordinate coding of path i . As is shown in Fig. 6, the line stands for the path k , and point $(0,0)$ is the initial point, the target point is $(20,20)$.

$$\begin{aligned} X_k(0, 1, 2, 3, 6, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20) \\ Y_k(0, 1, 3, 4, 7, 8, 9, 9, 11, 13, 16, 17, 18, 19, 19, 20) \end{aligned}$$

**Fig. 6.** Sample path

3.3 Fitness Function

In this paper, two conditions are considered as the evaluation criteria of fitness function: (1) the path length; (2) the path feasibility.

The Factor of Path Length

$$K1 = (L - \sum [(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2]) \cdot m_1 \quad (4)$$

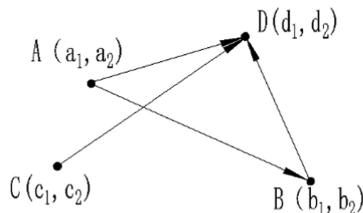
In formula, L is a relatively large constant, $\sum [(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2]$ means the length of path, use L minus it to get $K1$, the bigger the better. m_1 is the length operator that influences the path length.

The Factor of Path Feasibility

To judge the feasibility of the road, two measures are adopted in this paper:

1. When generating the initial population, determine if it falls into an obstacle area;
2. Detect if the adjacent path point connection interferes with the boundary of the obstacle.

To detect the interference between the path and the obstacle, using the vector method, shown as Fig. 7, is used to determine if the two segments AB and CD on the same plane are crossing. The coordinates of the points are $A(a_1, a_2)$, $B(b_1, b_2)$, $C(c_1, c_2)$, $D(d_1, d_2)$, then the question transformed to determine whether point A and point B straddle the two sides of segment CD .

**Fig. 7.** Path interference detection model

According to the basic law of vector, the following equations are obtained:

$$\begin{aligned}
 \overrightarrow{AD} &= (d_1 - a_1, d_2 - a_2) \\
 \overrightarrow{CD} &= (d_1 - c_1, d_2 - c_2) \\
 \overrightarrow{BD} &= (d_1 - b_1, d_2 - b_2) \\
 \vec{j} &= \overrightarrow{AD} * \overrightarrow{CD} \\
 \vec{i} &= \overrightarrow{BD} * \overrightarrow{CD} \\
 u &= \vec{j} \bullet \vec{i}
 \end{aligned} \tag{5}$$

Obviously, the positive and negative of u can judge whether the line segment AB crosses the line segment CD . When it is a positive number, it means segment AB has no intersection with CD . When it is negative, means segment AB intersected with segment CD . Then getting the factor of path feasibility $K2$, $K2 = m_2 \cdot n$ of which m_2 means the operator of obstacle interference, n means the number of interference between the path and the boundary of the obstacle.

The fitness function is:

$$F = K1 + K2 = (L - \sum \left[(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 \right]) \cdot m_1 + m_2 \cdot n \tag{6}$$

3.4 Niche Technique for Population Division

Niche technology can better preserve the diversity of progeny population [6], and divide the population according to the individual dissimilarity of the population. In this paper, different path selection is used as the basis of population division, such as Fig. 5 above, there can be a variety of path choices based on the starting point “start” and the target point “end”, such as “start-A2-C1-B1-D1-End” or “start-A3-B2-E1-D1-End” and so on.

3.5 Genetic Operation

Selection

In this paper, firstly sort the fitness, compare the individual probability between the new population and the original population, then select the better individuals according to the championship method, removes the individuals whose fitness is 0 or low. Ensure that the survival increases so that good individual can evolve to the next generation and form a new population.

Crossover

For the new population formed after selection, all individuals are divided into two parts: pre-chromosome and post-chromosome by single-point crossover method

according to the cross rate. A new individual is formed by splicing the pre-chromosome of one individual with the post-chromosome of another.

$$\begin{cases} P'_1 = r_c \cdot P_1 + (1 - r_c)P_2 \\ P'_2 = (1 - r_c)P_1 + r_c P_2 \end{cases} \quad (7)$$

P'_1 and P'_2 are the newly generated individuals. r_c randomly generates a number between 0 and 1 according to the roulette gambling method [7].

Mutation

In order to prevent genetic algorithm fall into precocious and local optimization, mutation is needed. Select an individual, randomly select a node for its chromosome, and replace it with another node randomly. If the mutation rate is more than 0.5%, the genetic algorithm will degenerate to random search [8]. In order to ensure that the better path in the later stage is preserved, the value of variation in this paper used an adaptive function. In this paper, the mutation rate is chosen by:

$$r_m = \begin{cases} k \left(1 - \frac{\arcsin\left(\frac{f_m}{f_{max}}\right)}{\pi/2} \right), & \arcsin\left(\frac{f_m}{f_{max}}\right) \geq \pi/6 \\ k \left(\frac{\arcsin\left(\frac{f_m}{f_{max}}\right)}{\pi/2} \right), & \arcsin\left(\frac{f_m}{f_{max}}\right) < \pi/6 \end{cases} \quad (8)$$

Where k is the mutation factor and f_m is the average of fitness, f_{max} is the maximum of fitness.

3.6 Simulated Annealing Operation

The new generation population is generated by genetic operation, then the newly generated individuals are adjusted to produce a new solution NS , calculates the objective function OB , and determines whether it accepts the new solution or not by the design criterion to realize the optimization of local optimal solution and global optimal solution in solution space, and to improve the convergence rate [9]. The actions are as follows:

1. Set the current solution OS and the current temperature T_i .
2. NS is the adjusted solution, and the objective function is OB , Δf represents the variable of the objective function. $\Delta f = abs(NB - OB)$. The design criterion is $E = \begin{cases} \exp\left(\frac{\Delta f}{T_i}\right) \\ 1 \end{cases}$, when $E = 1$, accept the new solution, or use $\exp\left(\frac{\Delta f}{T_i}\right)$ as probability to abandon the new solution.
3. Cooling operation with attenuation rate r , $T_{i+1} = T_i \cdot r$.
4. Determine whether the termination condition is met. If it is meet, jump to the next step, otherwise go to step 2.
5. Output result.

3.7 End Condition

In this paper, set the fitness function to be the termination condition, and select the shortest path which has the best fitness as result.

4 Simulations and Analysis

In order to verify the feasibility of the algorithm, the laboratory map is selected and simulated by software MATLAB. The initial population is 100; the number of genetic evolution is 100, produce 100 new populations at one time; the service robot is regarded as a particle, the starting point coordinates are (1,1), and the target point coordinates are (20,20), as shown in Fig. 8. The lines in the graph represent the path trajectory by improved genetic algorithm.

The final figure shows that the robot can circumvent obstacles under the control of the algorithm in this paper. Multiple experiments have taken out multiple tracks, and all the tracks are safe obviously, meeting the requirements of trajectory planning and also showing the uncertainty of navigation, as shown in Fig. 8, left. The right of Fig. 8 shows the path diagram obtained from the same starting point and the target point under the traditional general algorithm. Figure 9 shows the average fitness growth curve of GA and IGA. Table 1 shows the parameters of GA and IGA.

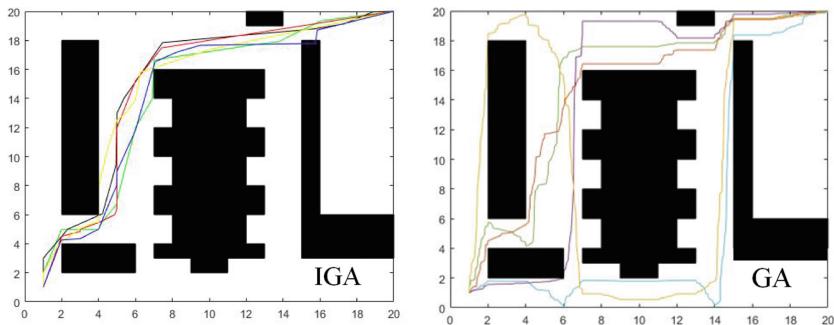
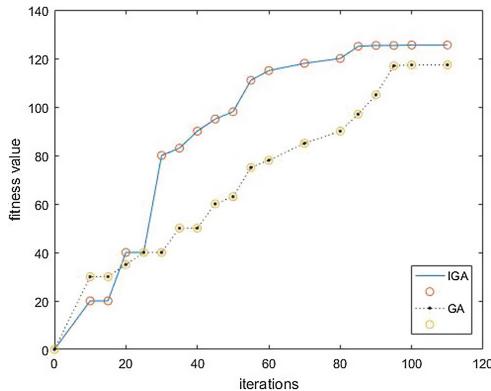


Fig. 8. Repetitive trajectories planning with different algorithms

As is shown in Table 1, the IGA improves the fitness value of the path, reduces the number of iterations, and optimizes the length of the path. The comparison results and the path graph show that the improved algorithm is feasible and effective, and has higher convergence efficiency and better optimization results than before.

**Fig. 9.** Average fitness growth curve**Table 1.** Parameters of IGA and GA

| | Path | 1 | 2 | 3 | 4 | 5 |
|--------------|------------------------------|----------|----------|----------|----------|----------|
| IGA | Length | 33.8220 | 32.8112 | 31.6743 | 32.6322 | 32.9052 |
| | Fitness | 124.6349 | 125.3910 | 126.2443 | 125.5259 | 125.3211 |
| | Convergence iteration number | 56 | 90 | 85 | 76 | 82 |
| GA Algorithm | Length | 44.4539 | 39.5081 | 37.9869 | 71.5301 | 34.145 |
| | Fitness | 116.6596 | 121.8670 | 121.5098 | 96.3525 | 124.3913 |
| | Convergence iteration number | 62 | 98 | 96 | 90 | 96 |

5 Conclusions

In this paper, the global optimization ability and the convergence rate of the algorithm are improved. This paper uses the graph theory to guide the path planning, and the niche technology to guarantee the diversity of population. After genetic operation, simulated annealing algorithm is used to improve the quality of the global optimal solution. The fitness function is improved, and removed the invalid path points, which improves the efficiency of evolution. The mutation rate is chosen by adaptive function in order to ensure the excellence of individuals in the later stage and to prevent the degradation of genetic algorithm due to the low mutation operator. Combined with simulated annealing algorithm, the new generation population is judged and selected to ensure the evolve direction of the offspring population. The simulation results show that the improved genetic algorithm is feasible and effective in the path planning of home service robot.

Acknowledgments. This work was supported by ‘Study on Key technologies of Parallel Robot for Minimally Invasive Spine Surgery’, Scientific Research Project of Shanghai Municipal Science and Technology Commission, (Projection No. 16090503700).

References

1. Wang R (2015) Research on path planning of mobile robot based on improved artificial potential field. In: Proceedings of 2015 joint international mechanical, electronic and information technology conference (JIMET). Global Union Academy of Science and Technology, Chongqing, p 4
2. Boone G, Chiang HD (1993) Optimal capacitor placement in distribution systems by genetic algorithm. *Int J Electr Power Energy Syst* 15(3):155–162
3. Song J, Meng T, Zhang H, Liu Y, Luo F (2017) Reactive power optimization of distribution network based on Niche genetic algorithm. *Electr Swichgear* 55(4):63–66
4. Li H (2017) Research of 2D-SLAM based on LIDAR. Zhejiang University of Technology
5. Bu Y, Wang W, Lv X (2015) Graph theory and its application. Southeast University Press, Nanjing, pp 15–17
6. Wang L, Li M, Cai J, Liu Z (2017) Research on mobile robot path planning by using improved genetic algorithm. *Mech Sci Technol Aerosp Eng* 36(5):711–716
7. Li L, Su Z, Xie Y, Liu F (2015) Path planning of robot fish's based on genetic algorithm. *Ordnance Ind Autom* 34(12):93–96
8. Zhao K, Wei Y, Xu L, Wang D (2015) Application of improved particle swarm optimization algorithm in path planning of mobile robot. *Fire Control Command Control* 31(9):233–236
9. Tang L, Guo G, Hu M, Yin J (2009) Path planning for mobile robot based on Niche genetic algorithm. *Modern Electron Tech* 32(24):95–99



Study on Dose Distribution Prediction of Esophageal Cancer Patients Using U-Net Model

Jing Wang¹, Ronghu Mao², Jiwei Liu¹, and Jianfei Liu^{3(✉)}

¹ School of Automation and Electrical Engineering,
University of Science and Technology, Beijing, China

² Department of Radiation Oncology, The Affiliated Cancer Hospital
of Zhengzhou University, Henan Cancer Hospital, Henan, China

³ Electrical Engineering and Automation, Anhui University, Hefei, China
liujf@ahu.edu.cn

Abstract. Accurate radiotherapy for esophageal cancer is very important for medical diagnosis and human health. In the course of radiotherapy, the distribution of dose in different organs is a problem that doctors and patients are very concerned about. In this study, the U-Net model was used to predict the desired dose distribution from the patient's contours of planning target volume (PTV) and organs at risk (OAR), so as to realize the automatic prediction of dose distribution. The object of this study was the CT data of esophageal cancer. The contours of patients' organs (PTV, left lung, right lung, heart) and dose distribution maps were extracted and processed, then we used U-net model to train the data. For a patient, the network model was trained on a single slice, and the output was the prediction of the dose for the patient slice. By training the network model, the radiotherapy plan was automatically formulated based on prior knowledge. Cosine similarity was used to evaluate the similarity between the predicted dose distribution maps and the true dose distribution maps. In addition, the dose volume histogram (DVH) curve of the true dose and the predicted dose was used to evaluate the accuracy of the prediction. It can be seen that the dose distribution predicted by U-Net is very close to the true dose distribution.

Keywords: Dose distribution · U-Net · Cosine similarity · DVH

1 Introduction

Esophageal cancer is a common digestive tract cancer, and about 300,000 people die of it every year in the world. It is a serious threat to human health. Therefore, precise radiotherapy for esophageal cancer is of great significance to medical diagnosis and human health. How to carry out precise radiotherapy accurately and rapidly is a long-standing problem in the medical field.

The fundamental goal of radiotherapy is to maximize the dose of radiation to the target area in order to kill cancer cells while keeping the surrounding normal tissues and key organs as little as possible from unnecessary irradiation. When developing a

radiotherapy plan, a large amount of calculation is required. The traditional calculation method is slow, time-consuming and costly, which seriously reduces the efficiency of doctors in formulating treatment plans. Therefore, it is of great practical significance to realize the automatic prediction of dose distribution. To a certain extent, it can reduce the burden of doctors and help doctors formulate high-quality treatment plans.

At present, the focus of dose prediction in radiotherapy is to use the patient's historical plans and information to predict the patient's dose distribution and dose volume histogram (DVH). Although it has made great success and progress and improved the accuracy and reliability of the results, it needs a lot of time and energy to select features, such as spatial location relationship between organs at risk (OAR) and planning target volume (PTV), overlapping volume histograms (OVH), distance-to-target histograms (DTH) [1–6], and hand-made for dose and DVH prediction models [7]. Mathematical models were also often used in dose and DVH prediction. For example, the spatial relationship between the target area and OAR and its correlation with DVH was established, then the principal component analysis (PCA) was performed and support vector regression was used to predict DVH [8, 9]. However, the features extracted by this method can only capture low-level information, which may not be enough for accurate prediction. Later, by comparing the dose models constructed by multiple linear regression and Deep Belief Nets (DBN) in the artificial neural network, it was found that the artificial neural network model was better than the multiple linear regression model [10].

With the rapid development of deep learning, the use of neural networks to deal with medical image problems has gradually come into our vision. The most famous deep learning model is the convolutional neural network (CNN). Based on CNN, fully convolutional network (FCN) was proposed [11], which opened up a new direction of image segmentation. Then, based on FCN, an improved model called U-Net was proposed [12], which was mainly applied to the semantic segmentation of biomedical images. The U-Net model had some advantages: (1) it supported a small amount of data to train the model, (2) higher accuracy can be obtained by predicting each pixel, (3) it required short training time. The model used most of the information in the image, which was a marked feature of this method compared with previous related work. Therefore, the accurate prediction of dose distribution can be realized by modifying the U-Net model and using the contour information of patients.

2 Methods

2.1 U-Net Architecture for Dose Prediction

Figure 1 depicted the U-Net model, which was modified to predict the dose distribution. The framework was based on the U-Net model to predict end-to-end dose distribution. The structure of the U-Net network was a U-shaped, and it was a fully convolutional neural network structure in which both the input and output were images.

There was no fully connected layer in the whole network. It contained two paths, the left side was the contraction path, following a typical convolutional network structure consisting of two repeated 3×3 convolution kernels, all using a rectified linear unit (ReLU) activation function and a 2×2 max-pooling operation with a step size of 2 for downsampling. In each downsampling step, the number of feature channels was doubled and the feature map size was reduced. The right side was the expansion path, and each step involved upsample of feature maps; then used the 2×2 convolution kernel for convolution, which was used to reduce the number of feature channels by half; and then cascaded the corresponding feature maps in the contraction path, followed by two 3×3 convolution kernels for convolution operation, using the ReLU activation function. Here, we constructed a seven layer network structure with six max-pooling operations and six upsampling operations. Convolution and ReLU operations of each layer were followed by batch normalization (BN), which acted as a regularizer for normalization feature. Seven CNN layers were added after the U-Net in order to smoothly reduce the number of filters to one, allowing for high precision prediction. The final output was restored to the same size as the input image. A prediction was generated for each pixel, and the pixel-by-pixel prediction was made on the upsampled feature map according to the probability map, thus achieving the goal of contour-to-dose mapping.

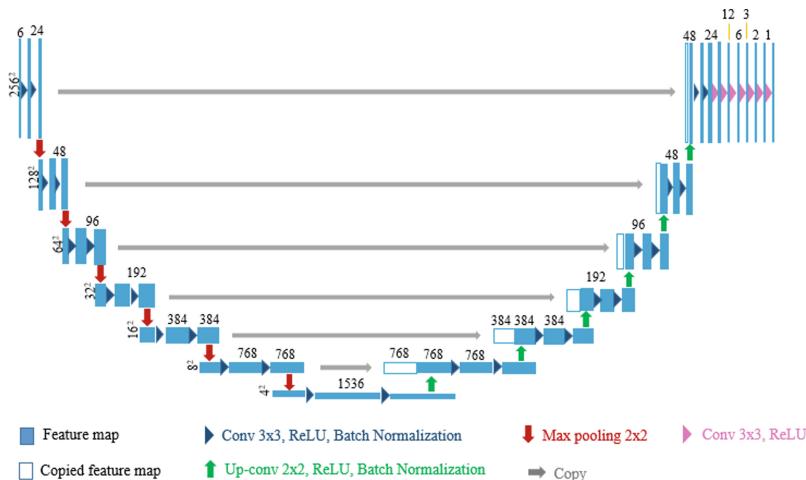


Fig. 1. The framework of the U-Net model used for dose prediction.

2.2 Data and Training

The input of the network was the image with structures, including PTV, the surrounding OAR (left lung, right lung, heart) and outline. The label was the true dose distribution map, and the output of the network was the corresponding predicted dose

distribution map. The network can be trained for end-to-end and pixels-to-pixels without manual intervention. Moreover, the training was more efficient because it avoids the problem of repeated storage and computational convolution caused by the use of pixel blocks.

All input images and dose images were normalized. The U-Net model was trained on single slices, and the loss function adopted the mean square error between the predicted dose and the true dose. The Adam algorithm was chosen as the optimizer to minimize the loss function and the Dice similarity coefficient was used to monitor the accuracy of the network.

3 Evaluation

3.1 Cosine Similarity

Cosine similarity was used to evaluate the accuracy of the predicted dose map. The cosine of the angle between the two vectors x and y is defined as Eq. 1.

$$\cos(x, y) = \frac{x'y}{[(x'x)(y'y)]^{1/2}} \quad (1)$$

The cosine similarity measures the difference between the two individuals by the cosine value of the angle between the two vectors in vector space. Here, the picture was represented as a vector, and the similarity of the two pictures was represented by calculating the cosine distance between the vectors.

3.2 DVH

The DVH curve of the predicted dose distribution was drawn and compared with the true DVH. The mean absolute error of DVH of each ROI was used to calculate the difference between predicted and true DVH parameters from D99% to D1%, thereby evaluating the accuracy of predicted DVH. The formula is shown in Eq. 2.

$$MAE_{DVH} = \frac{1}{99} \sum_{k=1}^{99} \left| (D_p)_{k\%} - (D_t)_{k\%} \right| \Bigg/ \text{Prescription Dose} \times 100\% \quad (2)$$

Where k is the volume percentage of DVH, D_p and D_t are the predicted and true dose at $k\%$ volume. The smaller the value of AME, the smaller the difference between the two curve parameters, that is, the similarity is higher, and the prediction result is more accurate.

4 Results

The loss function of the network varies with the epochs as shown in Fig. 2. After 300 epochs, the training loss was close to 0 and the test loss was 0.01. For test cases, the input images, true and predicted dose distribution maps were shown in Fig. 3. The cosine similarity of the test cases was calculated as shown in Table 1, which ranged from 0.71 to 0.95. The comparison between the predicted DVH curve and the true DVH curve of one of the test cases was shown in Fig. 4. It can be seen from the figure that the predicted DVH curve was similar to the true DVH curve with the same trend. The MAEDVH for each ROI were shown in Table 2. It can be seen that the MAEDVH of the left lung ranged from 1.4% to 7.5%, the right lung ranged from 2.1% to 5.9%, the heart ranged from 3.8% to 9.2%, and the PTV ranged from 2.8% to 5.2%.

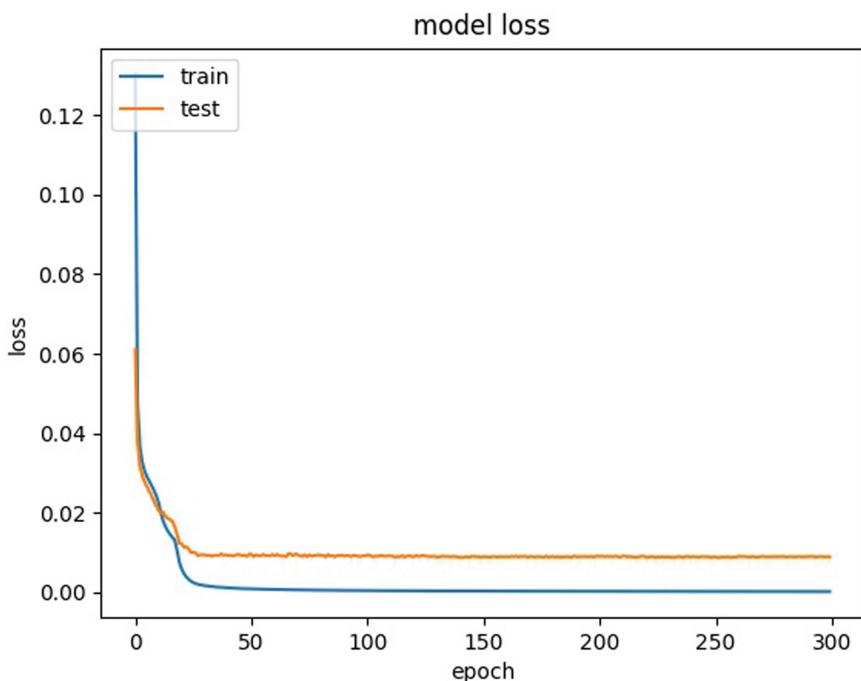


Fig. 2. The loss function of train and test varies with the epochs.

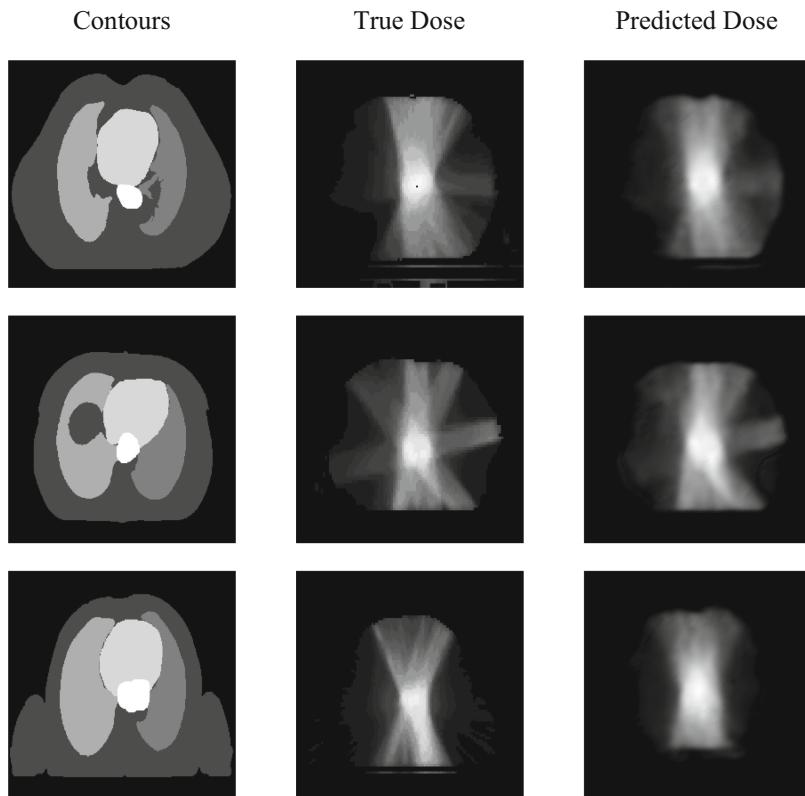


Fig. 3. The input images, true dose distribution maps and predicted dose distribution maps.

Table 1. The cosine similarity of test cases.

| Case | Cosine similarity | Case | Cosine similarity |
|------|-------------------|------|-------------------|
| 1 | 0.912 | 4 | 0.937 |
| 2 | 0.898 | 5 | 0.958 |
| 3 | 0.872 | 6 | 0.715 |

Table 2. Comparison of MAEDVH for each ROI.

| OAR | Case1(%) | Case2(%) | Case3(%) | Case4(%) | Case5(%) | Case6(%) |
|--------|----------|----------|----------|----------|----------|----------|
| Lung L | 7.5 | 4.3 | 3.8 | 2.1 | 4.4 | 1.4 |
| Lung R | 2.8 | 5.1 | 4.7 | 2.1 | 2.4 | 5.9 |
| Heart | 3.8 | 6.4 | 6.8 | 3.8 | 8.9 | 9.2 |
| PTV | 5.1 | 3.6 | 3.6 | 2.8 | 3.4 | 5.2 |

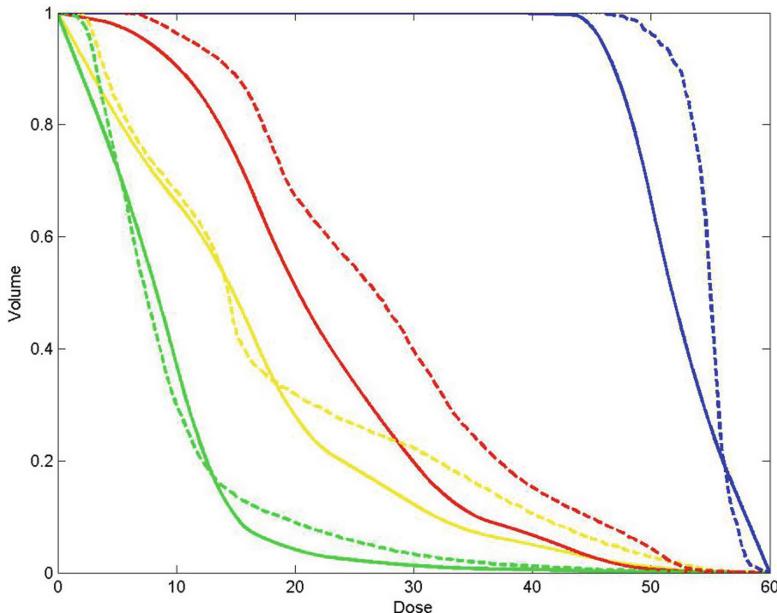


Fig. 4. Comparison of the DVH of a test case for each ROI.

5 Discussion

As far as we know, this is the first dose distribution prediction model for esophageal cancer, so it is difficult to directly compare with the existing models. In typical dose distribution or DVH prediction, task characteristics must be extracted. The manual extraction algorithm is complex and can only cover shallow features for modeling. In this study, we used the U-Net model to predict the dose distribution for esophageal cancer, effectively retaining the dose information and spatial location relationship of all voxels. We generated the input image with structures in which PTV and each OAR have a unique label, the dose distribution map extracted and processed as labels, and the output was the corresponding dose distribution map. Then the DVH curve was drawn according to the predicted dose distribution. For the results, the cosine similarity between the predicted and true dose distribution map ranged from 0.71 to 0.95, and the MAEDVH for each ROI ranged from 1.4% to 9.2%. The results showed that the U-Net model can be used to generate dose distribution for patients with esophageal cancer radiotherapy, and it can be applied to obtain the dose distribution layer by layer.

In the study of Fantu Kong for the prostate, a three-dimensional dose distribution prediction model of SVR was constructed. In his study, the difference between the predicted and the true DVH was -18.3% to 20.9% for the bladder. For head and neck cancer, some studies have implemented DVH prediction for knowledge-based planning. In the study of Junet and his colleagues, the difference between the predicted and true DVH was -17.7% to 15.3%, with the left parotid being -8.7% to 19.6% and the right parotid being -6.8% to 7.5%. Therefore, our study shows that the proposed method of using the U-Net model to predict accurate dose distribution is feasible.

The relatively small amount of data is a common problem for medical images. U-Net can fully combine the simple features of the shallow layer in the decode stage, and it is not easy to overfitting for small samples. Therefore, in the case of small samples, U-Net may be a more suitable model framework and may get better results.

However, this model has some limitations. Firstly, it was not possible for the model to develop a personalized treatment plan according to the specific circumstances of each patient. Secondly, the model was trained layer by layer, ignoring the relationship between the layers in the 3D structure of the patient's CT. Therefore, in the following study, it may be necessary to represent the patient's geometric structure in 3D as far as possible. Combining with the patient's 3D information, the accuracy of the dose distribution prediction can be improved.

With the development of deep learning, it is necessary to test and adopt a more accurate and effective network to predict dose distribution. Considering the advantages and disadvantages of each network model, we can try to integrate different network models. For example, combining U-Net with DenseNet or ResNet can deepen the network appropriately and extract higher-level features without too long training time, so as to achieve dose distribution prediction more accurately and effectively.

6 Conclusion

This study shows the feasibility of generating a dose distribution for radiotherapy with U-Net model. Using the contour of esophageal cancer patients as input, the U-Net model can predict the dose distribution by extracting contour features. Using the U-Net model predicted dose, the mean absolute error of dose difference from D99% and D1% for all ROI is within 10% of the prescription dose, and the cosine similarity between the predicted and the true dose distribution is more than 70%. In the following study, we want to improve the model by combining the patient's 3D information, and consider the integration of different network models to improve the accuracy, so as to ensure the quality of the plan and guide the automatic planning.

