

LNAI 6425

Honghai Liu  
Han Ding  
Zhenhua Xiong  
Xiangyang Zhu (Eds.)

# Intelligent Robotics and Applications

Third International Conference, ICIRA 2010  
Shanghai, China, November 2010  
Proceedings, Part II

2  
Part II

Lecture Notes in Artificial Intelligence 6425

Edited by R. Goebel, J. Siekmann, and W. Wahlster

Subseries of Lecture Notes in Computer Science

Honghai Liu Han Ding Zhenhua Xiong  
Xiangyang Zhu (Eds.)

# Intelligent Robotics and Applications

Third International Conference, ICIRA 2010  
Shanghai, China, November 10-12, 2010  
Proceedings, Part II

**Series Editors**

Randy Goebel, University of Alberta, Edmonton, Canada

Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Wolfgang Wahlster, DFKI and University of Saarland, Saarbrücken, Germany

**Volume Editors**

Honghai Liu

The University of Portsmouth, School of Creative Technologies

Portsmouth PO1 2DJ, UK

E-mail: honghai.liu@port.ac.uk

Han Ding

Shanghai Jiao Tong University, Robotics Institute

Shanghai 200240, China

E-mail: hding@sjtu.edu.cn

Zhenhua Xiong

Shanghai Jiao Tong University, Robotics Institute

Shanghai 200240, China

E-mail: mexiong@sjtu.edu.cn

Xiangyang Zhu

Shanghai Jiao Tong University, Robotics Institute

Shanghai 200240, China

E-mail: mexyzhu@sjtu.edu.cn

Library of Congress Control Number: 2010936828

CR Subject Classification (1998): I.4, I.5, I.2, I.2.10, H.4, C.2

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743

ISBN-10 3-642-16586-9 Springer Berlin Heidelberg New York

ISBN-13 978-3-642-16586-3 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

[springer.com](http://springer.com)

© Springer-Verlag Berlin Heidelberg 2010  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India  
Printed on acid-free paper 06/3180

# Preface

The market demand for skills, knowledge and adaptability have positioned robotics to be an important field in both engineering and science. One of the most highly visible applications of robotics has been the robotic automation of many industrial tasks in factories. In the future, a new era will come in which we will see a greater success for robotics in non-industrial environments. In order to anticipate a wider deployment of intelligent and autonomous robots for tasks such as manufacturing, healthcare, entertainment, search and rescue, surveillance, exploration, and security missions, it is essential to push the frontier of robotics into a new dimension, one in which motion and intelligence play equally important roles.

The 2010 International Conference on Intelligent Robotics and Applications (ICIRA 2010) was held in Shanghai, China, November 10–12, 2010. The theme of the conference was “Robotics Harmonizing Life,” a theme that reflects the ever-growing interest in research, development and applications in the dynamic and exciting areas of intelligent robotics. These volumes of Springer’s *Lecture Notes in Artificial Intelligence* and *Lecture Notes in Computer Science* contain 140 high-quality papers, which were selected at least for the papers in general sessions, with a 62% acceptance rate. Traditionally, ICIRA 2010 holds a series of plenary talks, and we were fortunate to have two such keynote speakers who shared their expertise with us in diverse topic areas spanning the range of intelligent robotics and application activities.

We would like to thank the International Advisory Committee for their guidance and advice. We are also especially grateful to the Program Committee members for their rigorous and efficient reviewing of the submitted papers, and the Organizing Committee for their enormous efforts and excellent work. In addition, we greatly appreciate the Ministry of Education, National Natural Science Foundation of China, Science and Technology Commission of Shanghai Municipality, Shanghai Jiao Tong University, State Key Laboratory of Mechanical System and Vibration, State Key Lab of Digital Manufacturing Equipment & Technology, and Springer for their support.

It has been our pleasure to organize the 2010 International Conference on Intelligent Robotics and Applications. We hope it is your pleasure to read these selected papers by some of the highest-quality researchers and practitioners in the world of intelligent robotics. We also hope that they provide inspiration for your own research and help advance the state of the art in the multidisciplinary field of robotics and its application.