Acknowledgments. This work was completed under the guidance of teacher Jiwei Liu and Jianfei Liu, and they provided me with many suggestions. I am very grateful to Jiwei Liu for his rigorous attitude towards academic and his strict requirements for us, which have benefited me greatly. At the same time, I would also like to express my heartfelt thanks to Jianfei Liu for his patience and guidance. This work is supported by the National Natural Science Foundation (NSF) of China (No. 61702001).

References

1. Tran A, Woods K, Nguyen D et al (2017) Predicting liver SBRT eligibility and plan quality for VMAT and 4π plans. Radiat Oncol 12(1):70
2. Wu B, Ricchetti F, Sanguineti G et al (2011) Data-driven approach to generating achievable dose–volume histogram objectives in intensity-modulated radiotherapy planning. Int J Radiat Oncol Biol Phys 79(4):1241–1247

3. Wu B, Pang D, Simari P, Taylor R, Sanguineti G, McNutt T (2013) Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study. *Med Phys* 40(2):021714
4. Yuan L, Ge Y, Lee WR, Yin FF, Kirkpatrick JP, Wu QJ (2012) Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans. *Med Phys* 39(11):6868–6878
5. Lian J, Yuan L, Ge Y et al (2013) Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: an intertechnique and interinstitutional study. *Med Phys* 40(12):121704
6. Folkerts MM, Gu X, Lu W, Radke RJ, Jiang SB (2016) SU-G-TeP1-09: modality-specific dose gradient modeling for prostate IMRT using spherical distance maps of PTV and isodose contours. *Med Phys* 43(6Part26):3653–3654
7. Nguyen D, Long T, Jia X et al (2017) Dose prediction with U-net: a feasibility study for predicting dose distributions from contours using deep learning on prostate IMRT patients. [arXiv:1709.09233](https://arxiv.org/abs/1709.09233)
8. Zhu X, Ge Y, Li T et al (2011) A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. *Med Phys* 38(2):719–726
9. Huang B, Zhu J, Yan X (2016) Development and evaluation of a predicting model of dose volume histograms of parotid in NPC IMRT planning. *Chin J Radiat Oncol* 25(2):150–154
10. Liu X (2017) Applied mathematical models for predicting the dosage of heparin in continuous renal replacement therapy. Dalian Medical University, Dalian
11. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. [arXiv:1411.4038](https://arxiv.org/abs/1411.4038)
12. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. [arXiv:1505.04597](https://arxiv.org/abs/1505.04597)



Primary Equipment Fault Diagnosis Strategy for Box Substation Based on PSO-BP Neural Network

Ao Hou^(✉), Fuzhong Wang, and Jili Ren

Henan Polytechnic University, Jiaozuo 454003, China
651575456@qq.com

Abstract. Box type substations are widely used in industrial, commercial and urban transmission and distributed system, which play an important role in power systems, and has considerable meaning in box type substation's fault diagnosis. Through the study of the box type substation's internal structure, working principle and the analysis of fault and fault characteristics. A fault diagnosis network model based on PSO-BP neural network is proposed to integrate the system data. To avoid the drawbacks of BP neural network algorithm learning and make the model have a good convergence and adaptability, the optimization of BP neural network is carried out by using the optimal characteristics of PSO. The simulation model results that the network has a good recognition effect, and it's very promising in the fault prediction of box type substation.

Keywords: Substation · Fault diagnosis · Particle swarm optimization · BP neural network

1 Introduction

Box type substation [1, 2] has been widely used in industrial and residential areas, its running status directly affects the user's electricity safety and electricity production efficiency. To reduce and prevent the occurrence of the box type substation faults, ensure the safe operation of power-consuming communities and electrical equipment. Domestic and foreign scholars have carried out research on the fault characteristics analysis and extraction, fault diagnosis models of the box type substation. In order to carry out fault diagnosis and prediction, Yang and Liu Xianxing used fault tree analysis technology [3] to constructing a fault tree diagnosis system by analyzing the mechanism of box type substation to complete the function of online fault monitoring. Wang Haoli, Wang Zhibin and others used a multi-neural network to develop smart box type substation fault diagnosis system and complete the fault online diagnosis of box type substation [4].

Cause the nonlinear fitting feature, BP neural network (hereinafter jointly called "BP") is becoming popular in fault diagnosis system in recent years. Comparing with other methods (such as Kriging interpolation, spherical harmonic method or nonlinear fitting), neural network (NN) could provide a more practical fault diagnosis model

because it has the inherent ability of nonlinear function approximation [5]. The search algorithm of the traditional BP is degraded along the gradient. There are many problems in slow training convergence rate and it's easy to fall into the local minimum value. In order to optimizing the BP neural network model, Qi Liwan, Liang Geng and others propose a new idea based on Drosophila algorithm to optimize BP. It reduced the risk of falling into local optimal, and the significantly enhances the generalization ability and global optimization ability of BP [6]. Li Yuyong and Wang Jiayang propose a BP optimization algorithm based on ant colony algorithm. The training error and test error of BP are used to guide the information update mechanism and selection mechanism on the ant path. The experimental results that the model has faster convergence and higher precision than the traditional BP algorithm. Improved the learning and generalization ability of BP [7].

Combine with the mechanism and characteristic of primary equipment failure in box type substation. This paper proposes a fault diagnosis model based on PSO-BP neural network. The BP is used to construct the fault diagnosis model, and particle swarm optimization (PSO) is used to optimize the weight and threshold of the BP. Which could improve the convergence speed of the network and the accuracy of the fault diagnosis, then get the global optimal solution to complete the fault diagnosis.

2 Primary Equipment's Failure Analysis and Feature Extraction of the Box Type Substations

This paper mainly studies the YB1-10/0.4-500 European-style box type substation. The schematic diagram of the main wiring diagram is shown in Fig. 1. The figure shows that the box type substation's primary equipment mainly includes high-voltage busbars, high-voltage vacuum circuit breakers, transformers, low-voltage busbars, low-voltage circuit breakers, capacitor compensators and low-voltage outlets. Different electrical equipment plays different roles and functions. The busbar collects and distributes electric energy. To prevent accidents expanding and ensure safely operating, the circuit breakers could cut off and switch on the load circuit or cutting off the circuit not available. The transformer completes the conversion between high and low voltage. The capacitor compensation cabinet completes the reactive power compensation of the circuit.

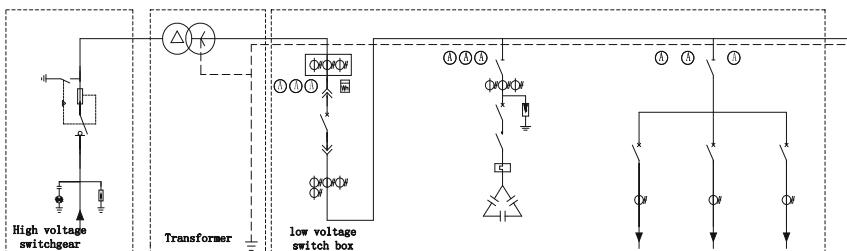


Fig. 1. Main wiring diagram

According to the primary equipment's main wiring diagram of the box type substation in Fig. 1, the faults and causes of each primary equipment are analyzed. The fault characteristic vector of each fault type is proposed by the analysing. The analysing results of box type substation's primary equipment fault characteristics are shown in Table 1.

Table 1. The list of equipment failure characteristics of box type substation

| Fault location | Fault type | Fault cause | Feature vector |
|-------------------------|---|---|--|
| High voltage switchgear | Busbar insulators reduce ground insulation | Insulator aging or looseness, etc. | Busbar insulator pairs low discharge signal |
| | Bad contact between busbar and breaker contact wiring | Partial oxidation, looseness or foreign matter, etc. | Temperature |
| | Circuit breaker vacuum reduction | After many operations, the waveform tube has a leak point, and the vacuum bubble quality is not closed | Circuit breaker vacuum |
| | Circuit breaker refuse to open/reject to close | The closing and closing coils are burned out, damp, etc. | Circuit breaker contact stroke |
| | Circuit breaker malfunction | The operating mechanism is out of order or the relay contacts are loose | Circuit breaker operating mechanical vibration signal |
| Transformer | Internal fault | Oil overheating Oil and paper overheating Influent moisture or bubbles in oil Natural aging Discharge in oil-paper insulation Spark discharge in oil Oil arc Oil and paper arc | H ₂ C ₂ H ₂ CH ₄ C ₂ H ₄ C ₂ H ₆ |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | External fault | Insulation bushing reduces ground insulation | Insulation aging or breakdown |

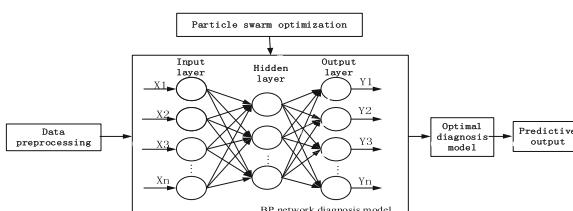
(continued)

Table 1. (continued)

| Fault location | Fault type | Fault cause | Feature vector |
|------------------------|--|---|---|
| Low voltage switch box | The low voltage circuit breaker cannot be closed or a phase contact cannot be closed | There is a problem with the control loop, the moving bridge is stuck or detached, and a connected rod is broken | Low voltage three-phase current |
| | Poor contact in the contact or terminal of the low voltage circuit breaker | Contact wear, loose joints or oxidation of the surface of the tap | Low voltage circuit breaker surface temperature |
| | Poor contact of the low pressure mother discharge line | The shell is aged, oxidized by moisture, or foreign matter is present | Low pressure busbar surface temperature |
| | Reactive power compensation capacitor internal fault | Poor process, insulation aging, moisture, etc. | Reactive compensation capacitor surface temperature |
| | The reactive compensation capacitor fuse is blown | Inferior devices or harmonic induced overcurrent | Three-phase current of capacitor bank outlet |

3 Box Type Substation's Primary Equipment Fault Diagnosis Model

Table 1 shows that many factors causing the failure of the box type substation. There is a quite complex nonlinear relationship on fault symptom and the fault type. The BP has a very strong nonlinear mapping, self-learning, generalized and fault tolerance ability. It also has a strong applicability on complex system's internal fault diagnosis. Therefore, this paper proposes a fault diagnosis model based on PSO-BP network showed in Fig. 2. The BP diagnoses the primary equipment's fault of the box substation and the PSO optimizes the weight and threshold of the neural network. Then the problems could be solved and improving the convergence speed and fault diagnosis accuracy of BP, then the global excellent solution could be obtained [8, 9].

**Fig. 2.** PSO-BP neural network schematic

3.1 Construction of Fault Diagnosis Model for Box Type Substation's Primary Equipment Based on BP Neural Network

Input and Output Layer Designing of Fault Diagnosis Mod. According to the primary equipment failure type and eigenvector of the box type substation listed in Table 1, 17 characteristic input vectors and 18 output vectors of the equipment fault diagnosis model are obtained, which are represented by X1–X17 and Y1–Y18. As shown in Fig. 3, it expected the output is 1 when the fault occurs. For example, when the vacuum of the box type substation breaker is reduced to a critical value, the high-voltage circuit breaker vacuum reduction failure will appear. That is the high-voltage circuit breaker vacuum reduction diagnostic network expects the output equal to 1.

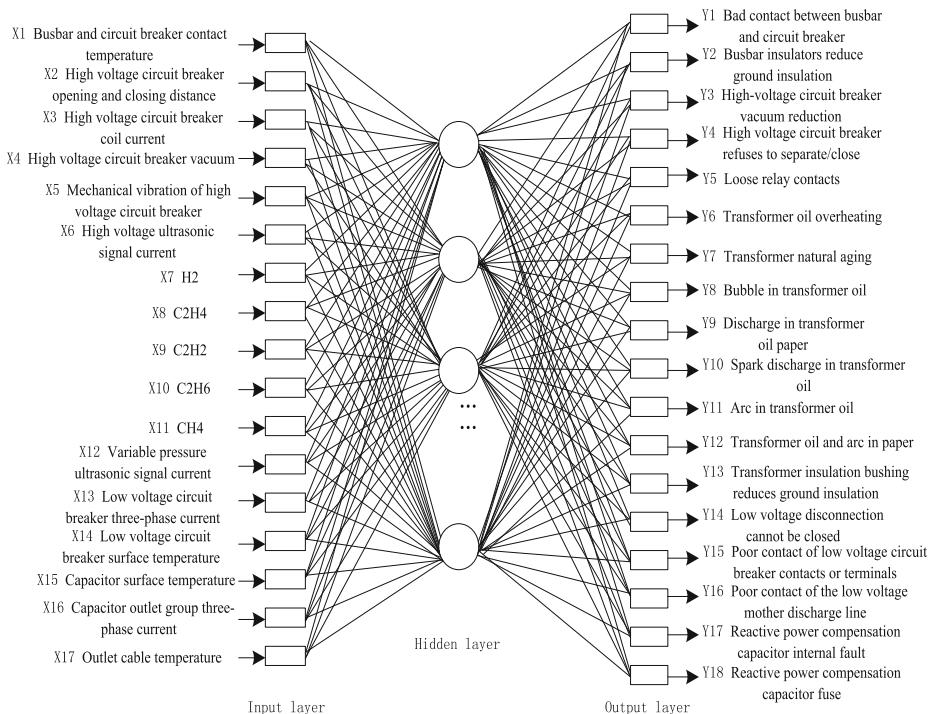


Fig. 3. BP neural network fault diagnosis model

Determining of the Hidden Layers for the Fault Diagnosis Model. The formula shown as (1) is often used to determine the hidden layers. The network training errors obtain by constantly trying and verifying the number of nodes, the errors are listed in Table 2. The “ n_1 ” is the number of hidden layer units, “ n ” is the number of input units, “ m ” is the number of output units, and “ a ” is a constant from [1, 10]. According to the comparison analysis of multiple experiments and network training errors, the network

performance is the best while “ a ” is 7, so the fault diagnosis network’s hidden layer is set to 13.

$$n_1 = \sqrt{n+m} + a \quad (1)$$

Table 2. Network training errors

| Number of hidden layer nodes | Mean square error |
|------------------------------|-------------------|
| 9 | 0.00912 |
| 10 | 0.00871 |
| 11 | 0.00894 |
| 12 | 0.00873 |
| 13 | 0.00705 |
| 14 | 0.00748 |
| 15 | 0.00792 |

Normalization. Because the fault characteristics of the primary equipment of the box type substation have different dimensions. Directly input fault features of different dimensions will seriously affect the accuracy of network diagnosis. In order to avoid the impact of different dimensions on fault diagnosis, it is necessary to preprocess the equipment fault characteristic data of the box type substation, which is normalizing the original fault characteristic parameters. The normalization original vector is $x_i = (x_{i1}, x_{i2}, x_{i3} \dots \dots x_{i17})$, the normalized data is $x'_i = (x'_{i1}, x'_{i2}, x'_{i3}, x'_{i4} \dots \dots x'_{i17})$, and the normalization formula is shown as 2.

$$x'_{ip} = x_{ip} / \sum_{p=1}^{17} x_{ip} (p = 1, 2, 3, \dots, 17) \quad (2)$$

3.2 Optimizing BP Neural Network with Particle Swarm Optimization

In order to solve the problem that BP has a slow convergence speed and easy to fall into local optimum, the PSO is introduced to optimize BP’s weights and thresholds [10–12]. The position of the particles represents the set of BP’s weights of the current iteration. The number of thresholds and the number of weights in the BP determines the dimension of each particle, respectively. The mean square error of the network output is gradually reduced by changing the search speed of the particles in the weight space. Particle swarms obtain optimal fitness and minimum error by continuously optimizing the weights and thresholds. The algorithm optimization flow chart is shown in Fig. 4.

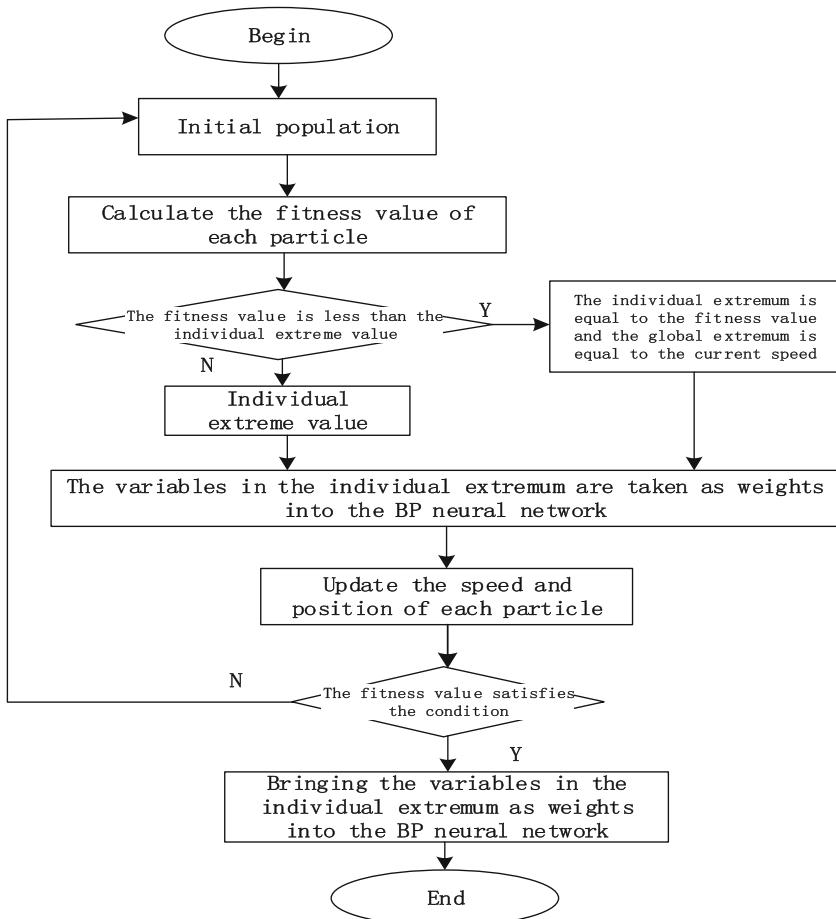


Fig. 4. PSO-BP flowchart

4 Simulation Analysis

4.1 Simulation Experiment Sample Selection

In order to improve the performance of the fault diagnosis model (taking 17 kinds of fault characteristics of box-type substation primary equipment as an example), set 200 groups of sample data with special representativeness and extensiveness as training samples to train the network. Table 3 shows some data after pre-processing of 200 sets of primary equipment in box-type substation. The preprocessed feature vector is represented by $x'_1-x'_{17}$.

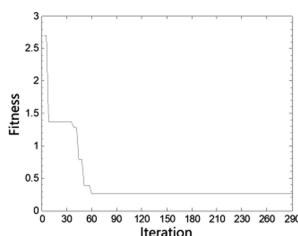
Table 3. Switch cabinet training sample data

| Numbering | 1 | 2 | ... | 18 | ... | 46 | ... | 200 |
|-----------|-----------|-------|-------|-----|-------|-----|-------|-------|
| Sample | x'_1 | 0.046 | 0.043 | ... | 0.042 | ... | 0.044 | ... |
| | x'_2 | 0.038 | 0.031 | | 0.041 | | 0.041 | 0.049 |
| | x'_3 | 0.044 | 0.047 | | 0.022 | | 0.035 | 0.034 |
| | x'_4 | 0.033 | 0.046 | | 0.031 | | 0.044 | 0.038 |
| | x'_5 | 0.047 | 0.056 | | 0.046 | | 0.039 | 0.051 |
| | x'_6 | 0.04 | 0.049 | | 0.048 | | 0.033 | 0.057 |
| | x'_7 | 0.053 | 0.062 | | 0.037 | | 0.045 | 0.047 |
| | x'_8 | 0.067 | 0.036 | | 0.055 | | 0.037 | 0.033 |
| | x'_9 | 0.065 | 0.037 | | 0.032 | | 0.034 | 0.032 |
| | x'_{10} | 0.056 | 0.096 | | 0.038 | | 0.026 | 0.034 |
| | x'_{11} | 0.051 | 0.032 | | 0.032 | | 0.048 | 0.042 |
| | x'_{12} | 0.048 | 0.034 | | 0.033 | | 0.046 | 0.026 |
| | x'_{13} | 0.055 | 0.041 | | 0.036 | | 0.042 | 0.052 |
| | x'_{14} | 0.096 | 0.048 | | 0.081 | | 0.054 | 0.032 |
| | x'_{15} | 0.052 | 0.038 | | 0.041 | | 0.037 | 0.043 |
| | x'_{16} | 0.047 | 0.06 | | 0.052 | | 0.04 | 0.012 |
| | x'_{17} | 0.035 | 0.021 | | 0.067 | | 0.032 | 0.031 |

4.2 Analysis of Simulation Results

BP's fault diagnosis model based on particle swarm optimization are designed by Matlab software. The first 150 groups of sample data are set as training in the simulation program and the last 50 groups are used as training sample test. Performance is evaluated by two criteria: convergence speed and accuracy.

The PSO will terminate if reaches the maximum number of iterations. The particle dimension is the number of weights and thresholds that need to be optimized. The particle group's particle number is 40, the maximum allowable iteration number is 290, $c_1 = c_2 = 1.5$ (acceleration factor), $w_{min} = 0.4$ (minimum inertia coefficient), $w_{max} = 1.4$ (maximum inertia coefficient). Limit the particle's speed range to $[-0.5, 0.5]$. The fitness curve of PSO optimize BP's weights and thresholds is shown in Fig. 5.

**Fig. 5.** Fitness curve

It can be seen from Fig. 5 that after about 60 evolutions, the fitness is closes and reaches about 0.3. The weight at this time is the optimal connection weight of the primary equipment of the box substation. This weight can be used to train the fault diagnosis network to obtain the optimal diagnostic model. In order to show the accuracy of the PSO-BP neural network fault diagnosis model for the diagnosis of box-type substation. Single BP model and the PSO-BP neural network are compared and the training error results are shown in Fig. 6.

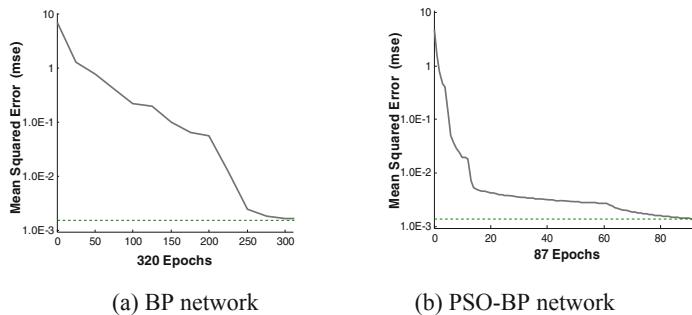


Fig. 6. Training error curve

It can be seen from Fig. 6 that under the premise of the same training samples and prediction samples. The output of the two network models for system fault diagnosis is very different. The BP training speed is slower than the BP optimized by PSO algorithm. To reflect the performance of the PSO-BP neural network fault diagnosis model, the remaining 50 groups of 17 fault feature quantities are set as two fault diagnosis models. The fault diagnosis accuracy is compared and analyzed and the results are shown in Table 4. It can be seen from the table that the PSO-BP neural network has higher correct rate of fault diagnosis and could reach the requirements of device fault diagnosis.

Table 4. Comparison of diagnostic accuracy of different algorithms for primary equipment in box substation

| Algorithm type | Total number of samples | Correct identification number | False judgment | Diagnostic accuracy |
|----------------|-------------------------|-------------------------------|----------------|---------------------|
| BP | 50 | 36 | 14 | 72% |
| PSO-BP | 50 | 47 | 3 | 94% |

5 Conclusion

This paper proposes a method of optimizing BP model by particle swarm. And it can be used to predict the fault of box-type substation. Particle swarms are introduced to avoid local minima and improve convergence speed, which is a defect found in classical

BP. By comprehensively evaluating the performance of the model from the same training sample and prediction samples and comparing the simulation and performance of the two models, we can realize that PSO-BP neural network is significantly better than single BP. This method improves the convergence speed, accuracy and enables excellent fault prediction for box-type substation fault data. It has a great application value in the box substation fault diagnosis system.

References

1. Chen C, Yan K, Chen Y (2013) Discussion on a novel medium voltage box type substation. *High Volt Appar* 49(05):124–127
2. Jiao X, Jing B, Huang Y et al (2017) Research on fault diagnosis of airborne fuel pump based on EMD and probabilistic neural networks. *Microelectron Reliab* 75:296–308
3. Yang T, Liu X (2012) Research on fault diagnosis system of box substation based on FTA. *Autom Instrum* 22(06):20–22
4. Wang H, Wang Z, Dan Y (2004) Research on intelligent box substation fault diagnosis system. *J Jiamusi Univ (Nat Sci Ed)* (03):320–322
5. Song R, Zhang X, Zhou C et al (2018) PredictingTEC in China based on the neural networks optimized by genetic algorithm. *Adv Space Res* 62(4):745–759
6. Qi L, Liang G, Tong G (2014) Gearbox fault diagnosis based on drosophila algorithm optimized BP neural network. *Power Syst Clean Energy* 30(09):31–36+42(2014)
7. Li Y, Wang J, Guo C et al (2010) BP network optimization algorithm based on ant colony algorithm. *J Comput Appl* 30(06):1513–1515+1518 (2010)
8. Tseng H-E, Chang C-C, Lee S-C et al (2018) A block-based genetic algorithm for disassembly sequence planning. *Expert Syst Appl* 96:492–505
9. Son DD, Abhary K, Marian R (2017) An innovative framework for designing genetic algorithm structures. *Expert Syst Appl* 90:196–208
10. Liu H, Zhao C, Li X et al (2016) A neural network optimization algorithm based on improved genetic algorithm. *Chin J Sci Instrum* 37(07):1573–1580
11. Liu P, Li X (2016) A fault segment location algorithm for distribution network with distributed power based on multi-population genetic algorithm. *Power Syst Prot Control* 44(02):36–41
12. Zhang Q, Hu N, Li H (2016) GA-BP fault diagnosis of mine hoist brake. *J Liaoning Tech Univ (Nat Sci Ed)* 35(02):155–159



Defect-Oriented Test Case Prioritization

Jinfeng Li^{1(✉)}, Yan Zhang¹, and Jilong Bian²

¹ College of Computer and Information Technology,
Mudanjiang Normal University, Mudanjiang 157011, China
allylili0453@163.com

² College of Information and Computer Engineering,
Northeast Forestry University, Harbin 150040, China

Abstract. In order to find more software defects during software testing within a short period of time, test case prioritization by ordering test cases is studied. This paper introduces four influence factors: defect detection number, defect detection rate, defect rank number and defect rank detection rate. In order to facilitate the tester to balance the four influence factors according to testing requirement, four weight factors are introduced. Meanwhile, we developed an on-line adjustment strategy to adjust the priorities of test cases. Experimental results show that compared with the traditional test case prioritization, the proposed method can find more defects within a short period of time and improve the testing efficiency.

Keywords: Test case · Prioritization · Software defect · On-line adjustment

1 Introduction

In the process of software development, it is necessary to test and correct the software according to the results of the test. It needs to constantly update and add new test cases. These test cases can be reused in the next test. In order to reduce the cost of testing, the following three methods are usually adopted: test case selection, test case reduction and test case prioritization. The third method is to assign each test case a priority and sort it without reducing the number of test cases, so that the test cases with higher priority are executed first.

The paper proposes a test case prioritization technique based on 4 factors including defect detection number (*DN*) [1], defect detection rate (*DR*), defect rank number (*DRN*) [2] and defect rank detection rate (*DRR*). In order to normalize these factors, we introduce the sigmoid [3] function. We adjust the proportion of each factor through the weight, and test case priority is on-line adjusted. It is verified that the proposed technique can detect more defects in limited time and improve the efficiency of test cases compared with other techniques.

2 Prior Studies

Different test cases have different effects on the software testing objectives [4], and can be sorted according to the contribution degree to the test objectives, so that the test cases with larger contribution are executed first. Wong et al. [5] introduced the test case

prioritization technique for the first time, giving priority to the test cases according to the code coverage ability. Rothermel [6], Elbaum [7] and Malishevsky [8] have comprehensively studied various test case prioritization techniques and made a large number of experiments to analyse the defect detection ability of each technique. In literature [6], prioritization techniques were divided into three categories, one was based on the code coverage, the second was based on the unsurfaced code coverage, and the third was based on the ability of test cases to detect defects to sort test cases. Walcott [9] and Chang [10] proposed a test case prioritization technology based on historical information. Literature [9] prioritized test cases based on the historical execution time of test cases, while literature [10] considered two factors: defect detection and the execution stability of test case.

In summary, although there have been many researches on test case prioritization techniques, the test results still need to be improved. We hope that the test cases executed first can detect as many defects as possible and the most destructive defects can be detected as soon as possible. We also need to consider the execution efficiency of test cases. Therefore, this paper proposes a test case prioritization technique based on 4 factors: defect detection number (DN), defect detection rate (DR), defect rank number (DRN) and defect rank detection rate (DRR).

3 Basic Concepts

To facilitate the description, we give the following definitions.

Definition 1 (Initial Test Case Set T). $T = \{t_1, t_2, \dots, t_n\}$ is a set of test cases generated for the testing objectives of software system. Taking the experimental data in literature [2] as an example, the initial test case set $T = \{t_1, t_2, \dots, t_{10}\}$ contains 10 test cases, as shown in Table 1.

Table 1. Test case along with detection defects and time

| | t_1 | t_2 | t_3 | t_4 | t_5 | t_6 | t_7 | t_8 | t_9 | t_{10} | β_j |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|-----------|
| f_1 | | | | | | | | ✓ | ✓ | | 6 |
| f_2 | | ✓ | ✓ | | ✓ | | | | | | 4 |
| f_3 | | | | ✓ | | ✓ | | | | ✓ | 2 |
| f_4 | | ✓ | ✓ | | | | | | | | 2 |
| f_5 | | | | | | | | ✓ | | | 6 |
| f_6 | | | | | | | | ✓ | ✓ | | 6 |
| f_7 | | | | ✓ | ✓ | | ✓ | | | | 4 |
| f_8 | ✓ | | | | | ✓ | | | | | 4 |
| f_9 | | | | ✓ | ✓ | | | | | ✓ | 4 |
| f_{10} | ✓ | | | | | | | ✓ | | | 2 |
| T_i | 9 | 8 | 14 | 9 | 12 | 14 | 11 | 10 | 10 | 13 | |

Definition 2 (Test Case Prioritization). Test cases are sorted according to some criteria so that test cases with high priority can be executed first. Rothermel et al. [6] formally defined the test case prioritization as: given a test suite T , a set PT containing all permutations of T , and a function f mapping each element in PT to a real number. The purpose of sorting test cases is to find a $T' \in TP$ such that $(\forall T'' \in TP)[f(T') \geq f(T'')]$. The function f is a quantitative description of sorting objectives, which is used to measure the effectiveness of sorting. Generally speaking, the higher the value, the better the sorting effect of the test case set, defect detection efficiency and coverage ability. The purpose of this paper is to find such a test case ordering and detect more defects in a short period of time.

Definition 3 (Defect Set F). Defects will lead to software products cannot meet the needs of users to some extent. In this paper, $F = \{f_1, f_2, \dots, f_m\}$ is the defect set of software system. The defect set shown in Table 1 contains 10 defects.

Definition 4 (Defect Rank Matrix DM). We divides software defects into five levels: the least severe, the less severe, the medium severe, the serious defect and the fatal defect, and quantifies them with numerical values. The severity value of least severe defects is 2, the severity value of less severe defects is 4, the severity value of medium severe defects is 6, the severity value of serious defect is 8 and the severity value of fatal defects is 10. DM is a matrix of $1 \times m$. Each element β_j in the matrix stores the severity value corresponding to this defect. As shown in the last column of Table 1, the severity value of defects $(f_1, f_2, \dots, f_{10})$ constitute defect rank matrix $DM = [6, 4, 2, 2, 6, 6, 4, 4, 4, 2]$.

Definition 5 (Coverage Matrix CM). CM is a Boolean matrix of $n \times m$, which is defined as a coverage matrix from test case T to defects set. The elements in the matrix are defined as follows:

$$\delta_{ij} = \begin{cases} 1 & t_i \text{ located } f_j \\ 0 & t_i \text{ unlocated } f_j \end{cases} \quad (1)$$

Definition 6 (Test Case Execution Time T_i). T_i represents the execution time of the i th test case running in the software system to locate defects in ms.

4 Test Case Prioritization

4.1 Influence Factors of Prioritization

In the process of prioritization of test cases, we introduce 4 influence factors:

- (1) Defect Detection Number (DN_i): DN_i depicts the total number of defects detected by the i th test case. Mathematically, it can be expressed as

$$DN_i = \sum_{j=1}^m \delta_{ij} \quad (2)$$

The defect detection number of each test case shown in Table 1 is the sum of each row in coverage matrix CM , respectively. $DN_1 = 0 + 0 + 0 + 0 + 0 + 0 + 0 + 1 + 0 + 1 = 2$. The larger the value, the more defects detected by the test case in the execution.

- (2) Defect Detection Rate (DR_i): DR_i refers to the number of defects detected by the i th test case in a unit time. That is, the ratio of defect detection number to the running time of test case. It can be mathematically represented as

$$DR_i = \frac{DN_i}{T_i} \quad (3)$$

Such as $DR_1 = 2/9 = 0.22$, $DR_2 = 0.25$. The larger the value, the more defects detected by the test case in a unit time and the higher the detection efficiency.

- (3) Defect Rank Number (DRN_i): DRN_i depicts the sum of severity values of all the defects located by the i th test case, the equation is as follows:

$$DRN_i = \sum_{j=1}^m \delta_{ij} \beta_j \quad (4)$$

Such as $DRN_1 = 0 \times 6 + 0 \times 4 + 0 \times 2 + 0 \times 2 + 0 \times 6 + 0 \times 6 + 0 \times 4 + 1 \times 4 + 0 \times 4 + 1 \times 2 = 6$. The larger the value, the higher the sum of severity values of defects detected by the test case, the stronger the damage ability of the defects.

- (4) Defect Rank Detection Rate (DRR_i): DRR_i depicts the sum of severity values of all the defects located by the i th test case in a unit time. That is, the ratios of defect rank number to the running time of test case. It can be mathematically represented as

$$DRR_i = \frac{SL_i}{T_i} \quad (5)$$

Such as $DRR_1 = 6/9 = 0.667$. The larger the value, the stronger the destructive ability of defects detected by the test case in a unit time.

4.2 Priority Calculation Method

The 4 factors (DN , DR , DRN , DRR) are used as defects detection criteria. In order to make these factors at the same order of magnitude, sigmoid function is applied to normalize each influence factor as shown in formula (6), so that the values are between 0 and 1.

$$S(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

In order to improve the efficiency of the test, software tester can adjust the proportion of factors by setting weight according to the specific testing requirements and

the importance of each factor. We define the priority value of the i th test case as shown in formula (7).

$$RT_i = w_1S(DN_i) + w_2S(DR_i) + w_3S(DRN_i) + w_4S(DRR_i) \quad (7)$$

where w_1, w_2, w_3, w_4 represent the weight of each factor, and satisfy $w_1 + w_2 + w_3 + w_4 = 1$, thus adjusts the proportion of the four factors.

4.3 Priority On-line Adjustment Strategy

T is the set of initial test cases and F is the set of defects detected by all test cases. The specific steps of the algorithm as follows:

Step 1: Set the weights (w_1, w_2, w_3, w_4) of 4 influence factors, initialize the defect rank matrix (DM), and initialize the coverage matrix (CM) according to the coverage relation between the test cases and the defect set.

Step 2: Calculate DN, DR, DRN, DRR of each unsorted test case by using the formulas (1)–(5).

Step 3: Calculate the priority of each test case by using formulas (6) and (7), and select the test case t_i with the highest priority and add it into the set T' , T' is the ordered set of test cases.

Step 4: Adjust the set T , delete the test case t_i from the set T and go to step 6 when the set T is empty. Otherwise adjust the elements in the coverage matrix CM , adjust the elements in the column of the non-zero element located by the test case t_i to 0, indicating that the defects covered by the test case have been found, and the unsorted test cases are no longer locate the defects. Adjust the defect set F , delete the defects detected by the test case t_i . If the defect set F is not empty and the unsorted test cases locate this defect set, go to step 2.

Step 5: Reinitialize the defect set F to the original set. Reassign the elements of the unsorted test case in the matrix CM according to the initial defect coverage relationship and go to step 2.

Step 6: Output the test case sorting set T' .

The relevant data are shown in Table 1 and the weights are $w_1 = 0.2, w_2 = 0.27, w_3 = 0.23, w_4 = 0.3$. The tester can adjust the weights according to the specific requirements in actual test. The values of the influence factors (DN, DR, DRN, DRR) and priorities (RT_i) of the test cases determined by Eqs. (1)–(7). According to the priority values, the test case t_8 with the highest priority is selected in this round and added to the set T' . Adjust the test case set, delete the test case t_8 from the set T and get unsorted test case set $T = \{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_9, t_{10}\}$. The test case t_8 can locate defects f_1, f_5, f_6, f_{10} . Adjust the defect set F , delete the 4 defects detected by the test case and get the undetected defect set $F = \{f_2, f_3, f_4, f_7, f_8, f_9\}$. Adjust the values of elements in the coverage matrix CM and assign the values of columns 1, 5, 6 and 10 to 0.

According to the undetected defect set F , the four influence factor values of the unsorted test cases are calculated by formulas (1)–(5), and the priority values of the unsorted test cases are recalculated by formulas (6) and (7). The test case t_4 is selected

by the second round. According to the algorithm proposed in this paper, we select the first 4 test cases $T' = \{t_8, t_4, t_2, t_1\}$ and the defects set F is empty at the same time. The defect set is reset to the initial value $F = \{f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9, f_{10}\}$, and the coverage matrix CM is adjusted to the initial matrix. Based on this, the priorities of unsorted test cases is calculated and sorted. When the first 8 test cases ($T' = \{t_8, t_4, t_2, t_1, t_9, t_6, t_5, t_3\}$) are obtained, the defect set is $F = \{f_5, f_{10}\}$, but the unsorted test cases t_7 and t_{10} do not locate the defects in the defect set. At this time, the coverage matrix CM should be reset to the initial matrix, and the defect set is reset to the initial value until all the test cases are sorted, and the sorting test case set is $T' = \{t_8, t_4, t_2, t_1, t_9, t_6, t_5, t_3, t_{10}, t_7\}$.

5 Analysis of Experimental Results

This part mainly verifies the effectiveness of the proposed method through experiments. We use average of the percentage of faults detected (APFD) as the evaluation standard of the test case ordering. The value varies between 0 and 1. The larger the value, the better the defect detection ability of the test case ordering. This can be formally stated by the following equation:

$$APFD = 1 - \frac{TF_1 + TF_2 + \cdots + TF_m}{nm} + \frac{1}{2n} \quad (8)$$

where n is the total count of test cases in the test case set T , m is the total count of defects detected by the test case set T , TF_i is the location of the first test case in test suite T that reveals defect i .

Table 2. The ordering of test cases for various methods

| Method | Ordering of test cases | | | | | | | | | |
|----------------|------------------------|-------|-------|-------|-------|-------|-------|----------|----------|----------|
| | t_1 | t_2 | t_3 | t_4 | t_5 | t_6 | t_7 | t_8 | t_9 | t_{10} |
| NOP | t_2 | t_9 | t_5 | t_7 | t_4 | t_6 | t_8 | t_{10} | t_3 | t_1 |
| ROP | t_8 | t_4 | t_9 | t_6 | t_5 | t_2 | t_1 | t_{10} | t_3 | t_7 |
| Literature [2] | t_8 | t_4 | t_2 | t_1 | t_9 | t_6 | t_5 | t_3 | t_{10} | t_7 |
| Our method | t_8 | t_4 | t_2 | t_1 | t_9 | t_6 | t_5 | t_3 | t_{10} | t_7 |

According to the data provided in Table 1, we compare the method proposed in this paper with NOP, ROP and the method proposed in literature [2]. Table 2 represents ordering of test cases obtained by different techniques. The APFD values obtained from the test case ordering of various techniques are shown in Fig. 1. The APFD values of NOP, ROP, literature [2] and this paper are 0.63, 0.69, 0.8, 0.85.

It can be perceived that the new technique needs only 40% of test cases to locate all the defects, whereas 80% of test cases are required in the NOP, 70% of test cases are required in the ROP, and 60% of test cases are required to detect all defects in literature [2].

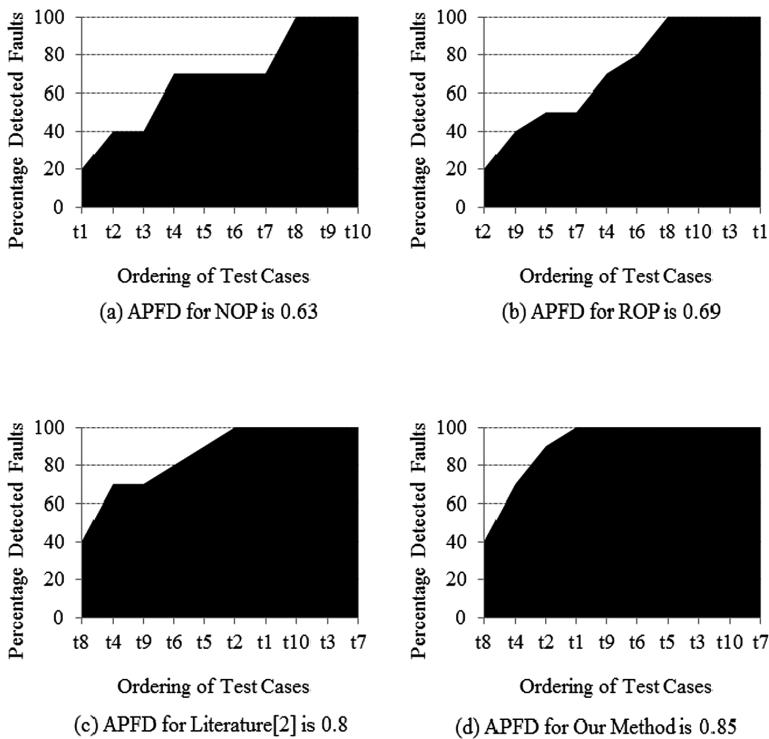


Fig. 1. APFD for various prioritization techniques

Table 3. Values of influence factors and priorities of test cases

| | t_1 | t_2 | t_3 | t_4 | t_5 | t_6 | t_7 | t_8 | t_9 | t_{10} | β_j |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|-----------|
| f_1 | | ✓ | | | ✓ | | | | ✓ | | 10 |
| f_2 | ✓ | | ✓ | ✓ | | ✓ | | ✓ | | ✓ | 8 |
| f_3 | | ✓ | | | ✓ | | ✓ | | | | 8 |
| f_4 | | | | | ✓ | ✓ | | ✓ | ✓ | | 2 |
| f_5 | | | | ✓ | | | ✓ | | | | 2 |
| f_6 | ✓ | | | ✓ | | | | ✓ | | | 4 |
| f_7 | | | | | | | ✓ | | ✓ | ✓ | 2 |
| f_8 | | | | | ✓ | | | | | | 10 |
| f_9 | | | | | ✓ | | | | ✓ | | 4 |
| f_{10} | | | | | | | ✓ | ✓ | | | 4 |
| T_i | 9 | 8 | 14 | 9 | 12 | 14 | 11 | 10 | 10 | 13 | |

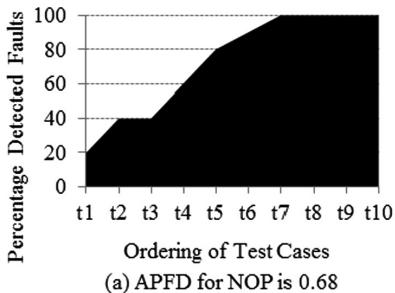
In order to further verify the effectiveness of the method proposed in this paper, we conduct the second experiment and select a software for test analysis. Design 10 test cases and get 10 software defects according to the first test, which are shown in

Table 3. The test case ordering obtained by the NOP, ROP, literature [2] and the method in this paper are shown in Table 4, and the corresponding APFD values are 0.68, 0.76, 0.85 and 0.87 as shown in Fig. 2. It can be perceived that the new technique needs only 30% of test cases to detect all the defects, whereas 70% in NOP, 80% in the ROP, 40% in the literature [2].

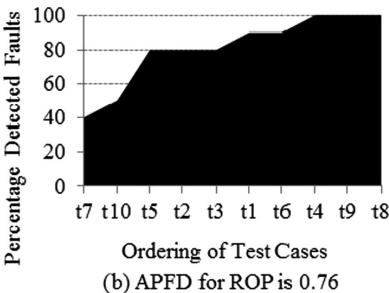
Through the analysis of the two groups of experiments, we can see that the technique proposed in this paper has the highest APFD compared with the NOP, ROP and literature [2], the generated test case ordering with the highest defect detection efficiency can detect all defects by executing fewer test cases.

Table 4. The ordering of test cases for various methods

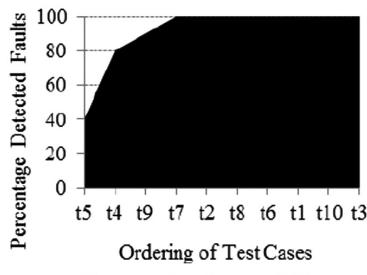
| Method | Ordering of test cases | | | | | | | | | |
|----------------|------------------------|----------|-------|-------|-------|-------|-------|-------|----------|----------|
| NOP | t_1 | t_2 | t_3 | t_4 | t_5 | t_6 | t_7 | t_8 | t_9 | t_{10} |
| ROP | t_7 | t_{10} | t_5 | t_2 | t_3 | t_1 | t_6 | t_4 | t_9 | t_8 |
| Literature [2] | t_5 | t_4 | t_9 | t_7 | t_2 | t_8 | t_6 | t_1 | t_{10} | t_3 |
| Our method | t_5 | t_4 | t_7 | t_9 | t_1 | t_2 | t_6 | t_8 | t_{10} | t_3 |



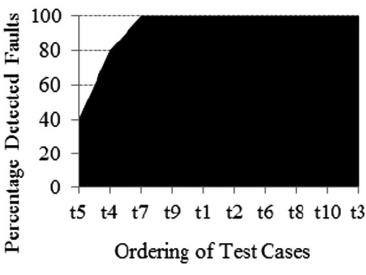
(a) APFD for NOP is 0.68



(b) APFD for ROP is 0.76



(c) APFD for Literature[2] is 0.85



(d) APFD for Our Method is 0.87

Fig. 2. APFD for various prioritization techniques

6 Conclusion

In this paper, the 4 influence factors are taken as the evaluation criteria of priority. The sigmoid function is used to normalize the factors, so that the values vary between 0 and 1. It is convenient for the tester to adjust the proportion of each factor according to the actual test requirements by setting the weight for each factor in the priority. In the process of test case priority calculation, the priority on-line adjustment strategy is used to adjust the priority of each unsorted test case. Through two sets of experiments, it is verified that the proposed method can detect defects in software as soon as possible and improve detection efficiency in a short period time.

Acknowledgments. This work was supported by the Foundation of Heilongjiang Education Department (No. 1354MSYYB003, No. 1353ZD003), the Natural Science Foundation of Heilongjiang Province (No. F2016039).

References

1. Nayak S, Kumar C, Tripathi S (2017) Enhancing efficiency of the test case prioritization technique by improving the rate of fault detection. *Arab J Sci Eng* 42(11):1–17
2. Kavitha R, Sureshkumar N (2010) Test case prioritization for regression testing based on severity of fault. *Int J Comput Sci Eng* 2(5):1462–1466
3. Han J, Moraga C (1995) The influence of the sigmoid function parameters on the speed of backpropagation learning. In: Proceedings of the international workshop on artificial neural networks. Springer, Berlin, pp 195–201
4. Yoo S, Harman M (2012) Regression testing minimization, selection and prioritization: a survey. *Softw Test Verif Reliab* 22(2):67–120
5. Wong W, Horgan J (1997) A study of effective regression testing in practice. In: Eighth international symposium on software reliability engineering. IEEE, Albuquerque, pp 264–274
6. Rothermel G, Untch R, Chu C: Test case prioritization: an empirical study. In: Proceedings of IEEE international conference on software maintenance. Oregon State University, Corvallis, pp 179–188 (1999)
7. Elbaum S, Malishevsky A, Rothermel G (2001) Prioritizing test cases for regression testing. *IEEE Trans Softw Eng* 10(27):929–948
8. Malishevsky A, Ruthruff JR, Rothermel G (2006) Cost-cognizant test case prioritization. Department of CSE, University of Nebraska-Lincoln (2006)
9. Walcott K, Sofya M, Kapfhammer G (2006) Time-aware test suite prioritization. In: Proceedings of the international symposium on software testing and analysis. ACM, New York, pp 1–12
10. Chang L, Miao H, Xiao L (2015) Self-adaptive test case prioritization based on history information. *Comput Sci* 42(9):154–158



A Positioning Method Using Teaching-Learning Based Optimization for GNSS

Jing Zhang, Jinheng Zhang, Zihang Ni, and Lan Cheng^(✉)

College of Electrical and Power Engineering, Taiyuan University of Technology,
Taiyuan 030024, China
taolan_1983@126.com

Abstract. Satellites based positioning has been widely used in our daily lives and thus become indispensable, which also leads to increasing demand on high-positioning accuracy. However, multipath error is one of the main sources in deteriorating positioning accuracy, and it is difficult to eliminate multipath error using differential techniques due to multipath uncertainty of occurrence and uncorrelation in different instants. To address this problem, a positioning method is proposed for global navigation satellite systems (GNSS) by adopting a teaching-learning based optimization (TLBO) algorithm after the positioning problem is formulated as an optimization problem. The TLBO algorithm simulating the teacher's teaching process and the student's learning process has the advantages of strong convergence ability, fast convergence speed and few parameters. Some experiments are conducted based on actual satellite data. The results show that the proposed positioning method outperforms other two methods, such as differential evolution (DE)-based positioning method and particle swarm optimization (PSO)-based positioning method in terms of accuracy and stability.

Keywords: GNSS positioning · Multipath error · TLBO algorithm

1 Introduction

For GNSS, multipath interference refers to the replica version of direct signal caused by the reflection of the surrounding obstacles near a receiver. Multipath error is difficult to eliminate by modeling, filtering and differential techniques due to the multipath uncertainty of occurrence and uncorrelation in different instants.

Different techniques for repressing the error caused by multipath interference have been proposed. For most of these algorithms, it is essential to model and estimate multipath signal according to its characteristics [1, 2]. However, there are always a certain multipath estimation error when these methods are adopted. In addition, the methods by modifying the receiver's antenna [3, 4] or the receiver's discriminator and correlator have also become a mode of multipath repression [5, 6]. Unfortunately, these methods have the drawbacks of being hardware-consuming or sensitive to noises. Therefore, it is necessary to come up with a new method to suppress multipath error.

On the other hand, it has been found that there is a strong correlation between multipath signal and signal-to-noise ratio (SNR) of the received signal, thus through

measuring the SNR, the multipath error can be estimated by measuring pseudorange [7]. Although this method is only limited to the case of one dominant specular reflection when it is first proposed, it indicates that multipath error can be eliminated by using pseudorange information. More research on this topic has also been done by other researchers [8–10].

In this paper, the pseudorange model is first transformed into an optimization problem. Then a teaching-learning based optimization (TLBO) algorithm is applied to solve the optimization problem. To further improve the positioning accuracy, a residual phase method to improve the accuracy of pseudorange measurement proposed in [11] is adopted before applying the proposed method.

The paper is organized as follows. Section 2 describes the TLBO algorithm. In Sect. 3, the positioning algorithm based on TLBO algorithm is introduced. Experiments and analysis are presented in Sect. 4. Finally, the conclusion of this paper is drawn in Sect. 5.

2 TLBO Algorithm

Teaching and learning optimization (TLBO) algorithm is proposed by Rao et al., which finds the optimal solution to an optimization by simulating the teacher's teaching process and the student's learning process [12]. The optimization process is divided into two phases: the teacher phase and the learner phase.

2.1 Teacher Phase

For the best individual in the teaching process, the teacher keeps the students' learning level close to themselves, thus raising the level of the whole class. The teaching operation is shown in Eq. (1).

$$x_i^{new} = x_i^{old} + rand(0, 1)(x^{best} - T_F \bar{x}) \quad (1)$$

where, x_i^{old} and x_i^{new} represent the grade of the i th student before and after learning from the teacher, respectively; $rand(0, 1)$ is a uniform random number between 0 and 1; x^{best} indicates the current optimal individual, that is, the teacher; \bar{x} represents the average of all individuals; T_F is a teaching factor, and $T_F = round(1 + rand(0, 1))$.

2.2 Learner Phase

After the teacher completes its teaching, the students can improve their grades by learning from each other. A learner interacts randomly with other learners through communication, question and discussion, etc. The teaching process can be described as follows.

$$x_i^{new} = \begin{cases} x_i^{old} + rand(0, 1)(x_i - x_p), f(x_i) < f(x_p) \\ x_i^{old} + rand(0, 1)(x_p - x_i), \text{otherwise} \end{cases} \quad (2)$$

where x_p is an individual randomly selected from the population, and $p \neq i$.

2.3 The Implementation of TLBO

According to the teacher phase and the learner phase described above, after setting the termination condition, the implementation of TLBO algorithm can be described as follows.

Step 1: Define the optimization problem $f(x)$ and set algorithm parameters: $k = 1$ (Iteration number), N_P (Size of population), D (Individual dimension), (L_{j_min}, L_{j_max}) (Value range for each dimension variable, $j = 1, 2, \dots, D$), G (Maximum number of iteration) and σ (Termination condition).

Step 2: Randomly generate initial population $X = \{x_1, x_2, \dots, x_{N_P}\}$, where $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,D})$, where

$$\begin{aligned} x_{i,j} &= L_{j_min} + rand(0, 1)(L_{j_max} - L_{j_min}) \\ i &= 1, 2, \dots, N_P, j = 1, 2, \dots, D \end{aligned} \quad (3)$$

And L_{j_min} and L_{j_max} represent the minimum and maximum values of the j -dimensional variable, respectively.

Step 3: Calculate the fitness value $f(X)$. Check if the termination condition ($f(x_i) \leq \sigma$ or the number of iterations k reaches the maximum) is satisfied or not. If yes, the iteration is terminated and the best individual is obtained as the optimal solution output; Otherwise, go to the next step.

Step 4: Teacher phase: update the individual according to Eq. (1), compare the fitness value of the new individual with the old individual, and select the better individual as the optimal solution.

Step 5: Learner phase: update the individual according to Eq. (2), compare the fitness value of the new individual with the old individual, and select the better individual as the optimal solution.

Step 6: $k = k + 1$, go to step 4.

3 A Positioning Method Based on TLBO

3.1 Problem Description

Although the multipath error is caused by the reflection of the surrounding obstacles, the addition path error is introduced to the pseudorange. Therefore, after the remaining errors except for the multipath error are corrected, the pseudorange ρ_i can be expressed as:

$$\rho_i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} + b_u + \varepsilon_i + v_i, i = 1, \dots, m \quad (4)$$

where m is the number of visible satellites, ρ_i is the pseudorange of satellite i , define $x = (x_u, y_u, z_u)$ is the user position. Denote $a_i = (x_i, y_i, z_i)$ as the position of satellite i , b_u is the clock offset, ε_i is the measurement error, v_i is the addition path error caused by multipath interference to the pseudorange.

Obviously, model (3) can be transformed into a minimization problem:

$$\min_{x, b_u, v_i} \left\{ \sum_{i=1}^m \left(\rho_i - \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} - b_u - v_i \right)^2, v_i \in \varphi_i \right\} \quad (5)$$

where φ_i is the interval of the value of v_i .

The (5) can be regarded as an objective optimization function, where the parameters that need to be optimized are x, b_u, v_i . According to the method in [11], the range of v_i can be obtained by measuring the SNR of satellite i . However, the intervals of x and b_u are not known, thus the approximated position of user needs to be first obtained without considering the multipath interference.

Denote

$$d_i(x) = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2}, i = 1, \dots, m \quad (6)$$

The approximated position of the user is obtained in an iterative way. After linearizing (4), we can express it as

$$\delta \rho_i = \frac{(x_i - x_u)\delta x_u + (y_i - y_u)\delta y_u + (z_i - z_u)\delta z_u}{\rho_i - b_u} + \delta b_u \quad (7)$$

At the k th iteration, yields

$$\delta \rho_i^{(k)} = \rho_i - d_i(x^{(k-1)}) \quad (8)$$

$$\begin{cases} \delta x_u = x_u^{(k)} - x_u^{(k-1)} \\ \delta y_u = y_u^{(k)} - y_u^{(k-1)} \\ \delta z_u = z_u^{(k)} - z_u^{(k-1)} \\ \delta b_u = b_u^{(k)} - b_u^{(k-1)} \end{cases} \quad (9)$$

Express (7) in the form of a matrix as follows.

$$\begin{pmatrix} \delta \rho_1^{(k)} \\ \delta \rho_2^{(k)} \\ \vdots \\ \delta \rho_m^{(k)} \end{pmatrix} = \begin{pmatrix} \alpha_{11}^{(k)} & \alpha_{12}^{(k)} & \alpha_{13}^{(k)} & 1 \\ \alpha_{21}^{(k)} & \alpha_{22}^{(k)} & \alpha_{23}^{(k)} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1}^{(k)} & \alpha_{m2}^{(k)} & \alpha_{m3}^{(k)} & 1 \end{pmatrix} \begin{pmatrix} \delta x_u \\ \delta y_u \\ \delta z_u \\ \delta b_u \end{pmatrix} \quad (10)$$

where

$$\alpha_{i1}^{(k)} = \frac{x_i - x_u^{(k-1)}}{\rho_i - b_u^{(k-1)}}, \alpha_{i2}^{(k)} = \frac{y_i - y_u^{(k-1)}}{\rho_i - b_u^{(k-1)}}, \alpha_{i3}^{(k)} = \frac{z_i - z_u^{(k-1)}}{\rho_i - b_u^{(k-1)}} \quad (11)$$

Then (10) can be expressed as

$$\delta\rho^{(k)} = A_k \begin{pmatrix} \delta x \\ \delta b_u \end{pmatrix} \quad (12)$$

where

$$\delta\rho^{(k)} = [\delta\rho_1^{(k)}, \delta\rho_2^{(k)}, \dots, \delta\rho_m^{(k)}]^T \quad (13)$$

$$\delta x = [\delta x_u, \delta y_u, \delta z_u]^T \quad (14)$$

$$A_k = \begin{pmatrix} \alpha_{11}^{(k)} & \alpha_{12}^{(k)} & \alpha_{13}^{(k)} & 1 \\ \alpha_{21}^{(k)} & \alpha_{22}^{(k)} & \alpha_{23}^{(k)} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1}^{(k)} & \alpha_{m2}^{(k)} & \alpha_{m3}^{(k)} & 1 \end{pmatrix} \quad (15)$$

is the measurement matrix.

Then, the (12) can be solved by the least square method

$$\begin{pmatrix} \delta x \\ \delta b_u \end{pmatrix} = [A_k^T A_k]^{-1} A_k^T \delta\rho^{(k)} \quad (16)$$

Then, the estimated value of the user position can be updated by

$$x^{(k)} = x^{(k-1)} + \delta x \quad (17)$$

Since the initial user position is not known, the initial value can be set as $x^{(0)} = 0$. The iteration process is terminated when $\sqrt{\delta x_u^2 + \delta y_u^2 + \delta z_u^2 + \delta b_u^2}$ is less than an arbitrary given threshold or the number of iterations is maximized. According to the steps mentioned above, the approximated user position \hat{x} , the final update value $\delta\hat{x}$ and $\delta\hat{b}_u$ can be updated.

After getting the approximated user position \hat{x} , we began to consider the impact of multipath interference. The (5) can be approximately by

$$\min_{\delta x, \delta b_u, v} \left\| \hat{A} \begin{pmatrix} \delta x \\ \delta b_u \end{pmatrix} - \delta\hat{\rho} + v \right\|^2 \quad (18)$$

where $v = (v_1, \dots, v_m)^T$, \hat{A} and $\delta\hat{\rho}$ can be obtained by using (8), (11) and (15) when the user position $x = \hat{x}$.

Equation (16) can be regarded as an optimization problem. δx , δb_u and v are the parameters that need to be optimized in the optimization problem. Once the optimal solution \hat{x} is found, the final user position can be obtained by

$$\tilde{x} = \hat{x} + \delta\hat{x} \quad (19)$$

3.2 The Positioning Method Based on TLBO

The pseudo code of the positioning method based on TLBO algorithm is given in Table 1.

Table 1. Pseudo code of positioning method based on TLBO

```

//Initial positioning//
Initialization
Set a threshold and maximum iteration number and set  $x^{(0)} = 0$ .
for  $k = 1 : \max k$ 
    Calculate  $A_k$  and  $\delta\rho^{(k)}$  according to (8), (11), (13) and (15);
    Calculate  $\delta x$  according to (12);
     $x^{(k+1)} = x^{(k)} + \delta x$  ;
    if  $\sqrt{\delta x_u^2 + \delta y_u^2 + \delta z_u^2 + \delta b_u^2} < \text{threshold}$ 
        break;
    end if
end for
Get an initial position  $\hat{x}$ .
//Correct the initial position//
Initialization
Formulate the optimization problem (18) and define  $y = (\delta x, \delta b_u, v)$  ;
Set up parameters:  $N_p$ ,  $D$ ,  $(L_{j_{\min}}, L_{j_{\max}})$ ,  $G$  and  $\sigma$  ;
Randomly generate initial population  $Y = \{y_1, y_2, \dots, y_{N_p}\}$  according to (3);
for  $k = 1 : G$ 
    1) Calculate the objective function value for all individuals according to (18);
    2) Output the optimal solution and break if there is an objective function value less than
         $\sigma$  or  $k > N_p$ , otherwise continue;
    3) Perform teacher phase according to (1) and get the new individual  $y_i^{new}$ . If  $y_i^{new}$  is
        better than  $y_i$  according to objective function (18) and then select  $y_i^{new}$  as the optimal
        solution, otherwise reserve  $y_i$  ;
    4) Perform learner phase according to (2) and get the new individual  $y_i^{new}$ . If  $y_i^{new}$  is
        better than  $y_i$  according to objective function (18) and then select  $y_i^{new}$  as the optimal
        solution, otherwise reserve  $y_i$  .
end for //Output optimal solution
Update the user position:  $\tilde{x} = \hat{x} + \delta\hat{x}$ .