August 2010

Xiangyang Zhu  
Han Ding  
Zhenhua Xiong  
Honghai Liu

# Conference Organization

## International Advisory Committee

Tamio Arai	University of Tokyo, Japan
Hegao Cai	Harbin Institute of Technology, China
Toshio Fukuda	Nagoya University, Japan
Huosheng Hu	Essex University, UK
Oussama Khatib	Stanford University, USA
Jurgen Leopold	Fraunhofer IWU, Germany
Ming Li	National Natural Science Foundation of China
Peter Luh	Connecticut University, USA
Jun Ni	University of Michigan, USA
Nikhil R. Pal	Indian Statistical Institute, India
Grigory Panovko	Russian Academy of Sciences, Russia
Shigeki Sugano	Waseda University, Japan
Michael Wang	Chinese University of Hong Kong, China
Bogdan M. Wilamowski	Auburn University, USA
Ming Xie	Nanyang Technological University, Singapore
Lotfi Zadeh	University of California, Berkeley, USA

## General Chairs

Zhongqin Lin	Shanghai Jiao Tong University, China
Youlun Xiong	Huazhong University of Science and Technology, China

## Program Chairs

Xiangyang Zhu	Shanghai Jiao Tong University, China
Honghai Liu	University of Portsmouth, UK

## Organizing Committee Chairs

Han Ding	Shanghai Jiao Tong University, China
Zexiang Li	Hong Kong University of Science & Technology, China
Hong Liu	Harbin Institute of Technology, China
Guang Meng	China Aerospace Science & Technology Co., China
Guobiao Wang	National Natural Science Foundation of China
Tianmiao Wang	Beihang University, China

## Area Chairs

H. Levent Akin	Bogazici University, Turkey
Kaspar Althoefer	King's College London, UK
Shuzhi Sam Ge	National University of Singapore, Singapore
Qiang Huang	Beijing Institute of Technology, China
Naoyuki Kubota	Tokyo Metropolitan University, Japan
Chunyi Su	Concordia University, Canada
Jun Wang	Chinese University of Hong Kong, China
Diedrich Wolter	Bremen University, Germany
Bin Yao	Purdue University, USA
Jie Zhao	Harbin Institute of Technology, China

## Tutorial/Workshop Chairs

Ryad Chellali	Institute of Italy Technologies, Italy
Qixin Cao	Shanghai Jiao Tong University, China

## Poster Session Chairs

James Hu	Kumamoto University, Japan
Min Lei	Shanghai Jiao Tong University, China

## Publicity Chairs

Mohamed Kamel	University of Waterloo, Canada
Yun Luo	Shanghai Jiao Tong University, China

## Publication Chairs

Rainer Palm	Orebro University, Sweden
Yanzheng Zhao	Shanghai Jiao Tong University, China

## Financial Chair

Zhenhua Xiong	Shanghai Jiao Tong University, China
---------------	--------------------------------------

## Registration Chair

Dingguo Zhang	Shanghai Jiao Tong University, China
---------------	--------------------------------------

## Local Chair

Lifeng Xi	Shanghai Jiao Tong University, China
-----------	--------------------------------------

## General Affairs and Secretariat

Xinjun Sheng

Shanghai Jiao Tong University, China

## Acknowledgment

Ministry of Education, China

National Natural Science Foundation of China

Science and Technology Commission of Shanghai Municipality, China

Shanghai Jiao Tong University, China

State Key Laboratory of Mechanical System and Vibration, China

State Key Lab of Digital Manufacturing Equipment & Technology, China

## List of Reviewers

We would like to acknowledge the support of the following people who contributed to peer review of articles from ICIRA 2010.