```

4 Experiments and Analysis

In this section, Actual data is collected by an IF signal sampler to verify the performance of the proposed positioning method. In this paper, static single point positioning is adopted. The environment of data collection is located at Mingxiang Campus of Taiyuan University of Technology. The data is about 100-s and the positioning interval is set as 1 s.

In order to verify the performance of proposed algorithm, the differential evolution (DE) algorithm and particle swarm optimization (PSO) algorithm are compared with it in this paper. For the three algorithms, the values of N_p , D , G and σ are set as 80, 8, 500 and e^{-7} , respectively. Other parameters of DE and PSO algorithm are given in Table 2. For the TLBO algorithm, according to a large number of simulation experiments, the ranges of δx and δb_u are recommended as $[\delta \hat{x} - 3, \delta \hat{x} + 3]$ and $[\delta \hat{b}_u - 10, \delta \hat{b}_u + 10]$, respectively.

Table 2. Other parameters of DE and PSO algorithm

| Parameter | F | CR | c_1 | c_2 | v_{\min} | v_{\max} |
|-----------|-----|------|-------|-------|------------|------------|
| DE | 0.8 | 0.1 | \ | \ | \ | \ |
| PSO | \ | \ | 1.5 | 1.5 | -0.5 | 0.5 |

It is known that there is a strong correlation between multipath error and SNR from the Sect 1. Referring to [11], we have the following setting. If the SNR of satellite i is bigger than 42 dB, $v_i = 0$; when it ranges from 40 to 42 dB, $\varphi_i = [0, 5]$; when it ranges from 38 to 40 dB, $\varphi_i = [0, 20]$; when it ranges from 36 to 38 dB, $\varphi_i = [0, 40]$; for other cases, $\varphi_i = [0, 100]$.

Figure 1 shows the deviation of the positioning results from the average of the positioning results in the east, north and up directions by TLBO algorithm, DE algorithm and PSO algorithm. From the comparison of the positioning results, we can see that the positioning accuracy of the TLBO algorithm is higher than the other two algorithms.

Table 3 shows the standard deviation (STD) of the three algorithms in the east, north, and up directions. For a set of observations $\{x_i, i = 1, \dots, N\}$, the STD is defined as

$$STD = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}} \quad (20)$$

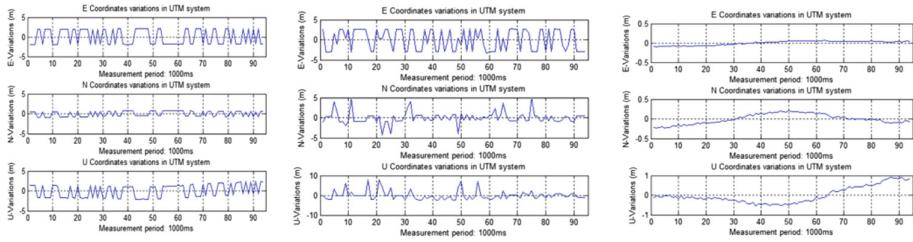


Fig. 1. The positioning results of the three algorithms. Left: DE. Middle: PSO. Right: TLBO

Here \bar{x} is the mean of $\{x_i, i = 1, \dots, N\}$. From Table 3, we can see that the STD of the TLBO algorithm in three directions is smaller than the other two algorithms, which means that the positioning result of the TLBO algorithm is more stable.

Table 3. The STD of the three algorithms

| Method | E(m) | N(m) | U(m) |
|--------|-------|-------|-------|
| DE | 1.976 | 0.683 | 1.639 |
| PSO | 2.558 | 1.484 | 2.252 |
| TLBO | 0.053 | 0.126 | 0.423 |

For a more comprehensive comparison, the experiment with two other different data samples is also conducted and the STD of the positioning results of the three algorithms is shown in Table 4. From the comparison of the above experimental results, we can see that the performance of the TLBO algorithm is better than the other two algorithms regarding positioning accuracy.

Finally, we use the Google Earth to match the positioning results of the TLBO algorithm. The matching results are shown in Fig. 2, indicating the matching results is about 7 meters deviation from the real position in the horizontal direction.

Table 4. Comparison of positioning results for different samples

| Sample | Method | E(m) | N(m) | U(m) |
|--------|--------|-------|-------|-------|
| 1 | DE | 1.993 | 0.664 | 1.615 |
| | PSO | 2.505 | 1.551 | 1.887 |
| | TLBO | 0.009 | 0.038 | 0.278 |
| 2 | DE | 2.050 | 0.837 | 0.738 |
| | PSO | 3.127 | 2.319 | 2.089 |
| | TLBO | 0.512 | 0.518 | 0.745 |

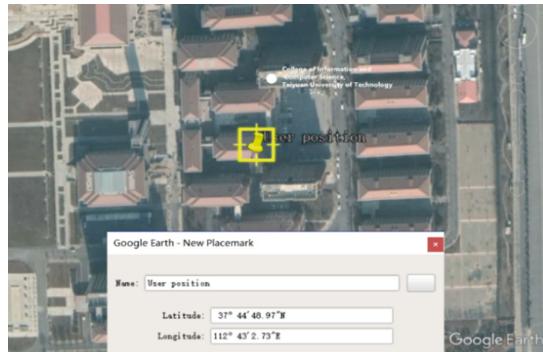


Fig. 2. The match results

5 Conclusion

In this paper, multipath suppression is performed based on the pseudorange model. The contributions of this paper are two folds: (1) the pseudorange model is transformed into an optimization problem; (2) the TLBO algorithm is used to solve the optimization problem, and its performance is compared with DE algorithm and PSO algorithm. The experimental results show that the TLBO algorithm has better positioning results in terms of accuracy and stability.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (No. 61603267, No. 61503271).

References

1. Ba B, Cui W, Wang D et al (2017) Maximum likelihood time delay estimation based on Monte Carlo importance sampling in multipath environment. *Int J Antennas Propag* 2017:1–7
2. Cheng L, Xing YJ, Ren MF et al (2018) Multipath estimation algorithm using ϵ constrained rank-based differential evolution. *Acta Electronica Sinica* 46(1):167–174
3. Xie L, Cui X, Zhao S et al (2017) Mitigating multipath bias using a dual-polarization antenna: theoretical performance, algorithm design, and simulation. *Sensors* 17(2):359
4. Yang H, Yang X, Sun B et al (2016) Global navigation satellite system multipath mitigation using a wave-absorbing shield. *Sensors* 16(8):1332
5. Ho MT, Krichene HA, Ricciardi GF et al (2017) Multipath mitigation in calibration range estimation algorithm. In: Radar Conference. IEEE, pp 1256–1261
6. Jia Q, Wu R, Wang W et al (2017) Multipath interference mitigation in GNSS via WRELAX. *GPS Solutions* 21(2):487–498
7. Breivik K, Forssell B, Kee C et al (1997) Estimation of Multipath error in GPS pseudorange measurements. *Navigation* 44(1):43–52
8. Angrisano A, Maratea A, Gaglione S (2018) A resampling strategy based on bootstrap to reduce the effect of large blunders in GPS absolute positioning. *J Geodesy* 92:81–92
9. Lesoult J, Robert T, Sahmoudi M et al (2018) Multipath mitigation for GNSS positioning in an urban environment using sparse estimation. *IEEE Trans Intell Transp Syst* 99:1–13

10. Chen GY, Gan M, Chen CLP et al (2018) A two-stage estimation algorithm based on variable projection method for GPS positioning. *IEEE Trans Instrum Meas* 99:1–8
11. Cheng L, Guo YF, Xie G (2016) Two improved signal synchronizing methods for GPS software receiver. In: International Symposium on Computational Intelligence and Industrial Applications
12. Rao RV, Savsani VJ, Vakharia DP (2011) Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems. *Comput Aided Des* 43(3):303–315



Opinion Leader Mining Algorithm Based on Double-Graph Model

Xunpu Yuan, Junping Du^(✉), Wanqiu Cui, Zhe Xue, 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. The opinion leader mining is critical to the orderly and healthy development of social networks. The traditional method mostly mines opinion leaders from the perspective of users, and ignores the relationship between Weibo content, so the performance of opinion leader mining algorithm is limited. To solve this problem, this paper proposes an opinion leader mining algorithm based on the designed double-graph model, namely DGRank. Based on the improvement of the LeaderRank algorithm, we combine the relationship between users in social network, the number of fans, the number of Weibo and the number of followers to build a user graph model. Then, we calculate the similarity between the Weibo published by the users and build a Weibo graph model. The correlation between Weibo supplements the relationship between users, and the two models are merged into a unified double-graph model. Finally, we calculate the user impact score to find the opinion leader. The experimental results show that the proposed DGRank algorithm is superior to the traditional methods, and can more effectively dig out opinion leaders.

Keywords: Social network · Weibo · Opinion leader · Double-graph model

1 Introduction

Social networks are becoming increasingly popular and accumulating a large amount of user-generated data. Due to the diversity of social tools, users can participate in the discussion of topics. Among these people, there will be such a group of people who will actively participate in the debate on the topic and have an energetic atmosphere, voice, and influence. These people are what we call opinion leader. And the content they send will generally occupy a prominent position under the whole topic. Therefore, for opinion leader mining, we should not only consider the relationship between users, but also the importance of the Weibo they send.

If opinion leaders spread or publish some illusory and false information on social networking platforms with special political purposes, they may induce users who do not know the truth to forward and spread. Such a situation will seriously endanger national security. Therefore, it is of great research significance to mine opinion leaders in the online social network and conduct real-time monitoring and tracking of opinion leaders to safeguard national security.

By investigating relevant literature, we found that the existing methods are mostly limited to users, only considering the relationship between users, and neglecting the correlation between Weibo content. Therefore, there remains room for improvements in the task of opinion leaders mining.

We propose the double-graph model, namely DGRank, which comprehensively considers the relationship between users, the relationship between Weibo and the relationship between users and Weibo. Experiments show that the proposed DGRank algorithm obtains better results in opinion leader mining compared to the existing works. The main contributions are as follows.

- (1) We optimize the calculation of user attributes by selecting the number of fans, Weibo and followers. The value of the user attribute is very different. For example, some users have only a few dozens of fan bases, and some users can reach tens of millions of fans, which is too large. To address this problem, we reduce the gap between them and assign corresponding coefficients, which leads to a better result.
- (2) We construct a unified double-graph model to mine the opinion leader. It considers not only the relationship between users, but also the relationship between Weibo sent by users, and uses the correlation between Weibo to supplement the relationship between users. Our method improves the effectiveness and accuracy of opinion leader identification.

2 Related Work

Online social network opinion leader mining methods have drawn many research interests in recent years. For example, Wang [1] combines user interaction information with topic similarity to build links between users. Liu [2] proposed a public emergency information dissemination model to analyze the opinion leaders. Chen [3] used community detection methods to solve the problem of overlapping and narrowing the size of candidate generation, and proposed a new cluster-based leadership analysis to find the opinion leaders in the social network. Cao [4] considers the user structure, behavior and emotions and so on, and combines them to propose a multi-feature opinion leader mining algorithm. In addition, the symbolic network is used as a research tool, and the opinion leaders are encouraged by giving corresponding representations to support or oppose the relationship symbols between users and combining traditional mining algorithms [5]. Li [6] analyzes the real information dissemination and combines the microblogging theme to realize the opinion leader mining.

At present, in the ranking of network node influence, the common measurement methods are degree centrality [7], median centrality, tightness centrality, local centrality, K-Shell [8], LeaderRank [9] and so on. As an improved algorithm of PageRank [10], LeaderRank controls the node jump information by adding the ground node in the network, which effectively improves the sorting accuracy and achieves fast convergence.

However, these methods generally only consider the user attributes and the relationship between users, and ignore the Weibo content, resulting in not very good

results. So we propose the double-graph model, namely DGRank. Experiments show that the proposed DGRank algorithm obtains better results in opinion leader mining compared to the existing works.

3 Implementation of Opinion Leader Mining Algorithm Based on Double-Graph Model (DGRank)

The DGRank algorithm is based on the improvement of the LeaderRank algorithm. It is mainly divided into three parts: the construction of user graph model, the construction of Weibo graph model, and the double-graph model which fuses the two graphs. In the construction of the user graph model, the user's attributes are first calculated based on the number of fans, Weibo and followers, as the initial weight of the user node. Then, based on the relationship between users, a user graph model is constructed. In the construction of the Weibo graph model, the similarity between the Weibo is first calculated according to the co-occurrence relationship. When the similarity exceeds the specified threshold, the two Weibo are connected to form a Weibo graph model. Finally, based on the user graph model and the Weibo graph model, combined with the relationship between the user and the Weibo, the two models are merged into a unified double-graph model. According to the calculation of the weight propagation algorithm, the impact scores of the user nodes are obtained and sorted. Finally, the opinion leaders are mined according to the rankings. The overall framework of DGRank algorithm is shown in Fig. 1.

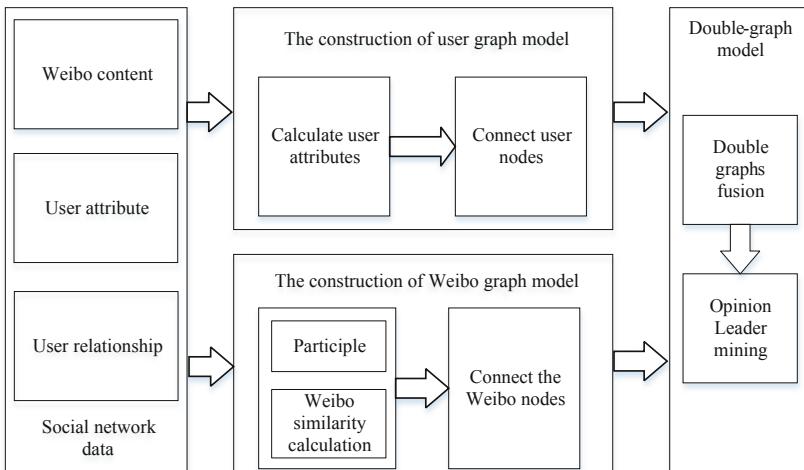


Fig. 1. The overall framework of DGRank algorithm

3.1 The Construction of User Graph Model

The user graph model is constructed based on the LeaderRank model. The user graph model is composed of two elements: node and directed edge. The node refers to the user node. The directed edge indicates the relationship between the users. For example, if user A pays attention to user B, then there is a directed edge from node A to node B. The node is represented by V and the directed edge is represented by E. Finally, we add a global node to connect all user nodes, so that you can get a fully connected graph. This ensures there is only one sort of the last user node. The user attribute selects the number of fans, Weibo and followers. We consider that the value of the user attribute is very different, for example, some users have only a few dozens of fan bases, and some users can reach tens of millions of fans, which is too large. So we reduce the gap between them and assign corresponding coefficients, which yields a better result. The specific calculation formula is as follows:

$$W_i = \alpha_1 \log_{10} N_1 + \alpha_2 \log_{10} N_2 + \alpha_3 \log_{10} N_3 \quad (1)$$

where W_i represents the user attribute. N_1, N_2, N_3 is the number of fans, Weibo and followers. The corresponding weights are 0.5, 0.3, and 0.2. Respectively, based on experience, and considering their importance is gradually reduced. Finally, we take the logarithm to reduce the gap and achieve better experimental results. The specific algorithm flow chart is as follows.

Algorithm 1 User graph model construction algorithm

Input: User list, user relationship, the number of fans, Weibo and followers.

Output: User graph model.

- Step 1: Read the number of fans, Weibo and followers, and then calculate the user weight w_i .
 - Step 2: Read the user list and combine the user weights to build a graph G containing only the user node V.
 - Step 3: Read the user attention relationship. If the user node V_1 pays attention to the user node V_2 , then add V_1 to one side of V_2 .
 - Step 4: Add a global node g, let all user nodes point to node g, and then output the user graph model.
 - Step 5: Return user graph model.
-

3.2 The Construction of Weibo Graph Model

This paper builds the Weibo graph model based on the Weibo similarity. First, each Weibo is segmented, and the stop words are filtered out. Second, we construct a Weibo model G that contains edge E and vertex V, where V is a collection of Weibo. Finally, we construct the edge between two nodes by the similarity between the Weibo. There is an edge between the two nodes only when the similarity between the two nodes is greater than the threshold. The threshold size is set to 0.5 empirically.

In this paper, the similarity calculation between Weibo is a co-occurrence relationship, which represents the corresponding vocabulary appears together in two microblogs. The specific formula is as follows:

$$S(M_i, M_j) = \frac{|w|_{w \in M_i \cap w \in M_j}}{\log(|M_i|) + \log(|M_j|)} \quad (2)$$

where M_i and M_j represent two microblogs, W represents the words, and the numerator part means the number of the same word appearing in two microblogs at the same time. The denominator is the sum of the amount of the word in the sentence. This design can curb the advantage of longer Weibo in similarity calculation and achieve better experimental results.

According to the formula, we can calculate the similarity between Weibo, and then the similarity threshold can be combined to construct the Weibo graph model.

3.3 Double Graphs Fusion and Opinion Leader Mining

Combining the relationship between the user and the Weibo, the user graph model and the Weibo graph model are merged to form a unified double-graph model. If there is an edge between the Weibo sent by the user node V_1 and the Weibo sent by the user node V_2 , there is a relationship between V_1 and V_2 , and it will affect each other. So the edge of the V_1 pointing to V_2 and the edge of the V_2 pointing to V_1 are added in the user graph model. Finally, the two graph models are merged to obtain a unified double-graph model.

The user's initial node weight is the user's property, and then the weight propagation calculation is performed using the following formula:

$$W_i = 1 - \varepsilon + \varepsilon * \sum_{j=0}^n \frac{W_j}{O(V_j)} \quad (3)$$

where W_i represents the weight of the user, ε represents the damping coefficient, which is 0.85, $O(V_j)$ represents the number of outlines of node j ; that is, the weight is divided equally among other user nodes.

Finally, the propagation calculation is performed by Eq. 3, and iterated a certain number of times. When the user weight is stabilized, the final weight of the user node can be obtained. Then sort in descending order, the higher the weight, the higher the ranking, so the top ranking is the opinion leader. The specific opinion leader mining algorithm is as follows:

Algorithm 2 Opinion leader mining algorithm

Input: User graph model, Weibo graph model.

Output: Opinion leader ranking.

Step 1: Read Weibo graph model G_1 , user graph model G_2 .

Step 2: Read each edge E_i of G_1 , and connect the user nodes corresponding to the Weibo nodes in E_i .

Step 3: Add a global node V_g to the fused model, so that all nodes point to V_g , thus constructing a fully connected double-graph model.

Step 4: The weight propagation calculation is performed using formula (3), and the iteration is ended 100 times or when the node weight change is less than 1.

Step 5: The weight of V_g is equally divided among all nodes to get the final node weight.

Step 6: Sort nodes by weight to get opinion leader ranking.

Step 7: Return opinion leader ranking.

4 Experiments

4.1 Dataset

We crawl data from Sina Weibo as our online social network dataset. We collected 5,595 Weibo data on the topic of “haze” including Weibo forwarding number, number of comments, likes and corresponding user id. The data set also contains corresponding 5285 user data, including the relationship between users, the number of fans, Weibo and followers.

4.2 Experimental Evaluation and Analysis

In the experiment, the LeaderRank algorithm [9], PageRank algorithm [10], User-WeightRank algorithm are compared with the proposed DGRank algorithm. The LeaderRank algorithm adopts the iterative idea of PageRank algorithm, and combines the user relationship network and the user’s attributes. Then it adds a global node to construct the user graph model into a fully connected graph. The UserWeightRank algorithm only considers the user’s attributes.

There are two evaluation indicators, namely coverage and manual evaluation. The coverage is based on the user topology map, and the influence is obtained by calculating the number of users covered by the opinion leader. The calculation formula is as follows.

$$C = \frac{1}{N} \sum_{j=0}^P C_j \quad (4)$$

where C stands for coverage, N stands for the total number of user nodes, P stands for the set of opinion leader nodes, and C_j stands for the number of users covered by each opinion leader node.

Manual evaluation refers to the evaluation method of comparing the ranking of opinion leaders calculated by experiments with the ranking of opinion leaders obtained by manual analysis. The main indicators are accuracy and average accuracy MAP values.

The coverage evaluation results of different algorithms are shown in Fig. 2, the DGRank algorithm achieves better results, and the average performance is increased by more than 60%. The DGRank algorithm covers more users in total, and the top 5 opinion leaders in the top rank cover half of the users under the topic, with a total user of 5285 users. When more and more opinion leaders are selected, the user growth rate of the covered users becomes slower. It is in line with the lower ranked users, and the scope of their influence will be smaller and smaller. Therefore, increasing the coverage of the opinion leaders will become less obvious.

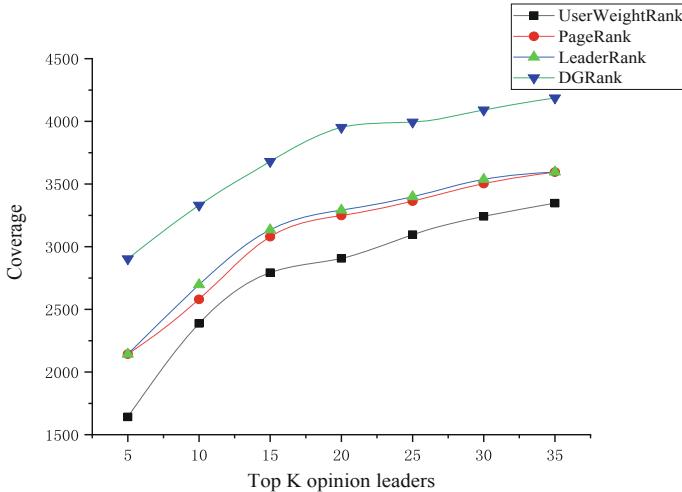


Fig. 2. Comparison of four algorithms on the coverage indicators

In the manual evaluation, the DGRank algorithm is also superior to the other three algorithms. As shown in Fig. 3, the accuracy rate is about 70%, which is increased by about 10%. As shown in Table 1, the MAP value is also increased by about 10%. Therefore, it can be seen that the DGRank algorithm is not only superior to the traditional method in terms of coverage, but also achieves better accuracy.

DGRank algorithm achieves better performances due to the following reasons. First, DGRank algorithm optimizes the user attribute calculation, and assigns the initial

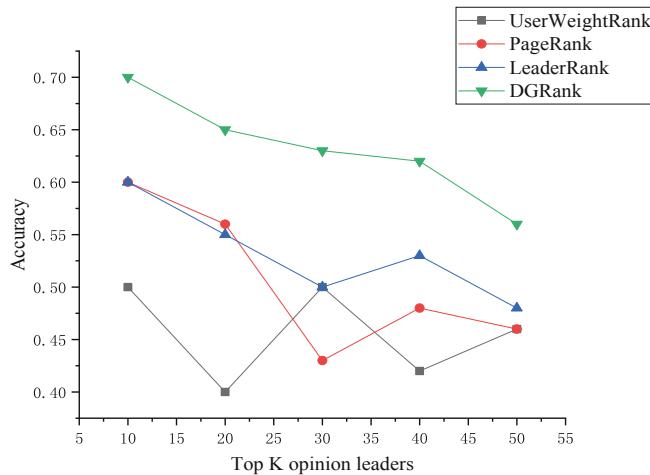


Fig. 3. Comparison of four algorithms on the accuracy indicators

Table 1. Comparison of four algorithms on the MAP evaluation indicators.

| Algorithms | @10 | @20 | @30 | @40 | @50 |
|----------------|------|------|------|------|------|
| UserWeightRank | 0.50 | 0.40 | 0.49 | 0.42 | 0.35 |
| PageRank | 0.60 | 0.55 | 0.43 | 0.47 | 0.46 |
| LeaderRank | 0.60 | 0.55 | 0.49 | 0.52 | 0.47 |
| DGRank | 0.69 | 0.64 | 0.63 | 0.62 | 0.55 |

value to the user node. Second, it absorbs the LeaderRank point, increases the global user node, and transfers the user graph model into a fully connected graph. Finally, the DGRank algorithm not only considers the simple relationship between users but also examines the relationship between the Weibo sent by users.

For the DGRank algorithm, we select the previous opinion leaders to display the case. As shown in Table 2, we can see that the overall performance is great. The number of fans, Weibo and followers are relatively large.

Table 2. Top 5 opinion leaders generated by DGRank algorithm

| Rank | Id | Nickname | Number of fans | Number of followers | Number of Weibo |
|------|------------|----------------|----------------|---------------------|-----------------|
| 1 | 2656274875 | CCTV News | 18249317 | 172 | 31901 |
| 2 | 2803301701 | People's daily | 19977043 | 272 | 28677 |
| 3 | 1618051664 | Headline News | 33787450 | 253 | 88798 |
| 4 | 1653689003 | New weekly | 9595566 | 446 | 18438 |
| 5 | 1644114654 | Beijing News | 5698798 | 894 | 34670 |

5 Conclusions

This paper explores the opinion leaders of online social networks based on the user attributes, the relationship between users, and increasing the relationship between Weibo. The proposed method can more accurately dig out the opinion leaders in the online social network. Through experimental comparison and analysis, the proposed DGRank algorithm can effectively mine the hidden opinion leaders under the topic.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant (No. 61772083, No. 61532006, No. 61877006), and in part by Science and Technology Major Project of Guangxi (GuikeAA18118054).

References

1. Wang C, Du YJ, Tang MW (2017) Opinion leader mining algorithm in microblog platform based on topic similarity. In: IEEE international conference on computer & communications. IEEE
2. Liu X, Liu C (2018) Information diffusion and opinion leader mathematical modeling based on microblog. *IEEE Access* 6:34736–34745
3. Chen Y-C (2019) A novel algorithm for mining opinion leaders in social networks. *World Wide Web* 22(3):1279–1295
4. Cao JX, Chen GJ, Wu JL et al (2016) Multi-feature based opinion leader mining in social networks. *Acta Electronica Sinica* 44:898–905
5. Cao L, Zheng M, Zhang Y et al (2018) An algorithm of opinion leaders mining based on signed network. In: American Institute of Physics Conference Series
6. Li Y, Tian Y, Jin L et al (2017) Identifying opinion leaders in social networks with topic limitation. *Cluster Comput* 20(3):1–11
7. Khan NS, Ata M, Rajput Q (2016) Identification of opinion leaders in social network. In: International conference on information & communication technologies
8. Wu H, Du JT, Sun H et al (2016) Identifying core users in social resource recommendation system with k-shell collapse sequences. *New Technol Libr Inf Serv* 32(9):58–64
9. Le VV, Nguyen HT, Snásel V et al (2016) Identify influential spreaders in online social networks based on social meta path and PageRank. In: International conference on computational social networks
10. Chen G, Fu K, Loza A et al (2017) PageRank tracker: from ranking to tracking. *IEEE Trans Cybern* 44(6):882–893



Model Reference Adaptive Control Based on Transformed Scalar Error for the Three-Tank Level Control System

Pei Zhou and Xuehui Gao^(✉)

Department of Mechanical and Electrical Engineering,
Shandong University of Science and Technology, Tai'an 271019, China
xhgao@163.com

Abstract. The model of water tank level control system with hysteresis nonlinearity is investigated in this paper. In order to improve the control precision of the tank level control system, the model is researched using mechanism method and especially considering the effect of the hysteresis nonlinearity which can be induced by valves. The Prandtl-Ishlinskii (PI) model is adopted to described the hysteresis nonlinearity and the error conversion algorithm is applied to reduce the complexity of the operation. Then, an adaptive controller is designed and the adaptive control law is given. Finally, the appropriate Lyapunov function ensures the error convergence between the system and the reference model and the effectiveness of the proposed approaches is verified by the simulations based on a third tank level model.

Keywords: Hysteresis nonlinearity · Prandtl-Ishlinskii model · Adaptive control · Error transform

1 Introduction

Adaptive control is a control method based on a mathematical model that automatically maintains the desired state of a system for unpredictable changes. By constantly measuring the input, state, output, or performance parameter of the system, adaptive control gradually understands and grasps controlled objects, and then, according to the process of the acquired information and the certain design method, makes control decisions to modify the structure parameters of the controller or control function, in order to achieve the optimal or approximate better control effect in some sense.

The model reference adaptive control system is based on the ideal system reference model. Its reference model is an ideal system with fixed structure and constant parameters. Under the action of the system reference input, the output (state) of the model is defined as the ideal output (state) of the controlled object of the system. Then the error $e(t)$ is obtained by comparing the output or state of the controlled object and the reference model. The adaptive link in the model reference adaptive control is usually hysteresis nonlinear, which sends adjustment signals to the adaptive controller according to the pre-designed adaptive law, which makes the error $e(t)$ constantly decrease or even approach to zero, so that the output or state of the controlled object is

close to the ideal model. The model reference adaptive control was not widely used due to technical difficulties in the early stages. With the development of microcomputer technology, the implementation of this system has been relatively easy. The technology of model reference adaptive control has been widely used in many aspects, such as aircraft autopilot, ship autopilot system, photoelectric tracking telescope follow-up system, SCR speed control system and manipulator control system.

In the water tank system, there are common links in the industrial field such as valve regulation, pipeline connection, sensor and actuator application, etc. It has the characteristics of time-delay, time-variability and non-linearity and can simulate a variety of typical nonlinear time-varying and multi-variable coupling systems in the industrial field. For example, industrial boilers, power station boilers that mainly used as power or heating equipment in thermal power plants and industrial enterprises, Storage tanks, the important equipment for crude oil production and transportation, fine feed tower and other industrial equipment can be simulated by the water tank system. Therefore, the water tank system has a very high engineering application value. With the continuous improvement of the degree of industrialization, the requirement of control accuracy is becoming higher and higher, and the research on hysteresis non-linearity is becoming more and more important. Therefore, it is of great practical significance to study the level control method of the water tank system.

Since the three-tank level control system can be used to simulate many container flow systems in the industrial process, and has a broad application background in practice, so it is of great research value. At present, many domestic and foreign companies and experts in the field of control have put the experimental target on the research and development of the water tank level control system. In the early, Join C proposed a systematic method to guarantee the synthesis of asymptotic nonlinear filters [1]. The effectiveness of the method was verified by a three-tank level system. Pham D T proposed a method to decompose the fuzzy relationship between two decision variables into a single decision variable [2], it could be seen that the three-tank system could not only truly verified the control effect of the new algorithm, but also promoted the development of the control field to a certain extent.