Absi Alfaro	Wen-Jie Chen	Nicola Ferrier
Cihan Acar	Xiao-Peng Chen	Kevin Fite
Mojtaba Ahmadi	You-Hua Chen	Zhuang Fu
Eric L Akers	Sonia Chernova	Antonios Gasteratos
H. Levent Akin	Yu-Zheng Chong	Shu-Zhi Sam Ge
Berkant Akin	Henrik Christensen	Akio Gofuku
Mohamed Al Marzouqi	Dimitrios Chrysostomou	Hu Gong
Redwan Alqasemi	Sadek Crisóstomo	Darwin Gouwanda
Kaspar Althoefer	Xavier Cufí	Dong-Bing Gu
Farshid Amirabdollahian	Ravinder Dahiya	En-Guang Guan
Cecilio Angulo	Konstantinos Dalamagkidis	Li-Wen Guan
Panagiotis Artemiadis	Greyson Daugherty	Lei Guo
Fernando Auat Cheein	Raoul de Charette	Wei-Zhong Guo
Joonbum Bae	Joris De Schutter	Alvaro Gutierrez
Feng Bai	Liyange De Silva	Norihiro Hagita
Malek Baklouti	Chandratilak	Hassan Haleh
Sven Behnke	Hua Deng	Tatsuya Harada
Nicola Bellotto	Jean-Yves Didier	Masaki Hasegawa
Roger Bostelman	Ming Ding	Kenji Hashimoto
Magdalena Bugajska	Hao Ding	Mitsuhiro Hayashibe
Darius Burschka	Can Ulas Dogruer	Vincent Hayward
Li-Xin Cao	Zhen-Chun Du	H. Hellendoorn
Jiang-Tao Cao	Damien Eynard	Patrick Henaff
Qi-Xin Cao	Bao-Jie Fan	Sophie Hennequin
Zhi-Qiang Cao	Hui-Jie Fan	Rafael Herrejon
Barbara Caputo	Janabi-Sharifi Farrokhi	Bernard Voon Ee How
Guillaume Caron	Ying Feng	Ming-Hui Hu
Dian-Sheng Chen	Manuel Fernandez-Carmona	Chuan-Zhen Huang

Yong-An Huang	Yun Luo	Ching-Hua Ting
Cong-Hui Huang	Xiang Luo	Michael John Tribou
Atsushi Imiya	Jie Ma	Ching-Chih Tsai
Xiao-Fei Ji	Li Ma	Cristina Urdiales
Li Jiang	Dominic Maestas	Akira Utsumi
Zhao-Jie Ju	Elmar Mair	Kalyana Veluvolu
Agung Julius	Adrian Martin	Qi-Yuan Wang
Takahiro Kagawa	Emanuele Menegatti	Furui Wang
Chen-Yu Kai	Krzysztof Mianowski	Zongyao Wang
Shun-Ichi Kaneko	Huasong Min	Zhenyan Wang
Kikuhi Kawasue	Xie Ming	Hao Wang
Erdal Kayacan	Seyed Mohamed Buhari	Guowu Wei
Chong-Hui Kim	Mohamed Ismail	Ang Wei Tech
Eun-Young Kim	Hyungpil Moon	Myra Wilson
Sang-Ho Kim	Behzad Moslehi	Hong-Tao Wu
Chyon-Hae Kim	Josiah L. Munda	Xiao-Jun Wu
Won-Jong Kim	Hyun Myung	Chao Wu
Denis Klimentjew	Jawahar N.	Xian-Bo Xiang
Akio Kosaka	Hiroyuki Nakamoto	Cai-Hua Xiong
Naoyuki Kubota	Duc Dung Nguyen	Ji-Jie Xu
Ki-Ju Lee	Hirotaka Osawa	Xi-Peng Xu
Bengt Lennartson	Mohammadreza Asghari	Xin Xu
Ales Leonardis	Oskoei	Zhao-Hong Xu
Gang Li	Nuno Otero	Bing Xu
Qin-Chuan Li	Chee Khiang Pang	Seiji Yamada
You-Fu Li	gaoliang Peng	Yong-Hua Yan
Yang-Min Li	Fernando Pereira	Wen-Yu Yang
Zhi-Jun Li	Anton Satria Prabuwono	Yue-Hong Yin
Hai-Zhou Li	Flavio Prieto	Ding-Guo Zhang
Bin Li	Yue Qiu	Jian-Guo Zhang
Geng Liang	M. Jayedur Rashid	Wen-Zeng Zhang
Yong-Cheng Lin	Florian Raudies	Xu-Guang Zhang
Chen-Liang Liu	Jorge Rivera-Rovelo	Yu-Ru Zhang
Hai-Ning Liu	Debanik Roy	Xian-Min Zhang
Hong-Hai Liu	Alexander Schmitz	Yu-Nong Zhang
Jin-Dong Liu	Desire Sidibe	Yan-jiang Zhao
Jia Liu	Ponnambalam Sivalinga	Yan-Zheng Zhao
Bing-Bing Liu	Kai-Tai Song	Yong Zhao
Wei Liu	Chun-Yi Su	Hui-Yu Zhou
Ping Liu	Anan Suebsomran	Li-Min Zhu
Chao Liu	Yue-Gang Tan	Chi Zhu
Xin-Jun Liu	Da-Wei Tan	Chun-Gang Zhuang
Yong-Huai Liu	Hideyuki Tanaka	Jun Zou
Benny Lo	Bo Tao	Wei Zou
Yi Lu	Michael Thomason	