In recent years, many domestic companies also began to research and explore the three-tank level system. For example, the university of Hong Kong and the solid High-tech company of Beijing jointly conducted research and development tests on the three-tank, and the test results showed that the three-tank system could simulate the real working condition from the aspects of liquid level and flow, which was more practical in verifying the control effect of the new algorithm than the simple simulation [3]. The cs-4000 four-capacity water tank designed and manufactured by Zhejiang central control company, which has the functions of real-time monitoring of liquid level, flow, temperature and other functions, involving communication, control, computer and other disciplines [4]. In production practice, PID method has also been widely used. Li designed an adaptive fuzzy PID controller to realize the level control of the double tank and had better robustness compared with the conventional PID controller [6]. Zhao divided the tank level control according to the flow characteristics of the resistance plate and the working condition, and then used the t-s fuzzy PID control strategy to control the tank level [5]. T-s fuzzy PID needs to be weighted for each group of parameters, and the test results show that it reduced the switching disturbance and

improve the control performance of the controller. Therefore, new control strategy need to be investigate for the three-tank systems.

The hysteresis phenomenon of a system refers to a certain state of the system (mainly physical system), which is not only related to the current system input, but also has different results due to the different path of the input process. In other words, a system operating through an input path cannot return to its original state even if the same input value is changed back to the original state. Common physical hysteresis includes: magnetic hysteresis, electrical hysteresis, elastic hysteresis, liquid-solid phase change hysteresis, contact Angle hysteresis, etc.

With the continuous improvement of the industrialization degree, the requirements of control accuracy are higher and higher, the influence of hysteresis nonlinearity becomes more and more difficult to avoid. Therefore, the research on the valve hysteresis nonlinearity in the three-tank level control system also become more and more important. Hysteresis models can be roughly divided into two categories: phenomenological models and physical models. PI model is one of the most important phenomenological models to describe hysteresis, and many scholars have conducted in-depth research on this issue. Liu et al. [7] proposed a robust adaptive control strategy for the inverse PI hysteresis model and gave strict proof; Mohammad [8] and Pavel studied the inverse compensation control of the frequency-independent PI model of the piezoelectric ceramic actuator, and proposed a feed forward compensation control strategy using the frequency-independent inverse PI hysteresis model to control the piezoelectric ceramic actuator; Tan et al. [9] proposed an extended PI operator to describe the hysteresis of piezoelectric actuators to overcome the hysteresis control problem under ill condition.

The hysteresis nonlinearity of valve is not considered in the above tank system studies. Different from the above studies, in this paper, considering the nonlinear hysteresis of valve, using the extended PI model, which is the most important phenomenological model, to describe the hysteresis. The “mechanism method” and “two-point method” will be adopted to model the three-tank level control system, and the model reference adaptive control will be adopted. In addition, the algorithm of error conversion is proposed to reduce the complexity of operation and an appropriate Lyapunov function is selected so as to ensure the effectiveness of the proposed control strategy.

2 System Description

The tank level control system is mainly composed of the tank body, power supply system, level sensor and actuator. The following model is established according to the system:

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bv(t) \\ y(t) &= Cx(t) \\ v(t) &= f(u(t)),\end{aligned}\quad (1)$$

Where, x are the state space state variables, y is the system output, u is the system input, and v represents the nonlinear hysteresis link output (valve).

A, B, C are the state space coefficient matrices of the linear element respectively, and $v(t) = f(u(t))$ means the hysteretic nonlinear element. In this paper, the extended PI model is used to express the hysteresis nonlinearity $v(t)$. The PI model can be defined as follows [3]:

$$v_m[u](t) = p_{m_0} m_0 u + \int_0^R p_m(r) G_{mr}[u] dr \quad (2)$$

Where, p_{m_0} is a normal number, $p_m(r)$ is a integrable density function, and $\int_0^R r p_m(r) dr < \infty$ when $p_m(r) > 0$ is satisfied.

The goal of the controller is to design a corresponding controller $u(t)$ so that the signal $x(t)$ in formula (1) can track the reference signal $x_m(t)$ and the tracking error converges. The reference signal $x_m(t)$ is defined as follows:

$$\dot{x}_m(t) = A_m x_m(t) + B_m m(t), \quad (3)$$

Where $m(t)$ is the expected input, $A_m \in R^{n \times n}$ is asymptotically stable and satisfies:

$$\det(sI - A_m) = R_m(s) = (s + k)R(s)$$

Where $k > 0$, $R(s)$ is Hurwitz.

3 Controller Design

3.1 Error Conversion

According to literature [10, 11], in order to design an adaptive controller, the following reasonable assumptions are made:

Assumption 1: $A_m \in R^{n \times n}$ is the unknown matrix, $B \in R^{n \times n}$ is known and there is an unknown vector $\theta \in R^{n \times 1}$ to make the following formula true:

$$A + B\theta^T = A_m, \quad (4)$$

where (A, B) is controllable.

The error between the reference model and the system can be defined as:

$$e(t) = x(t) - x_m(t) \quad (5)$$

According to Eqs. (1), (3) and (4):

$$\dot{e}(t) = A_m e(t) + B(v(t) - m(t) - \theta^T x(t)) \quad (6)$$

In order to simplify the design of the controller, Eq. (5) is simplified from the vector form to the scalar form. According to literature [11], the following lemma is established:

Lemma 1: If the asymptotically stable system is controllable, its characteristic polynomial is as follows:

$$\alpha(s) = \det(sI - A) = (s + k)R(s) \quad (7)$$

where $k > 0$, $R(s)$ is Hurwitz.

There exists h making the following equation to be true:

$$h^T(sI - A)^{-1}B = \frac{1}{s + k} \quad (8)$$

Lemma 1 provides a method to convert vector form into scalar form. According to the method provided by lemma, we can easily convert vector error (5) into scalar error.

According to the Eq. (8), there is h , which makes the following formula holding

$$h^T(sI - A_m)^{-1}B = \frac{1}{s + k} \quad (9)$$

Then, vector error can be converted into scalar error:

$$e_s = h^T e \quad (10)$$

Equation (6) is written by Laplace transform:

$$e_s(s) = \frac{1}{s + k} [v(s) - m(s) - \theta^T x(s)] \quad (11)$$

Equation (11) is written by Laplace inverse transform:

$$\dot{e}_s = -ke_s + [v(t) - m(t) - \theta^T x(t)] \quad (12)$$

3.2 Controller Design

For the tracked system, the controller design is as follows

$$u(t) = \frac{1}{p_{m_0}} \left\{ m(t) + \hat{\theta}^T x(t) - \int_0^R \hat{P}(t, r) G_{mr}[u] dr \right\} \quad (13)$$

where $\hat{\theta}$ and $\hat{P}(t, r)$ are estimated values of θ , $p(t, r)$, respectively, $p(t, r) = P_m(r)$, $\frac{1}{p_{m_0}} m(t)$ can be regarded as the special expected input, $\frac{1}{p_{m_0}} \theta^T x(t)$ can be regarded as the feedforward link, and $\frac{1}{p_{m_0}} \int_0^R \hat{p}(t, r) G_{mr}[u] dr$ is as the hysteresis compensation link.

To facilitate the design of the controller, define the following variables

$$\begin{cases} \tilde{\theta}^T = \theta^T - \hat{\theta}^T \\ \tilde{p}(t, r) = p(t, r) - \hat{p}(t, r) \end{cases} \quad (14)$$

Parameters θ and $p(t, r)$ adopt the adaptive law as follows

$$\begin{cases} \dot{\theta} = e_s x(t) \\ \frac{\partial}{\partial t} p(t, r) = -e_s G_{mr}[u](t) \end{cases} \quad (15)$$

Theorem 1. The nonlinear element of the system (1) is determined by Eq. (2). If the adaptive control (13) and the adaptive law (15) are adopted, the system error converges.

Proof:

The Lyapunov function is defined as follows

$$V(t) = \frac{1}{2} \left(e_s^2 + \tilde{\theta}^T \tilde{\theta} + \int_0^R \tilde{p}^2(t, r) dr \right) \quad (16)$$

There are:

$$\begin{aligned} \dot{V}(t) &= e_s \cdot \dot{e}_s + \frac{1}{2} \left(\dot{\tilde{\theta}}^T \tilde{\theta} + \tilde{\theta}^T \dot{\tilde{\theta}} \right) + \int_0^R \tilde{p}(t, r) \frac{\partial}{\partial t} \tilde{p}(t, r) dr \\ &= e_s \cdot \dot{e}_s + \tilde{\theta}^T \cdot \dot{\tilde{\theta}} + \int_0^R \tilde{p}(t, r) \frac{\partial}{\partial t} \tilde{p}(t, r) dr \end{aligned} \quad (17)$$

According to Eq. (14)

$$\dot{V}(t) = e_s \cdot \dot{e}_s + \tilde{\theta}^T \cdot \dot{\tilde{\theta}} + \int_0^R \tilde{p}(t, r) \frac{\partial}{\partial t} p(t, r) dr \quad (18)$$

Substitute Eq. (15) of the adaptive law into Eq. (18) to get:

$$\dot{V}(t) = e_s \cdot \dot{e}_s + \tilde{\theta}^T \cdot e_s x(t) + \int_0^R \tilde{p}(t, r) (-e_s) G_{mr}[u](t) dr \quad (19)$$

Substitute Eq. (12) into Eq. (19) to get:

$$\begin{aligned} \dot{V}(t) &= e_s \cdot [-k e_s + v(t) - m(t) - \theta^T x(t)] + e_s \cdot \tilde{\theta}^T \cdot x(t) - e_s \int_0^R \tilde{p}(t, r) \cdot G_{mr}[u](t) dr \\ & \quad (20) \end{aligned}$$

Substituting Eqs. (2) and (13) into Eq. (20), we have:

$$\dot{V}(t) = -ke_s^2 + e_s \left[p_{m_0} \times \frac{1}{p_{m_0}} \left\{ m(t) + \tilde{\theta}^T x(t) - \int_0^R \tilde{p}(t, r) G_{mr}[u] dr \right\} + \int_0^R p(t, r) G_{mr}[u] dr - m(t) - \theta^T x(t) \right] \quad (21)$$

Substitute Eq. (14) into Eq. (21) to get:

$$\begin{aligned} \dot{V}(t) = & -ke_s^2 + e_s \left[-\tilde{\theta}^T \cdot x(t) + \int_0^R \tilde{p}(t, r) \cdot G_{mr}[u](t) dr \right] + e_s \cdot \tilde{\theta}^T \cdot x(t) - e_s \int_0^R \tilde{p}(t, r) \\ & \cdot G_{mr}[u](t) dr \end{aligned} \quad (22)$$

There are:

$$\dot{V}(t) = -ke_s^2 \leq 0$$

Therefore, Lyapunov function is a non-increasing function, and the systematic error converges.

4 Simulation

Consider the following three-tank level control system:

$$\begin{aligned} \dot{x}_p(t) &= A_p x_p(t) + B_p v(t) \\ y_p(t) &= C_p x_p(t) \\ v(t) &= f(u(t)) \end{aligned}$$

$$\text{Where, } Ap = \begin{bmatrix} -4.33 & -6.2 & -3.14 \\ 3 & 0 & 0 \\ 0 & 3.35 & 0 \end{bmatrix} \quad Bp = \begin{bmatrix} 1.56 \\ 0 \\ 0 \end{bmatrix}$$

The PI model in (2) is adopted for the hysteresis link of the valve, and the parameter is: $p_{m_0} = 0.5$, $m_0 = 1.5$, $p_m(r) = e^{-0.00015r^2}$, $r \in [0, 100]$.

The corresponding reference model is:

$$A_m = \begin{bmatrix} -8.75 & -3.15 & -4.21 \\ 5 & 0 & 0 \\ 0 & 3.125 & 0 \end{bmatrix} \quad B_m = \begin{bmatrix} 1.47 \\ 0 \\ 0 \end{bmatrix}$$

The reference input is selected as $m(t) = \sin(t)$, and the system is tracked by the model reference adaptive control. The results are shown in the figure below:

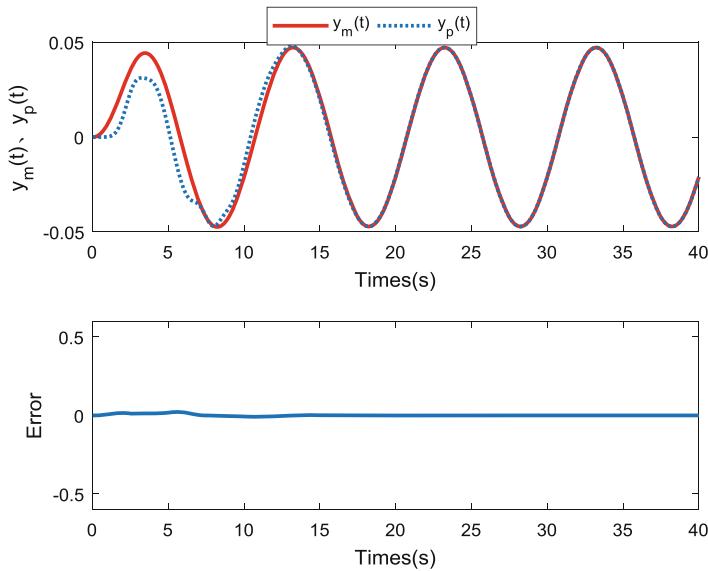


Fig. 1. Tracking trajectory and errors of the proposed approach

Figure 1 illustrates the tracking trajectory and errors with the proposed method and Fig. 2 shows the tracking trajectory and errors with reference adaptive control without

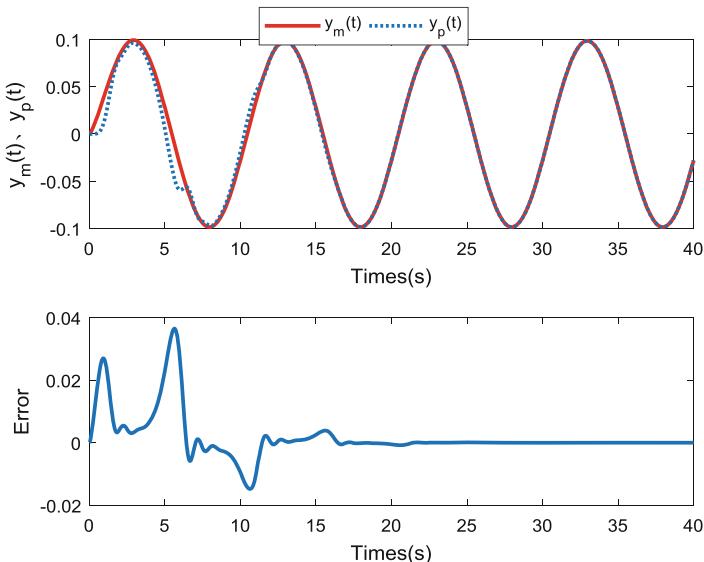


Fig. 2. Tracking trajectory and errors of model reference adaptive control without error conversion

error conversion. It can be seen from the figures that the use of model reference adaptive control enables the control output to track the reference signal well, and the tracking error converges to zero. It shows the superiority of the model reference adaptive control, verified the correctness and effectiveness of the control strategy designed in this paper, and realizes the high-precision control while effectively restraining the hysteresis of the valve. Meanwhile, from the Figures, we can find that the error conversion can simplify the calculation, but magnify the errors, degrade the control precision.

5 Conclusion

Aiming at the hysteresis of valves in the water tank level control system, this paper introduced the extended PI model describing the hysteresis nonlinear link. In order to improve the control precision, the model reference adaptive control method was adopted, and then the vector error was converted into a scalar error for simplifying the calculation, so that the error tends to zero to ensure the stability of the system. Finally, the convergence and stability of the designed controller were guaranteed by Lyapunov function. The validity of the method was verified by the simulations which used a third-order tank level model and the results were demonstrated the validity of the proposed approaches.

Acknowledgement. This work is Supported by Shandong Natural Science Foundation of China (ZR2017MF048), the National Natural Science Foundation of China (61803216), Tai'an Science and Technology development program (2017GX0017).

References

- Join C, Ponsart JC, Sauter D et al (2005) Nonlinear filter design for fault diagnosis: application to the three-tank system. *IEE Proc Control Theory Appl* 152(1):55–64
- Pham DT, Li D (2001) Fuzzy control of a three-tank system. *Proc Inst Mech Eng - Part C* 215(5):505–510
- Liu X, Kong X (2013) Power industry complex system model predictive control – current situation and development. *Chin J Electr Eng* 33(5):79–85 (in Chinese)
- Hou C (2016) Experimental device design and algorithm research of four - volume system. Master's thesis of Jilin University (in Chinese)
- Zhang J, Wang S, Zhao K (2007) Study of control strategies of the three-tank water system based on flow characteristic of piecewise-linear. *Control Eng China* 14(2):17–19 (in Chinese)
- Li Y (2016) Liquid level control of multi-tank tank based on adaptive fuzzy PID control. Master's degree thesis in automatic control, Jiangsu University of Science and Technology, Zhenjiang, pp 1–4 (in Chinese)
- Liu S, Su C, Li Z (2014) Robust adaptive inverse control of a class of nonlinear systems with Prandtl-Ishlinskii hysteresis model. *IEEE Trans Autom Control* 59(8):2170–2175
- Mohammad AJ, Krejci P (2013) Inverse rate-dependent Prandtl-Ishlinskii model for feedforward compensation of hysteresis in a piezo-micro positioning actuator. *IEEE/ASME Trans Mech* 18(5):1498–1507

9. Tan U, Latt W, Shee C, Cameron NR, Ang W (2009) Feedforward controller of ill-conditioned hysteresis using singularity-free Prandtl-Ishlinskii model. *IEEE/ASME Trans Mech* 14(5):598–604
10. Su CY, Wang QQ, Chen XK, Rakheja S (2005) Adaptive variable structure control of a class of nonlinear systems with unknown Prandtl-Ishlinskii hysteresis. *IEEE Trans Autom Control* 50(12):2069–2074
11. Wang XS, Hong H, Su CY (2003) Model reference adaptive control of continuous-time systems with an unknown input dead-zone. *IEE Proc-Control Theory Appl* 150(3):261–266



Research on Self-compensation Technology of Strapdown Inertial Navigation System

Jianqiang Zheng^{1(✉)}, Lin Gao², Qinghua Ma¹, and Shujun Yang¹

¹ Xi'an Modern Control Technology Research Institute, Xi'an, China
371016639@qq.com

² School of Engineering Machinery, Chang'an University, Xi'an, China

Abstract. To solve the problem of constant drift divergence in strapdown inertial navigation system, a self-compensation method using missile rotation modulation is proposed. The constant drift of gyroscope and accelerometer can be modulated into periodic signal by spin modulation of missile body without special rotating mechanism. The influence of constant drift can be effectively eliminated by integral operation. The simulation results show that the method can effectively suppress the error growth, improve the navigation accuracy of SINS.

Keywords: Strapdown inertial navigation system · Constant drift · Missile spin · Self-compensation

1 Introduction

The errors of inertial devices are generally divided into constant deviation and random deviation [1], which will cause attitude error and positioning error and affect the performance of inertial navigation system [2, 3]. For strapdown inertial navigation system, the constant drift of inertial devices seriously affects its navigation accuracy. Error compensation is an effective way to improve the navigation accuracy of strapdown inertial navigation system. In order to improve the accuracy of inertial devices and systems in practical application, The traditional error compensation method is extremely complex. Firstly, it needs to complete the error calibration work, that is, to establish the model equation of inertial instruments and inertial systems, Secondly, it needs to calibrate the various error items of inertial devices and systems with special precision testing equipment, and then to compensate the output of inertial devices in the actual application environment. Through complex and precise work, the constant deviation of inertial devices can be compensated theoretically, but in fact it will change with time. For example, the constant deviation will change with each start-up, that is, starting drift step by step, and other slow changes such as drift day by day.

Rotary modulation technology is an effective technology to compensate the constant drift of inertial devices. It has been developed for many years [4–6]. This technology can automatically adjust the constant drift by fixing strapdown inertial navigation system on the carrier through a special rotating mechanism and rotating in a specific way to offset the influence of constant deviation on system accuracy and realize self-compensation of constant drift [7–9]. This technique is different from the calibration compensation method mentioned above. Rotary modulation technology does not need to know the

estimated value of the inertial device error, but modulates the error into a periodic change form, so that its integral value is zero in a period, so that the inertial device error can be compensated. Based on the analysis of the principle of rotating auto-compensation, a new method of self-compensation is proposed in this paper. Instead of special rotating mechanism, the constant drift of inertial devices is modulated and suppressed directly by the spin of missile body.

2 Principle Analysis of Self-compensation Based on Missile Rotation Modulation

The traditional rotary modulation system is that strapdown inertial navigation system is fixed on the rotating platform, and the rotating platform rotates relative to the carrier, so that the drift changes periodically along the carrier coordinate system. The modulation system based on missile rotation is that strapdown inertial navigation system is directly attached to the missile body. The missile body rotates. Strapdown inertial navigation system rotates with the missile body. The rotating missile body and strapdown inertial navigation system constitute a new rotating coordinate system.

The strapdown navigation algorithm of missile body rotation modulation is basically the same as that of traditional strapdown navigation algorithm of rotation modulation, and a coordinate conversion relation is added. The difference is that the traditional strapdown navigation algorithm of rotation modulation is transferred from rotating coordinate system to missile body coordinate system, and the strapdown navigation algorithm of missile body rotation modulation is transferred from rotating coordinate system to quasi-missile body coordinate system.

The coordinate systems are as follows: n is the navigation coordinate system; I is the inertial coordinate system; e is the earth coordinate system; b is the quasi-projectile coordinate system; p is the rotation coordinate system, that is, the projectile coordinate system. The initial time b system coincides with p system.

The error model of conventional strapdown inertial navigation system is represented by the angular φ error equation.

$$\dot{\phi}^n = -\omega_{in}^n \times \phi^n + \delta\omega_{in}^n - C_b^n \delta\omega_{ib}^b \quad (1)$$

$$\delta\dot{v}^n = f^n \times \phi^n + C_b^n \delta f_{ib}^b - (2\omega_{ie}^n + \omega_{en}^n) \times \delta v^n - (2\delta\omega_{ie}^n + \delta\omega_{en}^n) \times v^n - \delta g^n \quad (2)$$

Among them,

ϕ denotes the misalignment angle of the mathematical platform, v and δv denote the velocity and velocity errors of the mathematical platform, ω and $\delta\omega$ denote the angular velocity and angular velocity errors respectively, f and δf denote the force and error measured by accelerometer, δg denotes the gravity deviation, and C_b^n denotes the transformation matrix from the missile coordinate system to the navigation coordinate system, namely the attitude matrix.

Because the output of inertial forward is measured in the rotating coordinate system, the strapdown inertial navigation system of missile body rotation modulation

should be converted to the navigation coordinate system for navigation calculation. The error equation of strapdown inertial navigation system based on missile rotation modulation is as follows: the rotating coordinate system (i.e. missile coordinate system) rotates relative to the quasi-missile coordinate system, so the error equation of strapdown inertial navigation system based on missile rotation modulation is as follows:

$$\dot{\phi}^n = -\omega_{in}^n \times \phi^n + \delta\omega_{in}^n - C_b^n C_p^b \delta\omega_{ip}^p \quad (3)$$

$$\delta v^n = f^n \times \phi^n + C_b^n C_p^b \delta f_{ip}^p - (2\omega_{ie}^n + \omega_{en}^n) \times \delta v^n - (2\delta\omega_{ie}^n + \delta\omega_{en}^n) \times v^n - \delta g^n \quad (4)$$

Among them,

C_p^b is the transformation matrix from the projectile coordinate system to the quasi-projectile coordinate system, and C_b^n is the transformation matrix from the quasi-projectile coordinate system to the navigation system.

Comparing the (1), (2) and (3), (4) formulas, it can be seen that the input error terms of inertial components of strapdown inertial navigation system with missile body rotation modulation need to be transformed twice.

In the formulas (3) and (4), the errors caused by the inaccuracies of gyroscope and accelerometer measurements are respectively, ω_{ip}^p and δf_{ip}^p . It can be seen that the errors of inertial elements ω_{ip}^p and δf_{ip}^p need to be multiplied by the attitude matrix $C_b^n C_p^b$ before they can be transformed into the quantities in the navigation coordinate system. Order,

$$\varepsilon^n = C_b^n C_p^b \delta\omega_{ip}^p \quad (5)$$

$$\Delta^n = C_b^n C_p^b \delta f_{ip}^p \quad (6)$$

Then, ε^n and Δ^n represent the errors of the angular velocity and force of the mathematical platform caused by the errors of the inertial components measured in the error equation of the inertial navigation system, respectively.

The essence of automatic compensation for missile-body rotating strapdown inertial navigation system is to periodically change the value of attitude matrix $C_b^n C_p^b$, so that the integral or mean value of ε^n and Δ^n , which are the error terms of mathematical platform in error propagation equation in a period is as close as possible to zero, so as to reduce the accumulation of system error and improve navigation accuracy.

3 Establishing Mathematical Model of Inertial Devices Error

In theory, the higher the order of the error model and the more the number of terms is, the higher the accuracy of its description is. However, the complexity and actual influence of the comprehensive model are selected and the following mathematical model of gyroscope and accelerometer errors are established.

The mathematical model of gyroscope error:

$$\begin{cases} V_x = E_{1x}(D_{0x} + D_{1x}A_x + D_{2x}A_y + D_{3x}A_z + \omega_x + E_{yx}\omega_y + E_{zx}\omega_z + T_x\Delta T) \\ V_y = E_{1y}(D_{0y} + D_{1y}A_x + D_{2y}A_y + D_{3y}A_z + E_{xy}\omega_x + \omega_y + E_{zy}\omega_z + T_y\Delta T) \\ V_z = E_{1z}(D_{0z} + D_{1z}A_x + D_{2z}A_y + D_{3z}A_z + E_{xz}\omega_x + E_{yz}\omega_y + \omega_z + T_z\Delta T) \end{cases} \quad (7)$$

Among them,

- T_x, T_y, T_z : the output coefficients of gyroscope in X, Y and Z directions are affected by temperature, respectively, $(^{\circ})/\text{s}/^{\circ}\text{C}$;
- ΔT : the difference between working temperature and standard temperature, $^{\circ}\text{C}$;
- V_x, V_y, V_z : the output voltage of gyroscope in X direction, Y direction and Z direction per unit time, respectively, mV;
- D_{0x}, D_{0y}, D_{0z} : the constant drifts of X, Y and Z channels of gyroscope are measured, respectively, $(^{\circ})/\text{h}$;
- D_{1x}, D_{1y}, D_{1z} : the output coefficients of gyroscope X-direction, Y-direction and Z-direction channels are affected by X-direction apparent acceleration, respectively, $(^{\circ})/\text{h} \cdot \text{g}$;
- D_{2x}, D_{2y}, D_{2z} : the output coefficients of gyroscope X-direction, Y-direction and Z-direction channels are affected by Y-direction apparent acceleration, respectively, $(^{\circ})/\text{h} \cdot \text{g}$;
- D_{3x}, D_{3y}, D_{3z} : the output coefficients of gyroscope X-direction, Y-direction and Z-direction channels are affected by Z-direction apparent acceleration, respectively, $(^{\circ})/\text{h} \cdot \text{g}$;
- E_{1x}, E_{1y}, E_{1z} : the scaling factors of the X, Y and Z channels of the gyroscope, respectively, $\text{mV}/(^{\circ}/\text{s})$;
- E_{yx}, E_{zx} : the installation error coefficients of the gyroscope axis ω_x perpendicular to the Y and Z axes of the body, respectively, 1;
- E_{xy}, E_{zy} : the installation error coefficients of the gyroscope axis ω_y perpendicular to the X and Z axes of the body, respectively, 1;
- E_{xz}, E_{yz} : the installation error coefficients of the gyroscope axis ω_z perpendicular to the X and Y axes of the body, respectively, 1;
- A_x, A_y, A_z : the apparent accelerations in X, Y and Z directions, respectively, g;
- $\omega_x, \omega_y, \omega_z$: the rotational angular rates of X, Y and Z directions, respectively, $^{\circ}/\text{s}$;

The mathematical model of accelerometer error is similar to that of gyroscope error, so it is not described here anymore.

The error model reflects the relationship between the output of inertial devices and various error terms. This paper focuses on the self-compensation effect of rotation modulation on constant drift.

4 Simulation Analysis

In order to verify the effect of missile rotation modulation on the performance of strapdown inertial navigation system, this paper simulates and compares the non-rotation modulation and rotation modulation of missile body, and verifies the feasibility of missile rotation modulation technology.

The simulation conditions: the constant drift of X-direction gyroscope is $1^\circ/\text{min}$, Y-direction gyroscope and Z-direction gyroscope is $10^\circ/\text{h}$, X-direction accelerometer is 5 mg, Y-direction accelerometer and Z-direction accelerometer is 0.5 mg, without considering other error factors; the simulation time is 100 s.

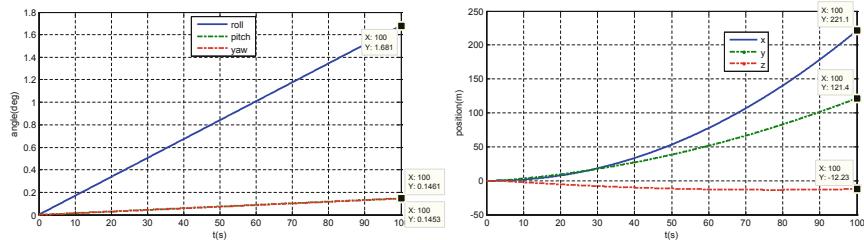


Fig. 1. Attitude error and position error of missile body without rotation and modulation

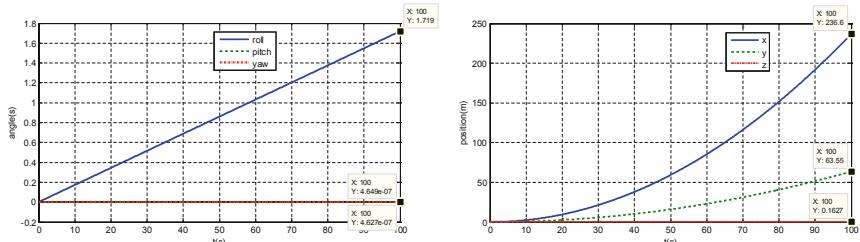


Fig. 2. Attitude error and position error of missile body under rotation modulation

Figure 1 shows the simulation results of attitude error and position error under the condition that the projectile body does not rotates and modulates. Figure 2 show the simulation results of attitude error and position error under the condition of rotating modulation of the projectile body. It can be seen from the simulation results that the error of inertial device is compensated by the automatic drift compensation technology of the projectile body modulation, and the attitude error and position error are obviously improve compared with that before the compensation, and the precision of the strapdown inertial navigation system is greatly improved.

5 Conclusion

Based on the analysis of the traditional rotational modulation method of strapdown inertial navigation system, this paper presents a method of directly utilizing missile rotation to modulate. This method eliminates the complex rotational mechanism, has simple implementation and improves the reliability and practicability of the system. The simulation results show that the scheme can effectively reduce the influence of inertial device constant drift on the inertial navigation system, which greatly improves the attitude and position accuracy of strapdown inertial navigation system. The scheme is a simple, practical and high-precision modulation and compensation method.

References

1. Huang WQ, Zhao GL, Tan ZF et al (2002) Research of Gyro case rotating monitor technique. *Ship Build China* 43(158):54–59 (1000-4882)
2. Titterton DH, Weston JL (2004) Strapdown inertial navigation technology, 2nd edn. The American Institute of Aeronautics and Astronautics and the Institution of Electrical Engineers, pp 71–73, 342–361
3. Huang D, Cheng L (1986) Inertial navigation system. National Defense Industry Press, Beijing, pp 74–92
4. Terry T, Emanuel L (2000) The AN/WSN-7B marine gyrocompass/navigator. In: Proceedings of the 2000 national technical meeting of the institute of navigation, pp 348–357
5. Levinson E, Majure R (1987) Accuracy enhancement techniques applied to the marine ring laser inertial navigation. *J Inst Navig* 34:64–86
6. Lahham JI, Brazell JR (1992) Acoustic noise reduction in the MK 49 ships inertial navigation system. In: IEEE position location and navigation symposium, pp 32–39
7. Zhang Y, Lu Q, Wong H (2009) Analysis and design of marine RLG navigation system base on IMU rotation. *Ocean Technol* 28(2):88–91
8. Levinson E, Willcocks M (1994) The next generation marine inertial navigation is here now. In: IEEE position location and navigation symposium. IEEE, pp 121–127
9. Gao Y, Wang T, Li G (2013) Error auto-compensation of constant drift of FOG. *J Syst Simul* 25(6):1161–1165



Scalable Observer Bandwidth Based Active Disturbance Rejection Control of Brushless DC Motor

Wei Wei, Zhiyuan Zhang, and Min Zuo^(✉)

Beijing Technology and Business University, Beijing 100048, China
zuomin@btbu.edu.cn

Abstract. A brushless DC motor (BLDCM) is a typical nonlinear and strong coupling plant. Unknown, complex, and time-varying external disturbances make the control of BLDCM be a big challenge. PID control, passive receiving tracking errors, cannot satisfy the requirements of high-precision speed control. To address it, a scalable observer bandwidth based active disturbance rejection control (SOADRC) is proposed. By adjusting observer bandwidth, high-precision speed control can be achieved. A numerical model of BLDCM is built to verify the SOADRC. Simulation results show that SOADRC is able to get better speed control accuracy.

Keywords: BLDCM · Speed control · LADRC · Scalable observer bandwidth

1 Introduction

Brushless DC motor (BLDCM) has numerous advantages, such as high efficiency, small size and less rotor loss. It has been widely used in aerospace, robotics and other fields [1]. However, nonlinearity, strong coupling, various unknown complex and time-varying external disturbances deteriorate closed-loop performance of a PID control, such as longer transient time, larger overshoots and unsatisfactory steady-state performance. Therefore, a series of related control strategies have been put forward. For example, fuzzy PID control [2, 3], which combines fuzzy control with traditional PID control, is able to achieve desired performance. However, fuzzy rules depend largely on personal experience, which make the rules be more subjective. A current hysteresis control [4] directly controls the current, which can obtain better torque control performance, but the switching frequency of an inverter is not fixed, which results in a larger pulse in output current. An improved particle swarm optimization algorithm was taken to optimize the fuzzy controller parameters [5], and it achieved desired performance. But the controller design is more complicated. Furthermore, a series of algorithms such as model adaptive control [1], feedback linearization control [6] and model predictive control [7] have been proposed and those approaches have obvious advantages over the traditional PID control. Nevertheless, dependence on the model information make those approaches be sensitive to internal uncertainties and external disturbances. Therefore,

a control technique, which is less dependent on the plant model and more robust, is in great need. Active disturbance rejection control (ADRC), proposed by Han [8–10], does not rely much on the model information. Based on input and output signals of a plant, system output, derivatives of the system output, and the generalized disturbance can be estimated and utilized to construct a control law. It is practical to get high-performance. However, there are lots of tuneable parameters. Gao proposed a linear ADRC (LADRC), and gave out a bandwidth-parameterization approach [11, 12]. LADRC accelerates the applications of active disturbance rejection control.

Although LADRC is effective, its parameters are fixed. It means that the bandwidth of an extended state observer is not adaptive to the time-varying disturbances. Obviously, it is difficult to get better disturbance estimation. Therefore, a scalable observer bandwidth is necessary so as to improve the closed-loop performance. In this paper, on the basis of the estimation error, a scalable observer bandwidth based ADRC (SOADRC) is proposed. It can adjust its observer bandwidth in real time to accommodate to the time-varying disturbances. Better disturbance estimation ability can be guaranteed, and more satisfied closed-loop performance can be desired. Finally, SOADRC is verified on a BLDCM simulation platform built by Matlab/Simulink. Numerical results show that performance of a closed-loop BLDCM system based on the SOADRC is significantly improved compared with both LADRC with fixed parameters and PID.

2 Linear Active Disturbance Rejection Control

A first-order LADRC system is shown in Fig. 1. r is the set value of the system, u_0 is the output of the controller, and G_p is the controlled plant. Linear extended state observer (LESO) is utilized to estimate the generalized disturbance f (including external disturbances d and internal uncertainties) in real time. Its inputs are the control signal u and the system output y . Its outputs are z_1 and z_2 , which are estimated values of the system output y , and the generalized disturbance f , respectively. k_1 and b_0 are tunable controller parameters.

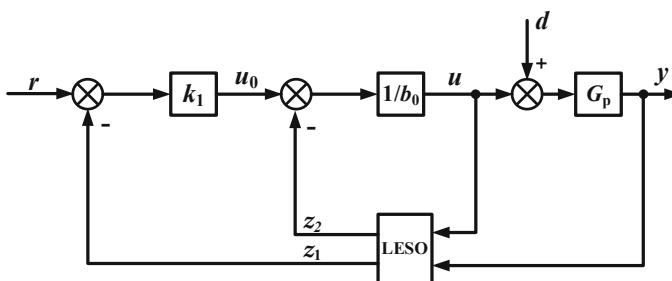


Fig. 1. A first-order LADRC control system.

A second-order LESO is designed as

$$\begin{cases} e = y - z_1 \\ \dot{z}_1 = z_2 + \beta_1 \cdot e + b_0 u \\ \dot{z}_2 = \beta_2 \cdot e \end{cases} \quad (1)$$

where β_1 and β_2 are observer parameters.

According to Fig. 1, the control law is

$$u_0 = k_1(r - z_1) \quad (2)$$

$$u = (u_0 - z_2)/b_0 \quad (3)$$

where u_0 is the nominal control, and u is the control signal of LADRC.

Consider a first-order plant G_p

$$\dot{y} = f + b_0 u \quad (4)$$

where f represents the generalized disturbance.

When parameters are properly determined, a LESO can estimate the output y and the generalized disturbance f accurately, that is, $z_1 \rightarrow y$, $z_2 \rightarrow f$.

By substituting (3) into (4), one has

$$\dot{y} = u_0 + (f - z_2) \approx u_0 \quad (5)$$

From Eq. (5), it can be seen that when generalized disturbance f is estimated accurately, a first-order plant can be approximated to be an integrator.

According to the bandwidth parameterization method [11], k_1 , b_0 , β_1 , and β_2 can be selected as

$$k_1 = \omega_c, \quad \beta_1 = 2\omega_o, \quad \beta_2 = \omega_o^2. \quad (6)$$

where ω_c is the controller bandwidth and ω_o is the observer bandwidth.

3 Scalable Observer Bandwidth Based ADRC

Estimation of the generalized disturbance directly affects the closed-loop performance of an ADRC. Therefore, improving the ability of an ESO is important. According to the bandwidth parameterization approach, adaptive adjustment of an ESO can be realized by a scalable observer bandwidth ω_o . Consequently, a scalable observer bandwidth based ADRC is proposed.

The transfer functions of z_2 can be obtained from Eq. (1)

$$z_2 = \frac{(\beta_2 s)y - (\beta_2 b_0)u}{s^2 + \beta_1 s + \beta_2} \quad (7)$$

From Eqs. (4) and (7), the disturbance observation transfer function between z_2 and f can be written as

$$\phi_1(s) = \frac{z_2}{f} = \frac{\beta_2}{s^2 + \beta_1 s + \beta_2} = \frac{\omega_o^2}{s^2 + 2\omega_o s + \omega_o^2} \quad (8)$$

From Eq. (8), it can be seen that the disturbance observation transfer function is similar to a second-order system's transfer function with $\zeta = 1$. Figure 2 is the frequency response of the disturbance observation transfer function $\phi_1(s)$. It can be seen that a smaller phase lag, faster and more accurate estimation of the generalized disturbance can be obtained by a larger observer bandwidth ω_o .

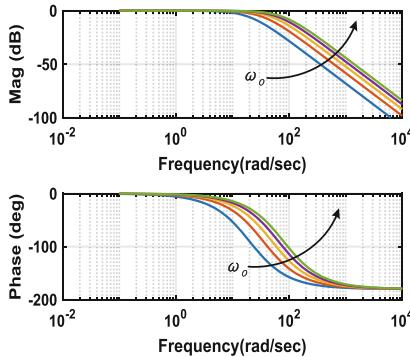


Fig. 2. Frequency characteristics of the disturbance observation transfer function.

To address it, a scalable observer bandwidth based ADRC is proposed. It adjusts the observer bandwidth adaptively according to the changes of estimation errors. Then, transient and steady-state performance can be ensured by an improving disturbance estimation and rejection ability.

An extended state observer with scalable observer bandwidth can be designed as

$$\begin{cases} e = y - z_1 \\ \dot{z}_1 = z_2 + \beta_1(e) \cdot e + b_0 u \\ \dot{z}_2 = \beta_2(e) \cdot e \end{cases} \quad (9)$$

where $\beta_1(e)$ and $\beta_2(e)$ are scalable observer gains, which are functions of the estimation error e . As can be seen later, both $\beta_1(e)$ and $\beta_2(e)$ provide scalable observer bandwidth.

Combined with the idea of bandwidth parameterization, $\beta_1(e)$ and $\beta_2(e)$ can be designed as

$$\beta_1(e) = 2(\omega_o + \Delta\omega_o), \quad \beta_2(e) = (\omega_o + \Delta\omega_o)^2, \quad \Delta\omega_o = k_p \cdot |e|. \quad (10)$$

where $\Delta\omega_o$ is the increment of an observer bandwidth, which is proportional to the absolute value of the estimation error e ; and k_p is a scalable gain.

Obviously, the scalable observer bandwidth is based on the bandwidth parameterization approach. However, the bandwidth is not fixed. There exists a scalable bandwidth increment.

According to Eqs. (8) and (10), an improved disturbance observation transfer function is

$$\phi_2(s) = \frac{(\omega_o + k_p \cdot |e|)^2}{s^2 + 2(\omega_o + k_p \cdot |e|)s + (\omega_o + k_p \cdot |e|)^2} \quad (11)$$

Let $\omega'_o = \omega_o + k_p \cdot |e|$, Eq. (11) can be rewritten as

$$\phi_2(s) = \frac{\omega'^2_o}{s^2 + 2\omega'_o s + \omega'^2_o} = \frac{\omega'^2_o}{(s + \omega'_o)^2} \quad (12)$$

If the denominator polynomial of (12) is set as zero, the characteristic equation is

$$(s + \omega'_o)^2 = 0 \quad (13)$$

Then, poles of disturbance observation transfer function are

$$s_{1,2} = -\omega'_o = -(\omega_o + k_p \cdot |e|) \quad (14)$$

According to Eq. (14), two equal real poles are placed on the negative real axis. The poles distributions of disturbance observation transfer function $\phi_2(s)$ are shown in Fig. 3.

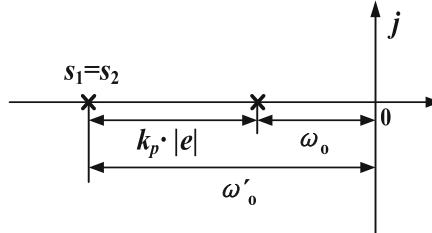


Fig. 3. Observer poles distribution.

According to Fig. 3, when estimation error e increases, observer poles move from ω_o to its left-hand side on the negative real axis. When observer bandwidth increases, the estimation error e decreases, which drives the observer poles approach ω_o . In such way, estimation error e can be kept within a relatively small interval. Observer bandwidth of an ESO is not a fixed parameter any more, but a scalable parameter that is able to be adaptively adjusted in $[\omega_o, \omega_o + k_p \cdot |e|]$. Additionally, initial observer bandwidth ω_o should be chosen properly so as to improve the system's disturbance rejection ability as well as the system's rapidity and stability.

4 System Simulation and Analysis

To verify the SOADRC, a BLDCM model in Simulink library is employed (see Fig. 4). Stator phase resistance $R_s = 2.875 \Omega$, stator phase inductance $L_s = 8.5 \text{ mH}$, inertia $J = 8 \times 10^{-4} \text{ kg m}^2$, viscous damping $F = 0.001 \text{ N m s}$, rated speed $n = 3000 \text{ r/min}$, pole pairs $p = 4$. The DC power supply is 500 V. Firstly, DC motor is started without load to check the dynamic performance. Then, a constant load (3 N m) is added at $t = 0.1 \text{ s}$, and the load is cancelled at 0.15 s. Next, a slope load ($60t \text{ N m}$) is added from 0.3 s and it is cancelled at 0.35 s. Control parameters are listed in Table 1, and simulation results are shown in Figs. 5 and 6.

Table 1. Control parameters.

| Parameter | PID | LADRC | SOADRC |
|-------------------|-------|-------|--------|
| b_0 | — | 9200 | 9200 |
| ω_c | — | 340 | 340 |
| ω_o | — | 1280 | 1280 |
| k_p | — | — | 27 |
| Proportional gain | 0.013 | — | — |
| Integral gain | 16.61 | — | — |
| Differential gain | 0 | — | — |

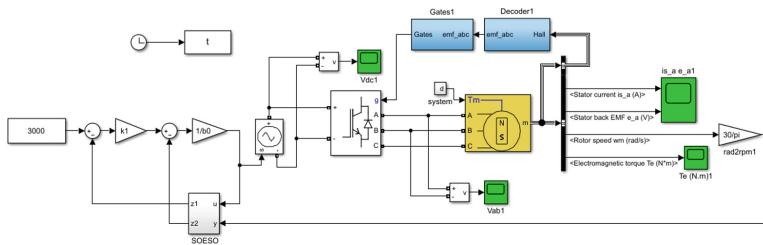


Fig. 4. Simulink module of a closed-loop BLDCM.

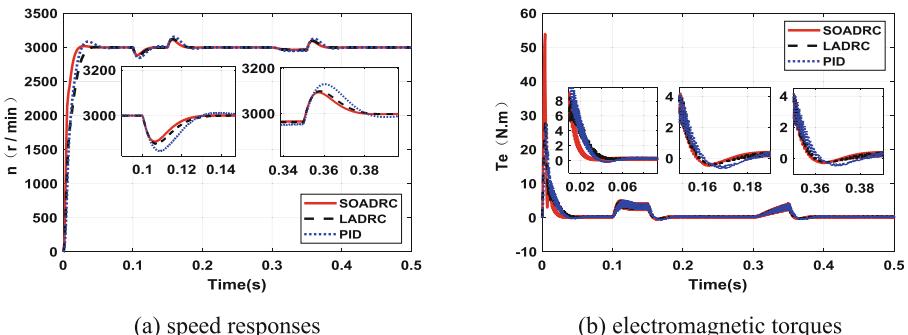


Fig. 5. Comparisons of the system responses.

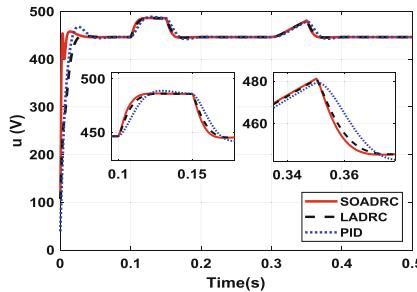


Fig. 6. Comparisons of the control signals.

Table 2. Comparisons of indexes.

| Controller | RMSE | MAE | ITAE |
|-------------|----------|---------|--------|
| PID | 334.3252 | 79.9156 | 2.6257 |
| LADRC | 323.7018 | 74.6752 | 1.8588 |
| SOADRC | 249.7657 | 45.8386 | 1.4700 |
| Improvement | 25.29% | 42.64% | 44.02% |

In Table 2, the improvement rate is the ratio of SOADRC to PID. Root-mean-square error (RMSE), average absolute error (MAE) and integrated time absolute error (ITAE) values listed in Table 2 are calculated based on the deviation between the desired speed and the actual speed.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N error^2}, \quad MAE = \frac{1}{N} \sum_{i=1}^N |error|, \quad ITAE = \int_0^t \tau |error| d\tau$$

where error is the deviation between the desired speed and the actual speed.

From the simulation results, we can see that smallest overshoot and shortest response time can be obtained by SOADRC. For the electromagnetic torque, response by SOADRC is fastest. In addition, for the disturbance rejection performance, we can see that when step load disturbance is added at $t = 0.1$ s and the ramp load disturbance is added at $t = 0.3$ s, the rotation speed controlled by PID changes greatly, the response time is the longest. However, the speed controlled by SOADRC changes smoothly and the response time is the fastest. From Table 2, the accuracy of SOADRC is better than the ones of both LADRC and PID.

5 Conclusion

In this paper, SOADRC is proposed. Since model parameters of a BLDCM are not necessary, system performance is not susceptible to external disturbances. Numerical results also show that closed-loop BLDCM system based on SOADRC is faster, more

accurate and robust than PID and LADRC with fixed parameters. It may be a feasible approach to improve the performance of a BLDCM.

Acknowledgments. This work is supported by Key program of Beijing Municipal Education Commission (KZ201810011012), National Natural Science Foundation of China (61873005), and Support Project of High-level Teachers in Beijing Municipal Universities in the Period of 13th Five-year Plan (CIT&TCD201704044).

References

1. Liu G, Minjie (2016) Research on reference model adaptive control system of brushless DC motor. *J Hubei Univ Technol* 31(02):73–76
2. Luo X, Fan G (2018) Simulation study of BLDCM fuzzy-PI adaptive control system. *Micromotors* 51(12):67–71
3. Shao B, Wang B, Shang J et al (2016) Application of fuzzy PID methods in brushless DC motor control system. *Micromotors* 49(08):31–34
4. Wang P, Jiang W, Wang J et al (2018) A current hysteresis control method for brushless DC motor in multi-state with commutation torque ripple reduction. *Trans China Electrotech Soc* 33(22):5261–5272
5. Guan X (2015) Research on control strategies of brushless DC motor based on particle swarm optimization. *Huazhong University of Science & Technology*
6. Liu Z, Xiao Q, Wang N (2018) Feedback linearization control in DFIG DC converter system. *Electron Plann* 47(01):11–14
7. Zhou Y, Zhang L, Wang H et al (2019) Research on FSC-MPC method of permanent magnet synchronous motor for vehicle. *Power Electron* 53(01):42–45
8. Han J (1998) Auto-disturbances-rejection controller and its applications. *Control Decis* 13 (1):19–23
9. Han J (2002) From PID to ADRC. *Control Eng* 9:13–18
10. Han J (2008) Active disturbance rejection control technology: the control technology for estimating and compensating the uncertainties. National Defense Industry Press, Beijing
11. Gao ZQ (2003) Scaling and bandwidth-parameterization based controller tuning. In: *Proceedings of American control conference*, pp 4989–4996
12. Gao Z (2013) Exploration of ADRC thought. *Control Theory Appl* 30(12):1498–1510



Tool Wear State Monitoring Based on Long-Term and Short-Term Memory Neural Network

Zhu Xiang and Xie Feng^(✉)

School of Electrical Engineering and Automation, Anhui University,
Hefei 230601, China
jeexf199@163.com

Abstract. The real-time detection of tool wear condition is the difficult point of machine tool processing status monitoring. The vibration signal of tool wear shows strong nonlinear and non-stationary phenomenon. The neural network model can be used to analyze the vibration signal of the tool to judge the wear state of the tool. To this end, an online monitoring method for tool wear status based on long-term and short-term memory neural network (LSTM) is proposed. Firstly, the tool vibration signal is collected by the acceleration sensor, and then the vibration signal is decomposed by wavelet packet transform to form the energy value corresponding to different frequency bands, which is used as the feature vector input of the long-term and short-term memory neural network. Finally, the long-short-term memory neural network model is used to process the tool vibration signal to judge the wear condition of the tool. In addition, the method is compared with BP neural network and multi-layer BP neural network fault diagnosis method. The results show that the LSTM network method is more effective for on-line monitoring of tool wear status.

Keywords: Tool wear · Long and short term memory neural network · Online monitoring · Wavelet packet transform

1 Introduction

With the extensive use of CNC milling machine in the machining process, tool wear becomes the main factor for the failure of CNC milling machine, while CNC milling machine are automatically programmed according to programming. Therefore, it is of great significance to monitor the cutting state of the tool in real time. The change of the tool wear state is due to the increase of the cutting force and vibration of the tool as the tool wear increases during the cutting process, and exhibits significant nonlinear and non-stationary features. In the actual cutting process, the vibration signal of the tool is not only affected by the degree of tool wear, but also by the processing parameters, machine tool transmission system, work piece material, tool structure and other factors. Due to the frequency bandwidth of the signal, the vibration signal caused by tool wear is difficult to identify. In addition, due to the limited number of experiments, the actual number of samples obtained is small, resulting in few experimental samples, resulting in a mathematical model that is not ideal. Therefore, tool wear signals are effective

recognition and finite model training samples are two prominent problems in tool wear state recognition, which increases the difficulty of tool wear state monitoring.

One of the methods for on-line monitoring of tool wear state is to determine the wear state of the tool by collecting the acceleration signal on the tool and then analyzing the acquired signal in the time or frequency domain. At present, the method of online monitoring of tool wear by using the acceleration signal of the tool mainly includes time domain analysis, frequency domain analysis, wavelet transform, Artificial Neural Network (ANN) and the like. Artificial neural networks [1–3] generally start from the input layer to the hidden layer, and then from the hidden layer to the output layer. The layers are fully connected, and the nodes between each layer are disconnected. To this end, Yu et al. [4–7] proposed a method of cyclic neural network (RNN), the nodes between the hidden layers of the cyclic neural network are no longer connected but connected, and the input of the hidden layer includes not only the input layer.

The output also includes the output of the hidden layer at the previous moment. The circulatory neural network memorizes the previous information and applies it to the calculation of the current output, thus avoiding the loss of the original information. However, it is difficult for the cyclic neural network to deal with the problem that the time series is too long. If the time series is too long, the information will disappear when feedback is generated.

In response to the above problems, Hochreiter et al. [4, 5] proposed a long-and short-term memory neural network method, which is a special cyclic neural network that suppresses the disappearance of gradients by increasing the threshold in the circulating neural network. It is a model for processing and predicting very long intervals and delays in time series. It is suitable for dynamic time series, especially for non-stationary, low repeatability and reproducible signals. It has strong pattern classification ability and it has discrete pairs. The input signal features have the advantages of high recognition efficiency and high speed. In this paper, the non-stationary vibration signal of the tool cutting process is first collected, and then the low-frequency vibration of the tool is removed by wavelet packet transformation. Only the high-frequency signal that affects the tool wear in the tool vibration signal is retained as the input of the long-term and short-term memory network, and finally the long-term and short-term memory is utilized. The neural network further extracts the characteristics of the wavelet packet energy, thereby better realizing the online monitoring of the tool wear state.

2 Long-Term and Short-Term Memory Neural Network Method

2.1 Cyclic Neural Network (RNN)

RNN is called cyclic neural network because the general neural network is the input layer to the hidden layer, and then from the hidden layer to the output layer and the hidden layer is disconnected, which easily leads to the loss of original information. The hidden layers of the cyclic neural network are related, and the previous information is memorized and applied to the output calculation of the current hidden layer, that is, the

nodes between the hidden layers are no longer connected but connected, and hidden. The input of the layer includes not only the output of the input layer but also the output of the hidden layer at the previous moment. Figure 1 is a block diagram of the RNN network.

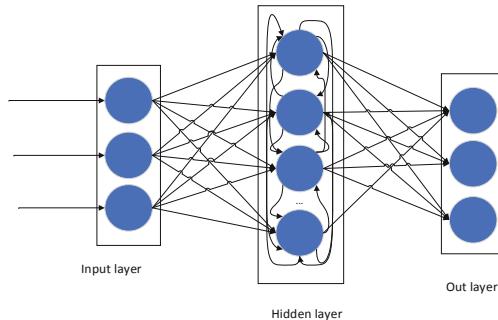


Fig. 1. RNN structure diagram

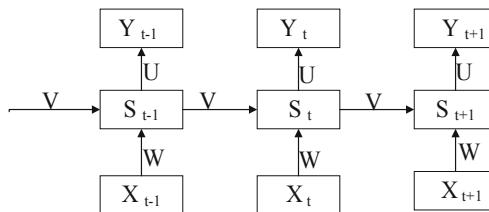


Fig. 2. Development of RNN neural network

In the RNN, the input indicating the time t , indicating the state of the hidden layer at time t , indicating the output, the aspect of the RNN different from the neural network in machine learning is that the input of the hidden layer is composed of two parts, one part is the current input, and the other part is the other part (Fig. 2).

$$S_t = g(W \times X_t + V \times S_{t-1}) \quad (1)$$

$$Y_t = \sigma[U \times S_t] \quad (2)$$

Where: V is the weight matrix of the output layer; g is the output layer activation function; W is the input weight matrix; U is the weight matrix input at the moment; σ is the hidden layer activation function.

The problem with RNN is that it is difficult to deal with nodes that are far apart in the processing time series, because the connection of nodes in the computer involves the multiplication of the matrix, which will cause the gradient to disappear. In order to solve this problem, many solutions have been made. The method of the problem: such

as ESN (Echo state Network), increase the leakage unit, etc., the most widely used is the long- and short-term memory neural network method by increasing the threshold on the RNN.

2.2 Long-Term and Short-Term Memory Neural Network Structure

The long- and short-term memory neural network is a special RNN network structure. By adding a “gate” structure in the RNN to remove or add information to the hidden layer, it prevents the gradient disappearing when the feedback is too long. Gates are a way to pass information through, they contain a sigmoid neural network layer and a bitwise multiplication operation. LSTM makes the weight of the self-loop change by increasing the input gate, the forget gate and the out gate. Thus, when the model parameters are fixed, the integration scale at different times can be Dynamically changing, thus avoiding the problem of gradient disappearance or gradient expansion. A unit structure of long- and short-term memory neural networks is shown in Fig. 3.

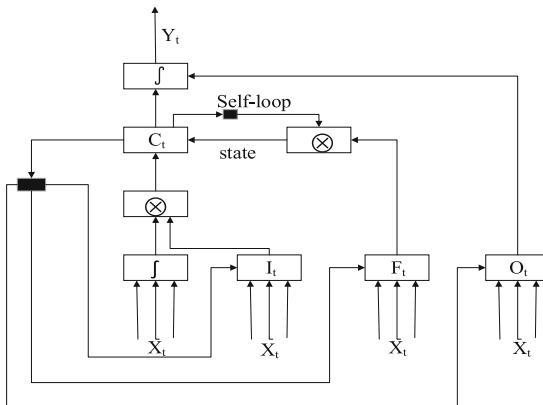


Fig. 3. Long-short-term memory neural network

According to the structure of the long-and short-term memory neural network, in the calculation of each long- and short-term memory neural network unit, where F_t represents the forgetting threshold, I_t represents the input threshold, and C_{t-1} represents the previous moment of LSTM. The cell state, C_t represents the LSTM cell state (where the loop occurs), O_t represents the output threshold, Y_t represents the output of the current cell, and h_{t-1} represents the output of the previous cell. The input gate is used to control the information input, the forget gate is used to control the retention of the historical state information of the LSTM unit, and the output gate is used to control the information output. The activation function causes the output value of the forgetting gate to be between [0, 1]. When the forgetting gate output is 0, it means that all the information of the previous state is discarded. When it is 1, the information indicating

the previous state is all retained. The following is a calculation formula for a unit of long-term and short-term memory neural networks;

$$F_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \quad (3)$$

The forgotten threshold is the control threshold between the long-term information of the previous moment and the long-term information of the current moment.

$$I_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \quad (4)$$

The input threshold is the control threshold between the current time and the short state information and the long state information.

$$O_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o) \quad (5)$$

The output threshold is the control threshold between the current memory state information and the output state information.

$$C_t = F_t * C_{t-1} + I_t * \bar{C}_t \quad (7)$$

The state information at time t is that the input information of the unit is composed of thresholds.

$$Y_t = O_t * \tanh(C_t) \quad (8)$$

The output of the time t is composed of the threshold information and the state information of the t time.

The output of the LSTM network will be obtained, hiding the layer unit. For the classification problem, it can be mapped to the linear output layer whose weight matrix is $W(s)$, and then the softmax function is used to calculate the probability distribution y'_t of the classification category, and then the cost function is calculated.

$$y'_t = \text{softmax}(W^s Y_t) \quad (9)$$

$$J^{(t)}(\theta) = - \sum_v y_{t,j} \times \log_2(y'_{t,j}) \quad (10)$$

Where: v represents the number of categories; $y_{t,j}$ represents the true probability of belonging to class j at time t; $y'_{t,j}$ represents the training probability of class j at time t.

3 Milling Test and Feature Extraction

3.1 Wear State of Tool for CNC Milling Machine

The degree of tool wear [6], usually the wear amount VB of the tool flank is selected to measure the degree of wear of the tool. In this experiment, 300 strokes were cut by an end mill. After each cutting process, the amount of wear of the flank of the tool was measured by a microscope and recorded. The variation curve of the flank wear of the three cutting edges is shown in Fig. 4.

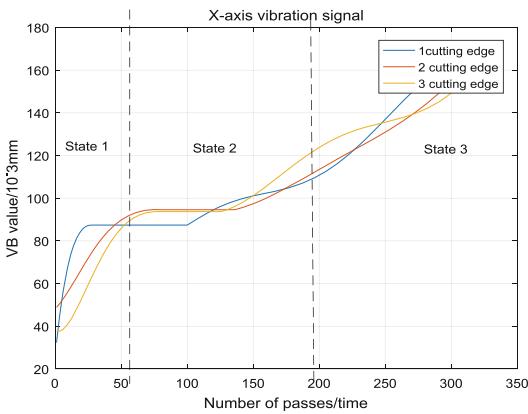


Fig. 4. Cutting edge flank wear variation curve

According to the VB value of the flank wear of the tool, the tool wear can be divided into three states: initial wear (state 1), normal wear (state 2) and sharp wear (state 3), where the tool is not worn at the initial stage and normal wear. The tool needs to be changed, and the tool needs to be replaced when the sharp wear state occurs. The state of the tool will be identified below.

In the process of machine tool cutting, different frequencies of vibration will occur when tool and work piece collide with each other in different degree of wear, so the vibration signal of tool can effectively reflect the real situation of tool wear. In this paper, the vibration signal of the CNC milling machine ZXJ7016 is taken as the processing object, and the vibration signal of the tool is preprocessed and the fault feature is extracted. Fault data of tool wear are acquired through experiments, and fault data of 1024 mm length are intercepted for data preprocessing. The experimental platform is built as shown in Fig. 5 below.