## Table of Contents – Part II

### Industrial Process Control, Manufacturing Processes, and Automation

A Behavior-Based Path Planning for Autonomous Vehicle <i>CaiJing Xiu and Hui Chen</i>	1
An Uncertainty Approach for Fixture Layout Optimization Using Monte Carlo Method <i>Xiaoping Zhang, Wenyu Yang, and Miao Li</i>	10
The Small-World Network Model of Mixed Production Line <i>Ting Yang, Dinghua Zhang, Bing Chen, and Shan Li</i>	22
Distributed Cooperative Control of Industrial Robotic Systems Using Petri Net Based Multitask Processing <i>Gen'ichi Yasuda</i>	32
A Graph Based Hybrid Approach of Offline Pre-planning and Online Re-planning for Efficient Assembly under Realtime Constraints <i>Daniel Ewert, Sebastian Thelen, Ralph Kunze, Marcel Mayer, Daniel Schilberg, and Sabina Jeschke</i>	44
Improved Inertial Pose Estimation Algorithms for On-Site Hydraulic Turbine Blade Repairing Robot <i>Xiande Ma, Qiang Chen, Zhenguo Sun, and Wenzeng Zhang</i>	56
Tool Path Generation Based on BCELTP for Maximizing Machining Strip Width <i>Hu Gong</i>	67

### Mobile Robotics

Dual-Quaternion-Based Variable Structure Control: A New Approach and Application <i>Hui Zhang, Xiangke Wang, and Dapeng Han</i>	75
A 3-D Simulation of Unicycle Robot Based on Virtual Prototyping Technology <i>Xiaogang Ruan, Qiyuan Wang, and Naigong Yu</i>	87
A Study of the Electric Wheelchair Hands-Free Safety Control System Using the Surface-Electromyogram of Facial Muscles <i>Hiroki Tamura, Takao Manabe, Takafumi Goto, Yuki Yamashita, and Koichi Tanno</i>	97

A New Triple-Based Multi-robot System Architecture and Application in Soccer Robots .....	105
<i>Xiangke Wang, Hui Zhang, Huimin Lu, and Zhiqiang Zheng</i>	
Pursuit Evasion in Dynamic Environments with Visibility Constraints .....	116
<i>Ibrahim Al-Bluwi and Ashraf Elnagar</i>	
Wall Following with a Single Ultrasonic Sensor .....	130
<i>Shérine M. Antoun and Phillip J. McKerrow</i>	
Kinematics Modeling and Analysis of A Five-Wheeled Rover with Caster and Camber .....	142
<i>He Xu and Zhenyu Zhang</i>	
Experimental Investigation of a Prediction Algorithm for an Indoor SLAM Platform .....	154
<i>Jung-Fu Hou, Yu-Shin Chou, Yau-Zen Chang, and Jing-Sin Liu</i>	
Trajectory Estimation of a Skid-Steering Mobile Robot Incorporating Free Wheels .....	166
<i>Tokuji Okada, Abeer Mahmoud, and Wagner Votelho</i>	
Kinematics Modeling and Analyses of an Articulated Robot .....	180
<i>Yang Zhang, Xinhui Liu, and Tongjian Wang</i>	
A Tracking Robot Based on Wireless Beacon .....	191
<i>Zhou Fang, Lei Deng, Yongsheng Ou, and Xinyu Wu</i>	
A Class of Biped Mechanisms with Three Links and Two Joints .....	203
<i>Chang-Huan Liu, Yao-bin Tian, Xi-ming Wang, and Yan-an Yao</i>	
Time Efficient Strategy to Explore Unknown Indoor Environments by Mobile Ground Robots .....	215
<i>Mirco Alpen, Stephan Stuedemann, and Joachim Horn</i>	
Overhead High-Voltage Transmission Line Deicing Robot System and Experiment Study .....	227
<i>Yucheng Bai, Gongping Wu, hua Xiao, yu Yan, Yingsong Li, and Xingwei Fu</i>	
<b>Theory and Applications of Heavy Manipulators</b>	
Projection on Convex Set and Its Application in Testing Force Closure Properties of Robotic Grasping .....	240
<i>Miao Li, Wenyu Yang, and Xiaoping Zhang</i>	
Buffering Indices for Robot Manipulators Based on the Energy Distribution .....	252
<i>Hao Wang, Zhongqin Lin, Kai Zhao, and Genliang Chen</i>	