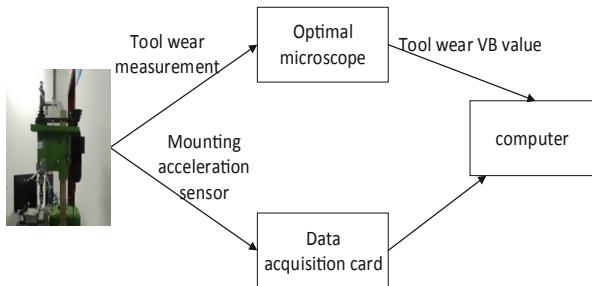


Fig. 5. Test system block diagram

3.2 Feature Extraction

In this paper, the vibration signal collected by each cutting test is decomposed by wavelet packet, and the signal is decomposed into different frequency bands according to the frequency band. Thus, the parameters such as energy value, power spectrum, and spectrum corresponding to different frequency bands reflect the characteristics of the frequency band. During the machining of the workpiece, the vibration signal of the workpiece is aggravated by the wear of the tool, and the cutting speed of the tool is fixed. Therefore, the vibration signal energy causing the tool wear is concentrated in a certain frequency band, and the frequency band includes the information on the wear of the tool, so the energy of the corresponding frequency band of the tool wear can be used as the characteristic parameter to monitor the wear state of the tool.

The collected vibration signal is subjected to wavelet packet decomposition by wavelet packet transform technique. The wavelet packet transform is a linear transform, and its expression is:

$$E_n(x^{k,m}(i)) = \frac{1}{2^{-k}N - 1} \sum_{i=1}^{2^{-k}N} (x^{k,m}(i))^2 \quad (11)$$

It can be seen from the above formula that the wavelet packet coefficient $C(jk)$ has the dimension of energy, and the wavelet coefficient of the vibration signal can be used to obtain the variation law of the energy value after the wavelet packet decomposition, so as to reflect the vibration condition of the cutting process. In this test, the vibration signal of the tool X direction is analyzed by 3-layer wavelet packet decomposition, and the wavelet function is selected as db3. After 3 layers of wavelet decomposition, the obtained 8 wavelet coefficients are normalized to the energy value, and this is used as the eigenvalue input into the long-term and short-term memory neural network. The LSTM network is used to classify the energy values of the unknown state wavelet transform. Discrimination to effectively identify the wear state of the tool. Figure 6 is the energy of each frequency band corresponding to the three-layer wavelet decomposition of the vibration signal in the feed direction collected by the tool in the initial state of initial wear, normal wear and rapid wear.

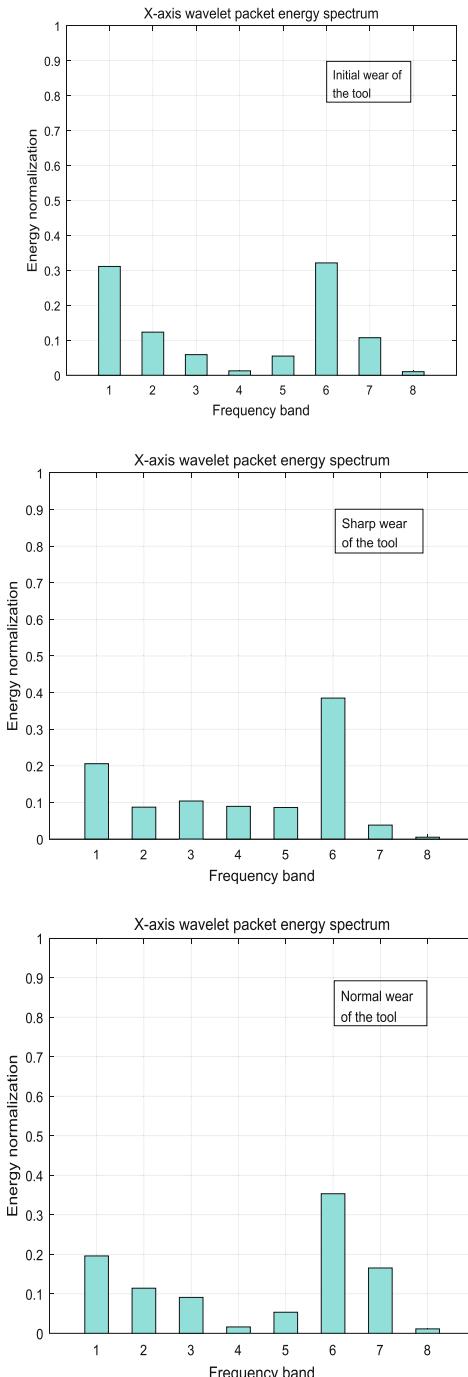


Fig. 6. Normalized histogram of energy in each band of tool wear

It can be seen from the figure that the energy of the vibration signal in each frequency band is different when the tool is in different wear states. In different wear states, the energy of the frequency bands in the 1st, 5th, 6th, and 7th bands is significantly different in the figure, while the energy in other frequency bands only has some slight differences. Therefore, Ex1, Ex5, Ex6, and Ex7 are selected as the feature vector D2 to input the long-term and short-term memory neural network.

In order to verify the effectiveness of the proposed algorithm, 231 data were randomly selected from 300 sets of tool wear state samples as training samples, and 99 groups were used as test data. In the long-term and short-term memory neural network model, there are 100 neurons in the hidden layer and 3 neurons in the output layer. The loss function uses Mean Absolute Error (MAE) and the optimization algorithm uses Adam. According to the characteristics of the tool wear state, the long-term and short-term memory god memory neural network expects to output a 3-dimensional vector, and when the expected output is sharply worn, the output [001].

In order to reduce the feature vector dimension, the fault feature vector is combined in different forms. Table 1 is the algorithm diagnosis result of LSTM under different tool wear feature vector combinations.

Table 1. Tool classification accuracy rate

| Different algorithms | Feature vector | Test set accuracy |
|----------------------|----------------|-------------------|
| BP | D1 | 82.67% |
| | D2 | 85.74% |
| Multi-layer BP | D1 | 87.23% |
| | D2 | 89.37% |
| LSTM | D1 | 91.345 |
| | D2 | 95.67% |

It can be seen from Table 1 that the wavelet packet energy value D1 of the vibration signal decomposition is selected as the feature vector of the tool wear and the identification accuracy is lower when the wavelet packet energy value D2 is selected as the tool wear characteristic vector. Therefore, the feature dimension is reduced, and the sample discrimination is improved to further improve the accuracy.

4 Test Conclusions

A long-term and short-term memory neural network model algorithm for tool wear state recognition is constructed. This method combines the advantages of wavelet packet transform and neural network self-learning ability, and introduces a “threshold” to overcome the tool wear state using traditional neural network. When it causes the disadvantage that the gradient disappears. The accuracy of the test set of the long- and short-term memory neural network is higher than that of the traditional BP neural network and the test set of the multi-layer BP neural network. Therefore, the LSTM

neural network can effectively avoid the neural network falling into the local optimal solution during training, and further improve the performance of the model. This is of great significance for the establishment of an on-line monitoring system which can timely and accurately evaluate the wear status of end milling cutters in the future. Because this research is aimed at off-line evaluation of collected data in the environment of MATLAB and PYTHON on PC, in the next step, the analysis code can be integrated into LabVIEW environment through MathScript nodes, and the on-line monitoring system of tool wear state can be developed to realize on-line monitoring of tool wear state in milling process.

References

1. Zhou Z (2017) Machine learning, vol 6. Tsinghua University Press, Beijing, pp 121–133
2. Ghate VN, Dudul SV (2010) Optimal MLP neural network classifier for fault detection of three phase induction motor. *Expert Syst Appl* 37(4):3468–3481
3. Zhang C, Zhang H (2016) Modeling and prediction of tool wear using LS-SVM in milling operation. *Int J Comput Integr Manuf* 29(1):76–91
4. Heng YY (2018) Research on the structure model of cyclic neural network. Zhejiang University
5. Da WL (2010) Theory and application of fault diagnosis based on vibration signal analysis. Northeast Petroleum University
6. Qing CR, Gen S (2005) Rotating machinery fault diagnosis method based on recurrent neural network. *Vibr Test Diagn* 03:70–72
7. Sai T, Wei H, Yue ZJ, Jun YA (2018) Bearing fault identification based on long and short term memory network. *J Autom Eng* 8(04):297–303



A New V/f Control Method for Permanent Magnet Linear Synchronous Motor Drives

Leilei Cui¹, Hongwei Zhang^{1(✉)}, and Mingren Wang²

¹ Henan Polytechnic University, Jiaozuo 454000, China

zhanghw@hpu.edu.cn

² Wuxi Tongfang Juneng Control Technology Co., Ltd., Wuxi 214000, China

Abstract. This paper presents a new V/f control method for permanent-magnet linear synchronous motor (PMLSM) drives. The proposed control strategy combines the merits of traditional V/f and field-oriented control strategies to achieve simple and high-performance motor drive. V/f control for PMLSM is able to get rid of the complex sensor position calculations, and it is of high research value. Two speed closed-loop are introduced to compensation the V/f voltage vector reference and improve overall system stability. Meanwhile, in order to enhance the efficiency of the system, an $i_d = 0$ strategy is proposed, which makes it more robust against transient disturbances. Finally, simulation and experimental results are provided to validate the effectiveness of the fidelity of the proposed strategy.

Keywords: Permanent-magnet linear synchronous motor (PMLSM) · V/f control · Stable control · Efficient control · Sensorless control

1 Introduction

The linear propulsion system with permanent-magnet linear synchronous motor as the power device does not need intermediate mechanisms such as gears, connecting rods and chain. It has the advantages of simple structure, high power density, high control precision, low noise, high efficiency and convenient maintenances [1]. However, it is difficult to control a synchronous motor because it is necessary to determine the initial position of the encoder in advance. Expensive sensor configuration limits the development of synchronous motors [2]. With V/f control, position sensors can be saved to reduce system cost and increase system reliability.

There are basically two major conventional sensorless control structures: field-oriented control (FOC) [3–5] and scalar control [6–8]. In the FOC, the rotor position information is obtained by designing the observer, and the error can be optimized by the current controller.

Some scholars have studied the V/f control method to realize the simple control scheme similar to the one for the asynchronous motor. It is known that the traditional V/f open loop control is unstable. In [7, 9, 10], the instability of open-loop V/f control is confirmed. Compared with vector control, traditional V/f control lacks effective control of voltage phase angle. Scholars have designed different closed-loop control to achieve voltage amplitude and phase angle control at the same time. In [11],

compensate voltage by balancing reactive power. Improve system efficiency by controlling power factor angle in [12]. A maximum torque per ampere plug in block is implemented to optimize torque generation efficiency without intensive power or power factor angle calculations in [9] and [10], on the basis of transient analysis, study the conditions required to approach steady state, the magnitude and phase angle of the generated vector voltage are compensated by comparing the actual output values with reference values.

The rest of this paper is organized as follows: the stability of V/f control for PMLSM is analyzed at first. Then the control method characterized by high stability and efficiency PMLSM V/f control is proposed. At last, simulation and experimental results are demonstrated to validate the fidelity of the control system.

2 Control Method and Stability Analysis

2.1 Analysis of PMLSM Stability

Permanent-magnet linear synchronous motors have similar characteristics to synchronous rotating electrical machines. Here we use rotational coordinates to facilitate analysis of the problem. To simplify the stability analysis, in this paper is considered only the surface magnetic motor, which has identical d-axis and q-axis inductances.

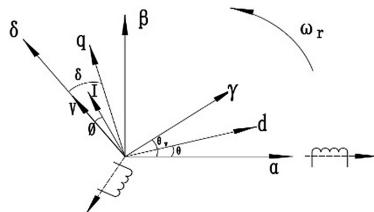


Fig. 1. Definition of coordinate reference frames.

Generally speaking, d-axis is defined along permanent magnets flux and q-axis is defined along motor electromotive force vector. However, in V/f control, the rotor position is unknown. In V/f control, the coordinate system needs to be re-established to illustrate the positional relationship. The d-q axis represents the rotor rotation coordinate axis, and the $\gamma-\sigma$ axis represents the voltage synthesis vector coordinate axis. Figure 1 shows defined reference coordinates in the proposed V/f method. Small signal analysis can be used to analyze the stability of the system. The voltage equation for the PMLSM under the d-q axis is as follows

$$\begin{cases} u_d = R i_d + L_d \frac{d i_d}{dt} - \frac{\pi v}{\tau} L_q i_q \\ u_q = R i_d + L_q \frac{d i_q}{dt} + \frac{\pi v}{\tau} L_d i_d + \frac{\pi v}{\tau} \psi_f \end{cases} \quad (1)$$

Stator flux linkage equation is

$$\begin{cases} \psi_d = L_d i_d + \psi_f \\ \psi_q = L_q i_q \end{cases} \quad (2)$$

The formula for calculating the electromagnetic thrust is

$$F_e = 1.5 P_n \frac{\pi}{\tau} [(L_d - L_q) i_d i_q + \psi_f i_q] \quad (3)$$

Motion equation is

$$M \frac{dv}{dt} = F_e - F_l - Bv \quad (4)$$

The angle σ represents the power angle, and the relationship between the voltage vector angle θ_v and the rotor rotation angle θ is as follows

$$\delta = \theta_v - \theta \quad (5)$$

Thus, the relationship between the combined voltage and the d-q axis voltage is expressed as

$$\begin{cases} u_d = -u_s \sin \delta \\ u_q = u_s \cos \delta \end{cases} \quad (6)$$

Combine the above equations and linearize to get the following expression

$$\begin{cases} p i_d = -\frac{R i_d}{L_s} + \frac{\pi v}{\tau} i_q - \frac{u_s \sin \delta}{L_s} \\ p i_q = -\frac{R i_q}{L_s} - \frac{\pi v}{\tau} \left(\frac{\psi_f}{L_s} + i_d \right) + \frac{u_s \cos \delta}{L_s} \\ p v = \frac{3}{2} P_n \frac{\pi}{\tau M} \psi_f i_q - \frac{F_l}{M} - \frac{B v}{M} \\ p \delta = p \theta_v - p \theta \end{cases} \quad (7)$$

In (1)–(7)

- id, iq Equivalent currents in rotor d-q axis coordinates
- Ld, Lq Equivalent inductance of rotor d-q axis coordinates
- Ls Ld = Lq = Ls Equivalent synchronous inductance
- v Linear motor speed
- M Drive system quality
- B Viscous friction coefficient
- Pn Number of pole pairs
- p Operator d/dt
- Fe Electromagnetic thrust
- F_l External disturbance set
- ψ_f Permanent magnet flux linkage
- σ Power angle

After the system reaches steady state, the derivative term becomes zero. In the small signal analysis of (7), the second sub-small is ignored, and the motor characteristic equation is as follows

$$p \begin{pmatrix} \Delta i_d \\ \Delta i_q \\ \Delta v \\ \Delta \delta \end{pmatrix} = \begin{pmatrix} -\frac{R}{L_s} & \frac{\pi v_0}{\tau} & i_{q0} & -\frac{u_{s0} \cos \delta_0}{L_s} \\ -\frac{\pi v_0}{\tau} & -\frac{R}{L_s} & -\left(\frac{\psi_f}{L_s} + i_{d0}\right) & -\frac{u_{s0} \sin \delta_0}{L_s} \\ 0 & \frac{3}{2} P_n \frac{\pi}{\tau M} \psi_f & -\frac{B}{M} & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \Delta i_d \\ \Delta i_q \\ \Delta v \\ \Delta \delta \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{1}{M} \\ 0 \end{pmatrix} \Delta F_l \quad (8)$$

Consider that the machine's mechanical time constant is much longer than the electrical time constant. The following equation can be derived from the above equation

$$\begin{bmatrix} p\Delta v \\ p\Delta \delta \end{bmatrix} = \begin{bmatrix} -\frac{P_n \psi_f^2 \tau R}{ML^2 \pi^2 v_0^2} & \frac{P_n \psi_f^2}{ML} \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \Delta v \\ \Delta \delta \end{bmatrix} + \begin{bmatrix} \frac{P_n \psi_f \sin \delta_0}{ML^2 \pi^2 v_0} & \frac{P_n \psi_f \tau^2}{ML \pi^2 v_0^2} \left(\frac{P_n \psi_f L^2 v_0^2 \delta_0 \pi i_{d0}}{\tau} - R \cos \delta_0 i_{\gamma_0} - \frac{v_0 \pi L \sin \delta_0}{\tau} i_{\gamma_0} \right) \\ 0 & 1 \end{bmatrix} \quad (9)$$

Equation (9) indicates that the motor is a second-order system. In order to analyze the stability and response of the open-loop V/f system, the damping coefficient and the natural angular frequency of the system are calculated, in the medium-speed and high-speed regions ($\omega L \gg R$)

$$\begin{cases} \zeta = 0 \\ \omega_e = \frac{P_n \psi_f \tau}{\pi \sqrt{ML}} \end{cases} \quad (10)$$

It can be concluded that the motor system in the medium and high speed state is in an undamped state, the response oscillation is strong, the oscillation frequency is high, and the change of the input amount cannot be accurately tracked, so the simple V/f open loop control of the permanent magnet synchronous motor is un-stable. To improve the stability of the V/f control system and achieve its efficient and stable control, the system must be compensated for feedback (Fig. 2).

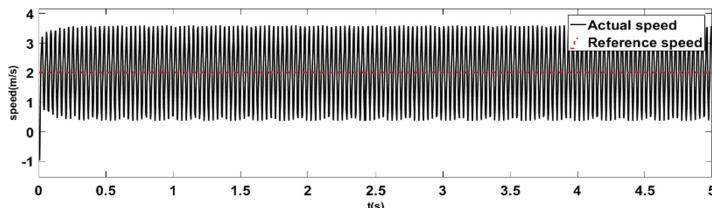


Fig. 2. V/f open loop control system speed response at high frequencies

2.2 Stable V/f Control of PMLSM

As in [9, 10] we need to achieve a stable system by introducing a closed loop in V/f control method. The reference vector voltage is first generated by the SVPWM inverter by introducing a reference vector voltage. Through the selection and comparison of the steady-state reference, the input vector voltage amplitude and phase angle are calibrated in real time, and finally the dynamic characteristics of the system are maintained to a stable state.

This paper chooses to discuss in the rotating coordinate system of the motor, and determines the appropriate d-q reference voltage and reference rotation angle through analysis. Finally, the d-q axis voltage and rotation angle are calibrated according to the steady state characteristics of the system. The steady state of the motor system is usually based on the rotating coordinate system. This analysis method not only does not require complex spatial coordinate angle transformation, but also avoids introducing new disturbance variables while simplifying the calculation.

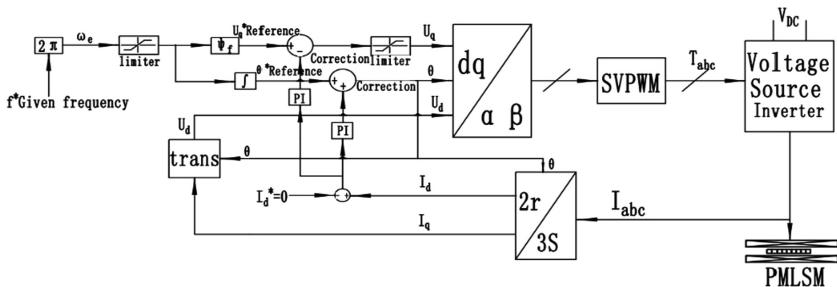


Fig. 3. Overall block diagram of the proposed algorithm

Figure 3 shows overall block diagram of the proposed algorithm. As in Eq. (1), the stator voltages V_d and V_q are typically used to generate the PWM signals for inverter. However, it should be noted that magnetic fluxes generated by $L_d * I_d$ and $L_q * I_q$ are typically very small (i.e., ≈ 0), and voltage generated by $I_q * R$ (even under full load) is very small (i.e., ≈ 0), compared with the flux linkage of the permanent magnet ψ_f . Thus, equations can be reduced as follows:

$$\psi_d = I_d L_d + \psi_f \approx \psi_f \quad (11)$$

$$\psi_q = I_q L_q \approx 0 \quad (12)$$

$$u_q = i_q R + L_q \frac{di_q}{dt} + \omega \psi_d \approx \omega \psi_f \quad (13)$$

The proposed V/f control strategy differs from the traditional open-loop V/f control strategy in that the voltage vector is calibrated in real time. The offset of the system stability reference is analyzed under transient disturbance, and the amplitude and angle of the generated voltage are calibrated by two PI controllers. The u_q reference value is provided by the product of the angular frequency and the flux linkage, and the inaccurate part is calibrated by the feedback link. The integral of the angular frequency is exactly the angle of rotation, u_d is calculated by the d-axis voltage equation.

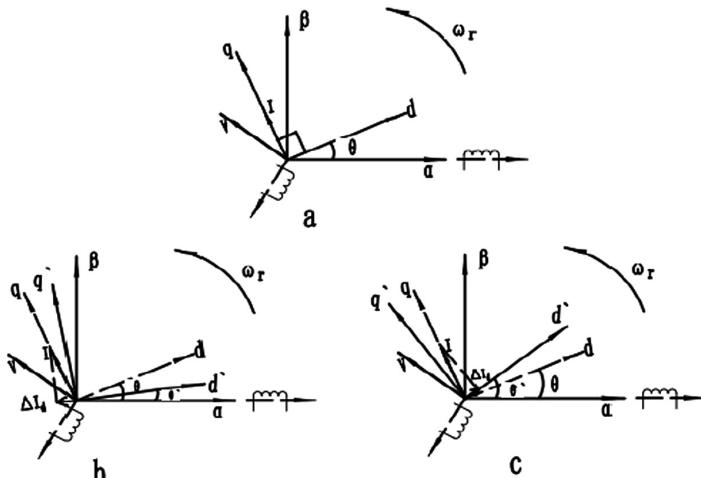
2.3 Efficiency Control and Transient Analysis

This paper chooses $i_d = 0$ control strategy to control the system to achieve higher efficiency. The two closed loops in the control block diagram provide system stability, and the $i_d = 0$ strategy is an efficiency control. In the transient analysis of the system, the state of the system is determined based on the error between the actual value of the i_d and the ideal value.

If the type of transient can be determined by the Δi_d , then we can obtain the steady-state effect by correcting the reference voltage vector by the steady-state loop.

Figure 4a shows the phasor diagram of the motor in steady state. In steady state, the vector current synthesized by the stator is aligned with the q-axis and the projection on the d-axis is zero. This also means that under the control strategy of $i_d = 0$, the projection current on the d-axis determines whether the system is in steady state. As a result, the voltage vector reference becomes inaccurate under such circumstances assuming that no correction is done by stabilizing loops.

Figure 4b and c show phasor diagrams for deceleration (suddenly reducing load) and acceleration (suddenly increasing load). During deceleration, the actual rotor position angle θ' becomes smaller than the commanded rotor position angle θ with respect to the voltage vector phase angle reference. The new rotation axis d'-q' is deflected counterclockwise by a certain angle from the previous d-q axis ($\Delta\theta > \approx 0$ in transient analysis). d-axis projection current in a new coordinate system $\Delta I_d < 0$. The feedback d-axis voltage u'_q transformed by using the nonoptimal angle reference becomes smaller than the d-axis voltage reference u_q ($u'_q < u_q$). The reference voltage u_d at the previous moment should be reduced, and the reference rotation angle u_d should be increased to bring the system to a new steady state. In the transient analysis of the acceleration process, this is exactly the opposite.



a Steady state phase diagram
b, c Two phase diagrams of out-of-step conditions

Fig. 4. The phasor diagram of the motor

3 Simulation and Experimental Results

In order to validate the fidelity of the proposed control scheme, both simulation and experimental tests are implemented. The PMLSM parameters are shown in Table 1.

Table 1. The PMLSM parameters of the test.

| Name | Symbol | Parameter |
|-------------------------------|----------|-----------|
| Stator resistance | R | 7.9 |
| Synchronous inductor | L_s | 0.052 |
| Number of pole pairs | P_n | 8 |
| Permanent magnet flux linkage | ψ_f | 0.655 |
| Motor quality | M | 5.1 |
| Pole distance | τ | 0.0025 |

3.1 Simulation Results

In Figs. 5 and 6, the electromagnetic thrust and the motor speed are compared by the conventional V/f open loop control method and the high efficiency V/f control method, respectively. Figure 7 shows a comparison of the actual operating speed of the motor with the reference speed obtained by the high-efficiency V/f control method ($v = \omega\tau/\pi$, Set $F_l = 300 \text{ N}$ $f = 10 \text{ Hz}$).

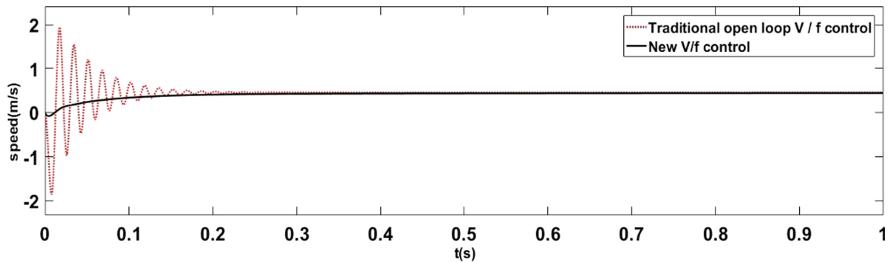


Fig. 5. Traditional V/f open loop control and high efficiency V/f control thrust simulation curve

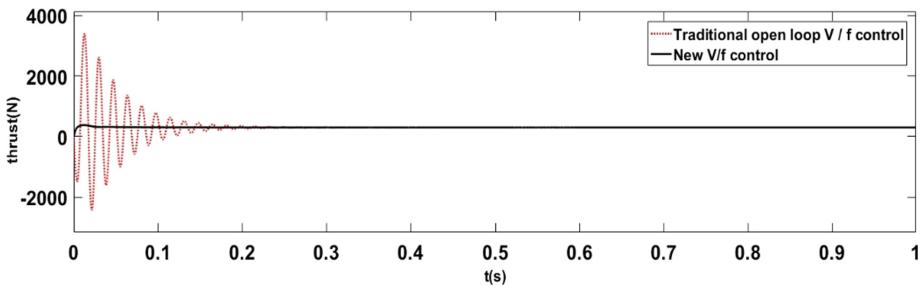


Fig. 6. Traditional V/f open loop control and high efficiency V/f control speed simulation curve

The results show that, although the region is stable in the end, the open-loop V/f control has a large oscillation in the initial stage. In contrast, efficient V/f control is a relatively smooth ascent process at the initial stage.

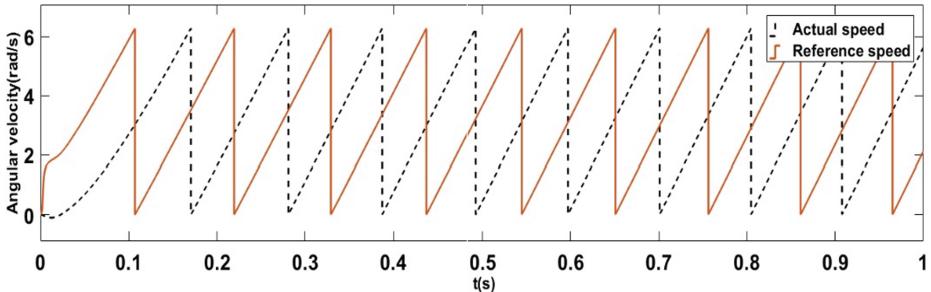


Fig. 7. Actual speed and reference speed comparison chart

Figures 8 and 9 respectively set different frequencies to observe the motor speed and electromagnetic thrust. New V/f control at different frequencies enables the motor to reach a steady state faster and more smoothly.

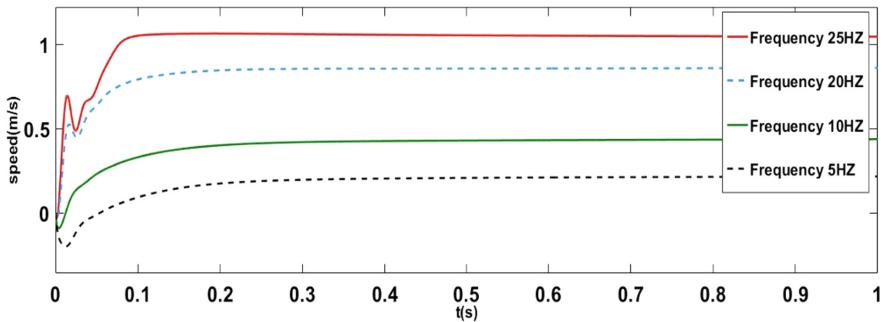


Fig. 8. Motor speed at different command electrical frequencies

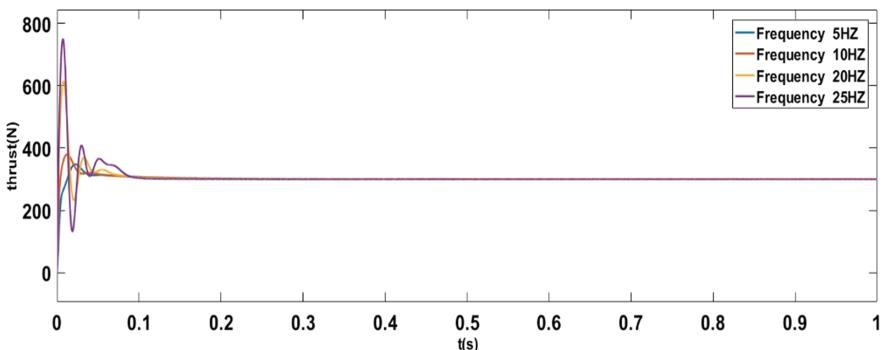


Fig. 9. Output thrust at different command electrical frequencies

3.2 Experimental Results

Finally, the experiment was carried out using the direct-drive home elevator experimental platform of the Process Control Laboratory of Henan Polytechnic University. The experimental platform is shown in Fig. 10. The experimental results are as follows (Fig. 11).