Robust Robotic Grasping Force Optimization with Uncertainty .....	264
<i>Wenyu Yang, Miao Li, and Xiaoping Zhang</i>	
Analysis of Forging Compliance Process and Design of the Forging Simulator .....	276
<i>Pu Zhang, Zhenqiang Yao, and Zhengchun Du</i>	
Field Calibration Method of Camera for Measuring the Dimensions of Large Forging .....	285
<i>Wei Liu, Jian Du, Bangguo Wang, Xinghua Jia, Shuangjun Liu, and Zhenyuan Jia</i>	
Velocity Control of the Horizontal Buffer System for Heavy Load Forging Manipulator .....	296
<i>Jiaoyi Hou, Hua Zhou, Jun Zou, and Xin Fu</i>	
Performance Optimization of Forging Manipulator during the Whole Forging Stroke .....	305
<i>Deshi Liu, Gang Li, Xinglin Guo, Yonggang Shang, and Dehao Liu</i>	
Structural Optimization for Wall Frame Design of a Forging Manipulator .....	317
<i>Kai Zhang, Gang Li, and Zhan Kang</i>	
Gripping Force Model and Influence Factors of Gripping Device for Heavy Forging Manipulators.....	329
<i>Lixin Yang, Hua Deng, and Zhen Xu</i>	
A Contact/Impact Analysis of Rigid Bodies Based on Newmark Direct Integration Method .....	339
<i>Zhen Xu and Hua Deng</i>	
RFID Enabled Remote Structural Health Monitoring for Heavy Lifting Appliances .....	345
<i>Xubing Chen and Fan Zhou</i>	
<b>Robot Vision</b>	
Compact Robot Vision for Moving Points .....	357
<i>Kikuhiro Kawasue, Satoshi Nagatomo, Masahiro Nishiyama, and Yuki Kawano</i>	
Motion Compensation Algorithm Based on Color Orientation Codes and Covariance Matching .....	368
<i>Xuguang Zhang, Weili Ding, and Peifeng Niu</i>	
An Advanced Algorithm for Illumination-Based Synchronization of High-Speed Vision Sensors in Dynamic Scenes .....	378
<i>Lei Hou, Shingo Kagami, and Koichi Hashimoto</i>	

## Biomedical and Rehabilitation Engineering, and Medical Robotics

A Method to Control Ankle Exoskeleton with Surface Electromyography Signals .....	390
<i>Zhen Zhang, Jiaxin Jiang, Liling Peng, and Hongchao Fan</i>	
A Resistance Compensation Control Algorithm for a Cable-Driven Hand Exoskeleton for Motor Function Rehabilitation .....	398
<i>Shuang Wang, Jiting Li, and Ruoyin Zheng</i>	
Kinematic Modeling of Bevel Tip Flexible Needle .....	405
<i>Yong-de Zhang and Yan-jiang Zhao</i>	
An Adaptive RBF Neural Network Control Strategy for Lower Limb Rehabilitation Robot .....	417
<i>Feng Zhang, Pengfeng Li, Zengguang Hou, Xiaoliang Xie, Xixiong Chen, Qingling Li, and Min Tan</i>	
Model Complexity Reduction of the Hand Musculoskeletal System in Tremulous Motion .....	428
<i>Ruwadi, Shee Cheng Yap, Philippe Poignet, and Ang Wei Tech</i>	
A Wearable Rehabilitation Robotic Hand Driven by PM-TS Actuators .....	440
<i>Jun Wu, Jian Huang, Yongji Wang, and Kexin Xing</i>	
A Cooperated-Robot Arm Used for Rehabilitation Treatment with Hybrid Impedance Control Method .....	451
<i>Jingguo Wang and Yangmin Li</i>	
Characteristics of the Robotic Arm of a 9-DoF Upper Limb Rehabilitation Robot Powered by Pneumatic Muscles .....	463
<i>Xianzhi Jiang, Caihua Xiong, Ronglei Sun, and Youlun Xiong</i>	
Optics Based Motion Measurement for a Catheter Navigation System: A Novel and Low Cost Approach .....	475
<i>Yonghua Yan, Daguo Chen, and Hang Yin</i>	