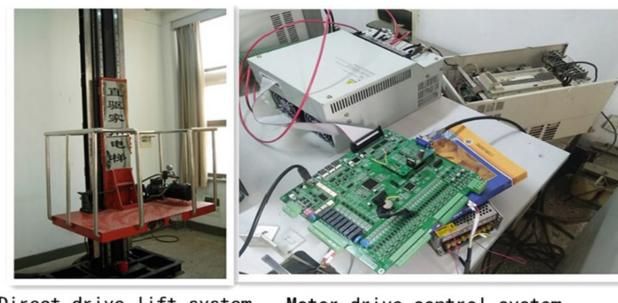


Fig. 10. The experimental platform

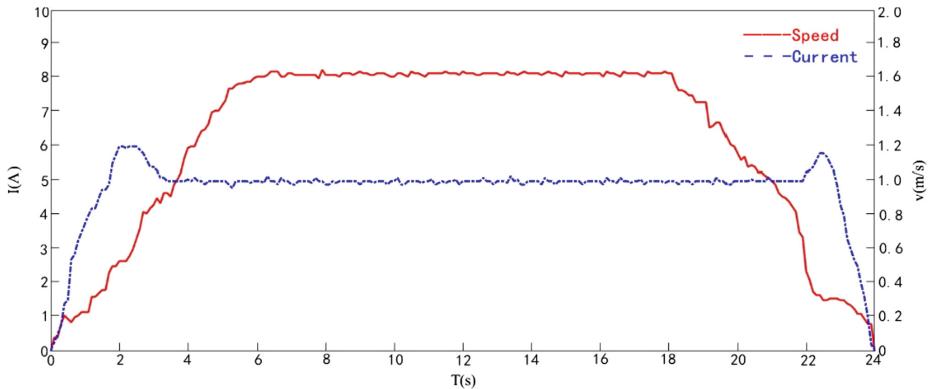


Fig. 11. Speed and current curves in manual mode

4 Conclusion

This paper presents a new V/f control scheme for permanent-magnet linear synchronous motor (PMLSM) drives. The proposed V/f control scheme not only achieves the purpose of stable control, but also achieves high efficiency control ($i_d = 0$). The control algorithm includes two closed-loop corrected u_q amplitudes and rotation angles, and an $i_d = 0$ strategy to drive the system to achieve efficiency. Its simple control structure saves cost and simplifies control algorithms. Its insensitivity to machine parameter changes and no speed feedback are key features of cost savings, which allows variable speed drives to operate in harsh environments. The advantages of the proposed V/f control method are verified through both simulation and experimental results.

References

- Cho K, Kim J, Park H, Choi S (2012) Periodic adaptive disturbance observer for a permanent magnet linear synchronous motor. In: Proceedings of IEEE 51st annual conference decision control, pp 4684–4689
- Paulus D, Stumper J-F, Kennel R (2013) Sensorless control of synchronous machines based on direct speed and position estimation in polar stator-current coordinates. *IEEE Trans Power Electron* 28(5):2503–2513
- Morimoto S, Kawamoto K, Sanada M, Takeda Y (2002) Sensorless control strategy for salient-pole PMSM based on extended EMF in rotating reference frame. *IEEE Trans Ind Appl* 38:1054–1061
- Rashed M, MacConnell PFA, Stronach AF, Acarnley P (2007) Sensorless indirect-rotor-field-orientation speed control of a permanent-magnet synchronous motor with stator-resistance estimation. *IEEE Trans Ind Electron* 54(3):1664–1675
- Genduso F, Miceli R, Rando C, Galluzzo GR (2010) Back EMF sensorless-control algorithm for high-dynamic performance PMSM. *IEEE Trans Ind Electron* 57(6):2092–2100

6. Agarlita SC, Coman CE, Andreescu GD, Boldea I (2013) Stable V/f control system with controlled power factor angle for permanent magnet synchronous motor drives. *Electr Power Appl IET* 7:278–286
7. Itoh J-I, Nomura N, Ohsawa H (2002) A comparison between V/f control and position-sensorless vector control for the permanent magnet synchronous motor. In: Proceedings of power conversion conference, pp 1310–1315
8. Moldovan A, Blaabjerg F, Boldea I (2011) Active-flux-based, V/f-with-stabilizing-loops versus sensorless vector control of IPMSM drives. In: Proceedings of International Symposium Industrial Electronics, pp 27–30
9. Tang Z, Li X, Dusmez S, Akin B (2016) A new V/f based sensorless MTPA control for IPMSM drives. *IEEE Trans Power Electron* 31(6):4400–4415
10. Tang Z, Akin B (2015) A robust V/f based sensorless MTPA control strategy for IPM drives. In: IEEE applied power electronics conference and exposition (APEC), pp 1575–1581
11. Jafari SH, Corzine KA, Huang J (2012) Efficiency optimization of a sensorless V/f control method for PMSM. In: 3rd IEEE international symposium on sensorless control for electrical drives (SLED 2012), pp 1–5
12. Andreescu G-D, Coman C-E, Moldovan A, Boldea I (2012) Stable V/f control system with unity power factor for PMSM drives. In: 13th international conference on optimization of electrical and electronic equipment (OPTIM), pp 432–438



Multi-branch Input Structure for Pyramid Scene Parsing Network

Chuanya Wang and Jin Chen^(✉)

Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
cjwoods@163.com

Abstract. Deep convolutional neural networks have been widely researched and have made outstanding achievements in the field of target recognition and image segmentation in recent years. In this paper, we propose a multi-branch input method for aggregating feature information at the early convolutional layer. At present, the popular neural network model uses single branch down-sampling to obtain feature maps, while we built four-branches structure with different dimension channels to segment the original image. Then, in the fusion unit, four feature maps are fused point by point and transmitted to the next convolution layer. Experiments show that the multi-branch input structure mentioned above can improves the system performance and saves the training time. In the experiment, we use the ADE20K dataset and the Cityscapes dataset, both of which are considered as high quality semantically segmented datasets.

Keywords: Semantic segmentation · Multi-branch input structure · Feature fusion · Early convolutional layer

1 Introduction

Image segmentation is the basic task of computer vision, which has two sub-problems: one is to predict the segmentation only at the class level, which marks one position of each pixel; the other is to distinguish individuals from different objects. According to different segmentation purposes, it can be divided into general segmentation, semantic segmentation and instance segmentation. Among them, general segmentation refers to the separation of pixel regions of different subordinates and different objects, such as the segmentation of foreground and background, the segmentation of dog's region and cat's region and background. Semantic segmentation is to classify the semantics of each area on the basis of general segmentation, such as pointing out the respective categories of each object in a picture. Instance segmentation is based on the semantic segmentation to set the number of each object, such as the number of dog A is set to 1, the number of dog B is set to 2. Semantic segmentation of images is a very important task in image understanding, whose goal is to classify every pixel in the image class label. Fast and accurate image segmentation can help solve many problems.

Neural networks, including fully convolution networks (FCNs) [1], AlexNet [2], Visual Geometry Group (VGG) [3] and now widely used Residual Neural Network (ResNet) [4], show outstanding performance in semantic segmentation. State-of-art

scenario analysis networks are based on FCN, which successfully use deep learning method to do semantic segmentation. AlexNet is a lightweight network structure consisting of five convolutional layers, a maximum pooling layer, Rectified Linear Units (ReLUs), three fully-connected layers and dropout. It uses fewer network layers, but achieves good results. The Conv1 to Conv5 of the semantics segmentation network consisting of all the above neural networks are convolution layers, followed by full-connected layers. Each convolution layer is downsampled by convolution and pooling. However, [5] and [6] propose to use deconvolution for up-sampling in each convolution layer, and [7] replaces all convolution layers with atrous convolution layers, so as to improve the receptive field index while keeping the number of parameters unchanged. Sequential extraction of feature information from conv1 to conv5 reduces the size of the image but increases the depth of the feature map, which makes the feature information more detailed. However, in the early stage of the network, because of the rapidly gradient decreases [8], the extracted feature map will be very rough. Therefore, we propose a multi-branch input structure for early convolution layers, that is, the original image is input into four branches at the same time, and the feature information is extracted separately. Then all the feature information needs to be fused and transmitted to the next layer. Experiments show that this method can improve the quality of rough semantic information of the network, accelerate the convergence speed of the deep network, and save computing time.

2 Related Work

In recent years, convolutional neural networks (CNNs) has become a very successful method in the field of semantic segmentation. FCN proposed an effective end-to-end training method, which has become the main choice in the field of semantic segmentation. In semantics segmentation, CNN mainly adopts the following basic operation modes: (1) down-sampling, up-sampling or the combination of the two; (2) multi-scale feature fusion, that is, adding features point by point or splicing feature channel dimensions; (3) getting a pixel-level segmentation image, that is, determining the category of each pixel. In fact, most of the network models in semantic segmentation follow the above methods, even complex networks like [9] adopt this basic method. By using atrous convolution, [10] enlarges the field of perception, so that it can grasp more global information of the image when the feature map is reduced to the same multiple. [11] scales the feature map to several different sizes, which looks like the multi-branch method, making the feature map have better global information and multi-scale information, which is very useful for improving accuracy. [6] uses the symmetrical structure of convolution network and deconvolution network, and uses both pooling and de-pooling. However, all the above networks adopt single-way operation, which can not aggregate multi-branch feature information. Therefore, we propose to adopt multi-branch input structure to process the input image in the early stage of the network, so as to extract more information. Semantic segmentation network has two ways of feature fusion: (1) point-by-point addition similar to FCN [1]; (2) channel dimension mosaic similar to U-Net [12]. The fusion unit of the network proposed in this paper adopts a point-by-point addition method similar to FCN, in order to reduce the

computational load and memory consumption. Through the multi-branch input structure, the whole system can be effectively trained end-to-end.

3 Method

In the field of neural networks, high-performance real-time model architecture is the goal of current researchers. Based on this goal, we propose a structure of multi-branch input in the early convolution layer, which enables the network system to effectively complete the training from low-level semantic information to high-level semantic information, thus improving the semantic information of low-level segmentation graph and shortening the training time.

3.1 Basic Network Model

As shown in Fig. 1, we input the original image into four branches to segment the image separately, and then fuse the segmented images and send it to the pooling layer. Our network model is based on PSPNet (Pyramid Scene Parsing Network), whose core contribution is a global pyramid pooling. PSPNet scales the feature map to several different sizes, so that the feature can contain better global and multi-scale information, which is very useful for improving accuracy. PSPNet network model architecture is based on the ResNet, which is a very useful network model. Resnet adds jump connections to all layers of a common network, and adds a shortcut to each two layers to form a residual module. Five residual modules can form a residual network when connected together. The residual modules is helpful to solve the problem of gradient disappearance and gradient explosion, ensuring good performance while training deeper network. Because too deep network structure will take too long time to train and consume too much GPU memory resources, we choose PSPNet-50 based on ResNet-50, which can already meet our needs.

As shown in Fig. 1, when the image is input into the multi-branch structure, a series of operations such as convolution, batch normalization, ReLU [13, 14] will be performed. After the maximum pooling of the four branches, the 1×1 convolution will increase the feature dimension to 256, so that the segmentation graphs of the four branches can be fused point by point. We built four independent branch units in the network, and the four feature maps are fused and output.

We use conventional convolution to preserve boundaries and details as much as possible. Convolution with size of 1×1 and maximization pooling are adopted in the network. The former is only for improving dimensionality, but not for downsampling. The latter is 3×3 in size and 2 in step, which can speed up the operation.

After the input image passes through the four branches, the network uses the method of point-by-point addition [1] to fuse the features, which improves the feature information of the shallow coarse segmentation maps, speeding up the convergence speed of the deep network and shortening the training time.

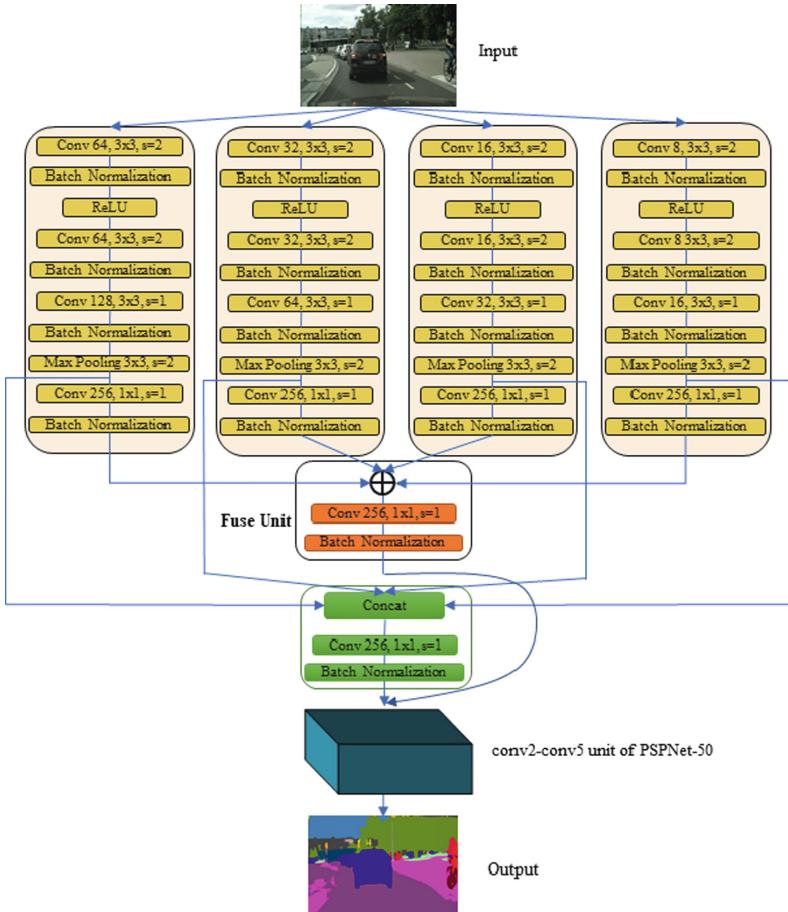


Fig. 1. The multi-branch input structure in the early stage of the network. The structure consists of four independent branches, each of which feeds back its segmentation features to concat. In the fusion unit, the segmentation graph of maximum pooling of each branch is fused and sent to the next layer of the neural network.

4 Experiment

We evaluated our network on two open datasets, including scene analysis dataset ADE20K [15] and urban scene understanding dataset Cityscapes [16]. Experiments show that our proposed method is very successful in scene parsing and semantics segmentation.

4.1 Implementation Details

In this section, we describe the details of the experiment and the parameter settings. Our experiment is based on the public platform TensorFlow, which is also one of the

most popular platforms at present. In experiment, we set the basic learning rate to 0.001, power to 0.9, momentum to 0.9, and weight attenuation to 0.0001. (MIoU) [17] is used as the standard for segmentation quality and performance is improved by increasing the number of iteration steps during the training process. These settings are designed to ensure that the network model is not over-fitting in training, and in order to improve the performance of the network, some special training techniques are needed. In the course of the experiment, we found that large “cropsize” can improve the performance of training, and larger “batchsize” in the Batch Normalization [18] layer can better improve the performance of training too. However, due to the physical limitation of GPU card, we can only set the “batchsize” to 6 during the training process. If the GPU memory is large enough or there are multiple GPU cards, you can set the “batchsize” to 16, 32 or larger, which will get better expectations.

4.2 Training Based on ADE20K Dataset

The ADE20K dataset is a newly released scene parsing dataset that provides 150 categories of dense labels and a variety of scenes with a total of 1038 image-level labels. In the whole database, 20K scene images are used for training, 2K for evaluation and 2K for testing. These categories include so wide variety of objects and stuff that are harder to parse than other datasets. The Table 1 is the experimental results of the ADE20K dataset without using the pre-training model. It can be seen that MIOU of multi-branch input structure is a little higher than that of PSPNet-50, and its training time is reduced by 19 h.

Table 1. Segmentation results on ADE20K dataset

| Model | Training time (h) | Training memory (M) | Val time (s) | Valuation memory (M) | MIoU (%) |
|--------------|-------------------|---------------------|--------------|----------------------|----------|
| Multi-branch | 24.5 | 15625 | 324 | 7675 | 0.0908 |
| PSPNet-50 | 43.58 | 15625 | 267 | 7659 | 0.0856 |

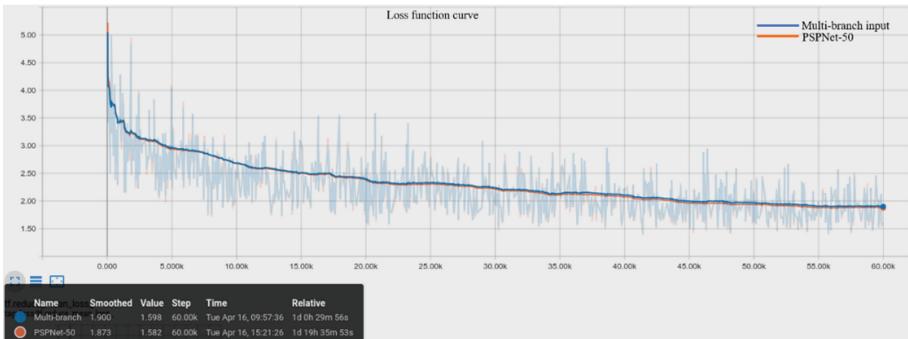


Fig. 2. Loss function curve on ADE20K dataset

As can be seen from Fig. 2, in each iteration, the loss of multi-branch input structure is almost the same as that of PSPNet-50, but it saves a lot of time and resources. Figure 3 shows the segmentation effect of multi-branch input structure and PSPNet-50.

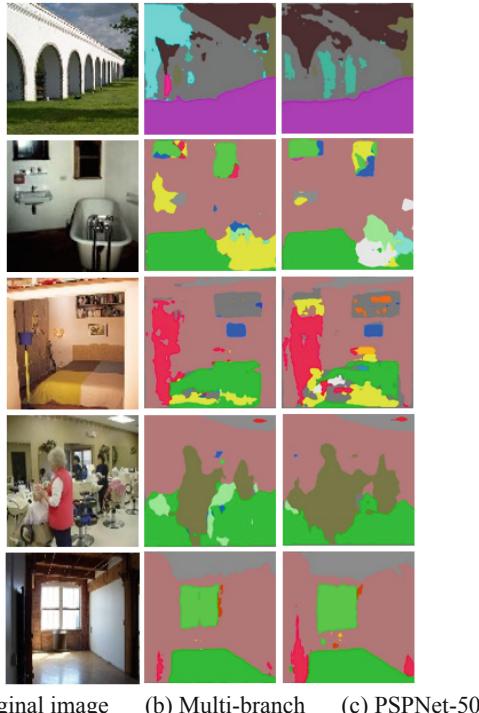


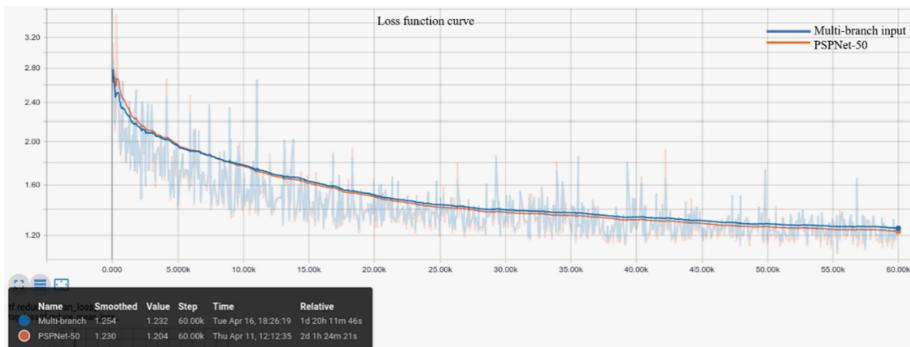
Fig. 3. Prediction examples on ADE20K dataset

4.3 Training Based on Cityscapes Dataset

Cityscapes is a very new dataset for urban road scene understanding which collects street images from 50 European cities in different seasons. It provides 5000 pixel-level finely annotated images, including roads, cars, pedestrians, bicycles, sky, and more. There are 2975 images for training, 500 images for evaluation and 1525 images for testing in this dataset. The Cityscapes dataset consists of a total of 19 categories, including objects and targets, which is much smaller than 150 classes of the ADE20K dataset, so it is easier to train and analyze. Table 2 shows the results of our experiments on Cityscapes dataset. It can be seen that the MIoU of multi-branch input structure transfer is a little higher than that of PSPNet-50, and the training time is also reduced by 5 h. It can also be seen in Fig. 4 that the loss of Multi-branch input structure is almost the same as that of PSPNet-50 in each iteration, but Multi-branch input structure saves computing time. Figure 5 shows the segmentation effect of multi-branch input structure and PSPNet-50.

Table 2. Segmentation results on Cityscapes

| Model | Training time (h) | Training memory (M) | Val time (s) | Valuation memory (M) | MIoU (%) |
|--------------|-------------------|---------------------|--------------|----------------------|----------|
| Multi-branch | 44.18 | 15625 | 347 | 7723 | 50.68 |
| PSPNet-50 | 49.04 | 15625 | 317 | 7675 | 50.53 |

**Fig. 4.** Loss function curve on the Cityscapes dataset**Fig. 5.** Prediction examples on Cityscapes dataset

5 Conclusion

In this paper, we propose a multi-branch input structure for complex scene analysis, which can improve the quality of shallow semantic information in early convolution layer, and also reduce the training time of the system. And we have improved the network model to make its application more real-time.

References

1. Long J, Shelhamer E, Darrell T (2017) Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell* 39:640–651
2. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: *Neural information processing systems*. Harrahs and Harveys
3. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In: *International conference on learning representations*, San Diego
4. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *The IEEE conference on computer vision and pattern recognition*, Las Vegas, no 1, pp 770–778
5. Badrinarayanan V, Kendall A, Cipolla R (2017) Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 39:2481–2495
6. Noh H, Hong S, Han B (2015) Learning deconvolution network for semantic segmentation. In: *The IEEE international conference on computer vision*, Santiago, Chile, pp 1520–1528
7. Yu F, Koltun V (2016) Multi-scale context aggregation by dilated convolutions. In: *International conference of learning representation*, San Juan, Puerto Rico
8. Ruder S (2016) An overview of gradient descent optimization algorithms. [arXiv:1609.04747](https://arxiv.org/abs/1609.04747)
9. Chen L-C, Papandreou G, Schroff F, Adam H (2017) Rethinking Atrous Convolution for Semantic Image Segmentation. [arXiv:1706.05587](https://arxiv.org/abs/1706.05587)
10. Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2018) DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans Pattern Anal Mach Intell* 40:834–848
11. Zhao H, Shi J, Qi X, Wang X, Jia J (2017) Pyramid scene parsing network. In: *The IEEE conference on computer vision and pattern recognition*, Honolulu, Hawaii, pp 2881–2890
12. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: *Medical image computing and computer-assisted intervention*, Munich, Germany, pp 234–241
13. Jarrett K, Kavukcuoglu K, Ranzato M, LeCun Y (2009) What is the best multi-stage architecture for object recognition? In: *IEEE 12th international conference on computer vision*, Kyoto, Japan
14. He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In: *The IEEE international conference on computer vision*, Santiago, Chile, pp 1026–1034
15. Zhou B, Zhao H, Puig X, Fidler S, Barriuso A, Torralba A (2017) Scene parsing through ADE20K dataset. In: *The IEEE conference on computer vision and pattern recognition*, Honolulu, HI, pp 633–641

16. Cordts M, Omran M, Ramos S, Rehfeld T, Enzweiler M, Benenson R, Franke U, Roth S, Schiele, B (2016) The cityscapes dataset for semantic urban scene understanding. In: The IEEE conference on computer vision and pattern recognition, Las Vegas, pp 3213–3223
17. Everingham M, Van Gool L, Williams CKI, Winn J, Zisserman A (2010) The Pascal visual object classes (VOC) challenge. *Int J Comput Vis* (88):303–338
18. Lin G, Milan A, Shen C, Reid I (2017) RefineNet: multi-path refinement networks for high-resolution semantic segmentation. In: The IEEE conference on computer vision and pattern recognition, Honolulu, HI, pp 1925–1934

Author Index

A

Ai, Yongle, 31

B

Bai, Jing, 233
Bian, Jilong, 651
Bian, Peng, 41, 83

C

Cao, Sijia, 270
Chen, Jiannan, 601
Chen, Jin, 159, 724
Chen, Qiang, 260
Chen, Qingqing, 117
Chen, Sai, 75
Chen, Xi, 582
Chen, Xianzhong, 520
Chen, Xin, 192
Chen, Xinlong, 260
Chen, Ya-nan, 621
Chen, Zengqiang, 75, 107
Chen, Zheyi, 41
Cheng, Hangyang, 183
Cheng, Lan, 660
Cheng, Pan, 429
Cui, Leilei, 713
Cui, Wanqiu, 670
Cui, Yan, 589

D

Dai, Fengzhi, 24
Ding, Nan, 610
Ding, Renqiang, 183
Du, Junping, 601, 670

Du, Xianjin, 370

Duan, Huida, 135

F

Fan, Yimin, 610
Fang, Shuang, 286
Fang, Yisheng, 513
Feng, Xie, 703
Fu, Xuesong, 504
Fu, Ziyi, 58

G

Gao, Caixia, 58
Gao, Haitao, 107
Gao, Lin, 689
Gao, Xiaofan, 324
Gao, Xuehui, 679
Ge, Yunsong, 573
Gong, Shenjian, 66
Guan, Sheng, 159
Guo, Jianming, 10
Guo, Lin, 75
Guo, Qiaochu, 270
Guo, Tan, 201
Guo, Xuemei, 243
Guo, Yongqi, 504

H

Han, Wei, 359
Hou, Ao, 641
Hou, Junyi, 66
Hou, Qingwen, 520
Hu, Chenyang, 446
Hu, Daoqing, 582

Hu, Zhongzhi, 91, 100, 254
 Huang, Fengguang, 582
 Huang, Ning, 563
 Huang, Xu, 243
 Huang, Zhuang, 295

J

Ji, Feng, 168
 Jia, Hongchen, 75
 Jia, Zhiqiang, 314
 Jiang, Min, 223
 Jiang, Yunxia, 117
 Jiang, Zhiguo, 306, 314
 Jie, Dangyang, 610
 Jin, Wemma, 461
 Jing, Xuedong, 513, 621

K

Kou, Feifei, 601, 670

L

Li, Bosen, 582
 Li, Bowen, 481
 Li, Changdi, 66
 Li, Chenglin, 18
 Li, Da, 343
 Li, Fengyi, 388
 Li, Hezheng, 233
 Li, Hua, 490
 Li, Hui, 388
 Li, Jinfeng, 651
 Li, Kaikai, 159
 Li, Lin, 343, 370, 419
 Li, Qinghua, 589
 Li, Shuai, 31
 Li, Wenyan, 306
 Li, Xiaobin, 279, 429
 Li, Xue, 270
 Li, Xueqi, 10
 Li, Yanxi, 233
 Li, Yunlong, 324
 Li, Zhonghua, 243
 Liang, Yueqian, 306, 314
 Lin, Wei, 437
 Liu, Hongyu, 388
 Liu, Jianfei, 632
 Liu, Jiwei, 632
 Liu, Quanpan, 550
 Liu, Qunfeng, 31
 Liu, Shengyi, 254
 Liu, Weimin, 254
 Liu, Xiangsheng, 1, 117, 175

Liu, Yang, 520, 610
 Liu, Ying, 100
 Liu, Yunxiao, 91
 Lu, Fengshun, 223, 334
 Lu, Lixin, 143
 Lu, Na, 41, 83, 286
 Lu, Qi, 582
 Luo, Qiang, 542
 Luo, Yansha, 582

M

Ma, Qinghua, 689
 Ma, Tao, 279
 Ma, Xianghua, 286
 Mao, Ronghu, 632
 Meng, Yue, 520

N

Ni, Zihang, 660
 Nie, Yanjie, 58

P

Pang, Yufei, 223, 334
 Peng, Jiansheng, 201
 Peng, Kaixiang, 192
 Peng, Xin, 135

Q

Qi, Long, 223, 334, 359
 Qin, Yahang, 201
 Qiu, Kepeng, 504
 Qu, Jiarui, 286

R

Ren, Di, 41
 Ren, Jili, 641
 Ren, Yajing, 260

S

Shao, Shuai, 437
 Shen, Zhe, 214
 Shi, Lei, 601
 Song, Qi, 260
 Song, Yue, 18
 Song, Yunzhong, 24
 Sun, Mingwei, 107
 Sun, Qian, 214
 Sun, Qinglin, 75
 Sun, Shiwen, 407

T

Tang, Xiaoping, 201
 Tian, Lin, 582

Tian, Shu, 324

Tu, Shuguang, 446

W

Wang, Bulai, 50, 83, 175

Wang, Chaoli, 352, 397

Wang, Chuanya, 724

Wang, Fuzhong, 641

Wang, Guoli, 243

Wang, Hao, 407

Wang, Huan, 550

Wang, Jia, 470

Wang, Jian, 529

Wang, Jianlin, 504

Wang, Jing, 632

Wang, Jinhai, 456

Wang, Jiqiang, 91, 100, 254

Wang, Li, 214, 407

Wang, Lin, 490

Wang, Longhong, 573

Wang, Mingren, 713

Wang, Qi, 456

Wang, Qingbin, 306, 314

Wang, Rutong, 504

Wang, Suda, 18

Wang, Wei, 582

Wang, Weihong, 470

Wang, Xiangfeng, 314

Wang, Xiaoyi, 214

Wang, Yafei, 446

Wang, Zhengjie, 550

Wang, Ziteng, 306, 513

Wang, Ziyi, 419

Wei, Wei, 695

Wu, Jiangxue, 520

Wu, Jiwei, 10

Wu, Qian, 470

Wu, Qiuzhen, 1

Wu, Xiangwei, 563

X

Xia, Chengyi, 407

Xiang, Zhu, 703

Xiao, Huimin, 24

Xiao, Sumei, 334

Xiao, Zhongyun, 359

Xing, Xiaolin, 582

Xu, Bing, 41

Xu, Jiping, 214

Xu, John, 168

Xu, Kan, 563

Xu, Min, 397

Xu, Ming, 542

Xu, Qiangqiang, 50, 83

Xu, Yuneng, 446

Xu, Zhengguang, 126

Xu, Zhishu, 270

Xue, Yuquan, 621

Xue, Zhe, 601, 670

Y

Yan, Fei, 352

Yan, Lutao, 456

Yang, Jianjun, 295

Yang, Meiling, 126

Yang, Shujun, 689

Yang, Yuanyuan, 1, 41, 50, 83, 117

Yao, Kanzhong, 151, 379

Yin, Zhide, 388

Yu, Jiabin, 214

Yu, Lei, 66

Yu, Rongchuan, 201

Yu, Tianyang, 279, 429

Yuan, Xunpu, 670

Yuan, Yuyang, 175

Yue, Zongshuai, 270

Z

Zeng, Xu, 481, 573

Zhai, Xiaoman, 582

Zhang, Bin, 456

Zhang, Hongwei, 713

Zhang, Huiyan, 214

Zhang, Jing, 660

Zhang, Jinheng, 660

Zhang, Mengtian, 324

Zhang, Qianyao, 243

Zhang, Qiao, 18

Zhang, Sen, 529

Zhang, Sijing, 286

Zhang, Weicun, 143

Zhang, Yan, 651

Zhang, Zhiyuan, 695

Zhao, Liqiang, 504

Zhao, Xin, 286

Zhao, Xuedong, 481, 573

Zhao, Yanjie, 306, 314

Zhao, Yuqi, 159

Zhao, Zhiyao, 214

Zhao, Zihe, 481

Zheng, Chaohang, 490

Zheng, Jianqiang, 689
Zheng, Yaru, 542
Zheng, Yu, 481, 573
Zhong, Jiale, 175
Zhou, Jiaomei, 334
Zhou, Pei, 679

Zhou, Wuneng, 183
Zhou, Zhengxin, 1, 117
Zhu, Mengping, 260
Zhu, Xukun, 1
Zhu, Yanjuan, 151, 379
Zuo, Min, 695