## Human-Friendly Robot Partners

Visual-Audio Integration for User Authentication System of Partner Robots .....	486
<i>Jiangtao Cao, Naoyuki Kubota, Ping Li, and Honghai Liu</i>	
Synthetic Finger Phalanx with Lifelike Skin Compliance .....	498
<i>John-John Cabibihan and Shuzhi Sam Ge</i>	

Study on Educational Application of Android Robot SAYA: Field Trial and Evaluation at Elementary School .....	505
---	-----

*Tskuya Hashimoto, Naoki Kato, and Hiroshi Kobayashi*

Essential Technology for Team Teaching Using Robot Partners .....	517
---	-----

*Akihiro Yorita and Naoyuki Kubota*

The Style of Information Service by Robot Partners .....	529
--	-----

*Aiko Yaguchi and Naoyuki Kubota*

## Surface Inspection, Modeling and Machining

Improved Full-Discretization Method for Milling Chatter Stability Prediction with Multiple Delays .....	541
---	-----

*XiaoJian Zhang, CaiHua Xiong, and Ye Ding*

An Algorithm to Generate Compact Dual NURBS Tool Path with Equal Distance for 5-Axis NC Machining .....	553
---	-----

*Qingzhen Bi, Yuhan Wang, Limin Zhu, and Han Ding*

Prediction of Cutting Forces Integrated Run-Out Effect for Five-Axis Peripheral Milling with a Cylindrical Cutter .....	565
---	-----

*Qiang Guo, Yuwen Sun, Feifei Xu, and Zhenyuan Jia*

Measurement Error Compensation Using Data Fusion Technique for Laser Scanner on AACMMs .....	576
--	-----

*Licheng Wang, Han Ding, and Shaojun Wang*

## Modelling for Robotics and Manufacturing Automation

Evaluating Q-Learning Policies for Multi-objective Foraging Task in a Multi-agent Environment .....	587
---	-----

*M. Yugeswaran and S.G. Ponnambalam*

A Hybrid GA-AIS Heuristic for Optimization of Multipass Turning Operations .....	599
--	-----

*Lee Yi Zheng and S.G. Ponnambalam*

Robot Assisted Stroke Rehabilitation: Joint Torque/Force Conversion from EMG Using SA Process .....	612
---	-----

*S. Parasuraman, Arif Wicaksono Oyong, and Veronica Lestari Jauw*

Genetic Goal Oriented Path Planning Algorithm for Acute Obstacle Avoidance in Mobile Robot Navigation .....	624
---	-----

*Soh Chin Yun, S. Parasuraman, and V. Ganapathy*

Services Integration Framework for Vehicle Telematics .....	636
---	-----

*S. Veera Ragavan, S.G. Ponnambalam, Velappa Ganapathy, and Joshua Teh*

Manipulating a Micro Stream by ‘Hydro Tweezers’ for Integration of Nanodevices .....	649
<i>Mei Liu, Yan Peng, Qiuquan Guo, Jun Luo, and Jun Yang</i>	
Multi-modality – EMG and Visual Based Hands-Free Control of an Intelligent Wheelchair .....	659
<i>Lai Wei and Huosheng Hu</i>	
<b>Robot and Investigation</b>	
A Novel Hand-Gesture Recognition Method Based on Finger State Projection for Control of Robotic Hands .....	671
<i>Wenzhen Yuan and Wenzeng Zhang</i>	
Adaptive Compliance Control for Collision-Tolerant Robot Arm with Viscoelastic Trunk .....	683
<i>Chungwei Chin, Yanan Li, Shuzhi Sam Ge, and John-John Cabibihan</i>	
CAN-Bus Based Distributed Control System for Hydraulic Turbine Blade Repairing Robot .....	695
<i>Kangtai Xiang, Zhenguo Sun, Hongjun Dai, Qiang Chen, and Jiajun Liu</i>	
Visual Assistance to an Advanced Mechatronic Platform for Pick and Place Tasks .....	705
<i>Rigas Kouskouridas and Antonios Gasteratos</i>	
An Approach to Multi-robot Site Exploration Based on Principles of Self-organisation .....	717
<i>Lyuba Alboul, Hussein S. Abdul-Rahman, Paul S. Haynes, Jacques Penders, and Julien Tharin</i>	
Analysis of Dynamic Characteristics of Single Piston Hydraulic Free Piton Engine .....	730
<i>Haoling Ren, Haibo Xie, and Huayong Yang</i>	
Carbon-Based Nanostructured Coatings on NiTi Shape Memory Alloy for Biomedical Applications .....	742
<i>Takanori Takeno, Hiroyuki Shiota, Hiroyuki Miki, Toshiyuki Takagi, and Yun Luo</i>	
Closed-Loop Stiffness Modeling and Stiffness Performance Analysis for Multi-axis Process System .....	754
<i>Rong Yan, Fangyu Peng, Wei Chen, Jianzhong Yang, and Linsen</i>	
<b>Author Index .....</b>	767

# A Behavior-Based Path Planning for Autonomous Vehicle<sup>\*</sup>

CaiJing Xiu and Hui Chen

School of Automotive Studies, Tongji University, 1239 Siping Road,  
Shanghai 200092, China  
xcj2\_11@163.com,  
hui-chen@tongji.edu.cn

**Abstract.** The path planning is one of the key issues of autonomous vehicle. In this paper, a behavior-based method is used for path planning. This method can manage complexity environment and is easy to design and test behaviors. An autonomous electric vehicle is tested in a driving course including the behaviors of road following, turning, obstacle avoidance and emergency braking. The test results show that the path planning algorithm can run in a real-time, while the unknown obstacles can be avoided and the autonomous vehicle can be driven toward the destination in a smooth path and continuous motion.

**Keywords:** Autonomous vehicle, Driver behavior characteristics, Path planning, Behavior-based.

## 1 Introduction

Autonomous vehicle becomes a hot research topic due to the demand of the intelligent transportation systems (ITS) and the military applications in the recent decades, while the path planning is one of the key issues of the vehicle. The path planning can be divided into the global path planning or navigation with fully known environmental information and the local path planning with uncertain environmental information. Goal-based path planning should avoid obstacles in a dynamic environment and reach the destination automatically. In the study of path planning, a top-down traditional method consists of the modules of sensing, modeling, planning and action. This method is difficult to establish an accurate map of environment model, so it is difficult to cope with real-time dynamic environment especially urban environment [1]. The traditional method based on artificial potential field (APF) is improved for its disadvantages such as local minimum [2][3][4], but APF does not consider some constraints or regulations of road and traffic, such as stop line and intersection. It means that APF can not be used to deal with all traffic in urban environment. Intelligent control methods, such as fuzzy control [5][6], neural networks[7], swarm[8][9]and so on, are applied in the study of path planning, but most of them are remained in simulation study or used in specific unstructured environment.

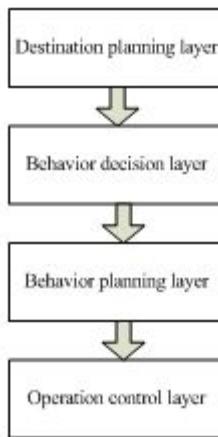
---

\* Supported by Tongji University, Shanghai 200092, China, TRW Automotive R&D (Shanghai) Co., Ltd. and Science and Technology Commission of Shanghai Municipality(07dz05901).

It can be seen that the current path planning methods have the disadvantages mainly in real-time performance, universal property and flexibility for different traffic conditions, especially for urban environment. In consideration of the above issues, the behavior-based path planning is applied for autonomous vehicle in this paper. The behavior is decomposed into a series of relatively independent small behaviors and the behavior decision decides the behavior of autonomous vehicle by the environmental characteristics [10][11].

## 2 Architecture of Autonomous Vehicle

In the process of driving, the structure of the driver's brain activity can be divided into four layers: destination planning, behavior decision, behavior planning and operation control [12], as showed in Fig.1. Destination planning layer is the first activity of the driver to be carried out according to current vehicle status (including the location and pose), driving destinations and pre-stored road map. It is global path planning of autonomous vehicle related to all possible routes from the current status to destination location considering some constrains such as shortcut, optimal time and so on. A series of sub-goals are outputted from this layer [12]. Behavior decision layer is carried out based on the environment characteristics, including vehicle status, traffic rules and other information in the memory in order to determine the vehicle's driving behavior. Behavior plan layer is the planning of driving trajectory and vehicle speed according to driving behavior which is outputted by behavior decision layer. Operational control layer is the operation of vehicle driving, braking, steering that is tracking vehicle trajectory which is given by behavior planning layer.

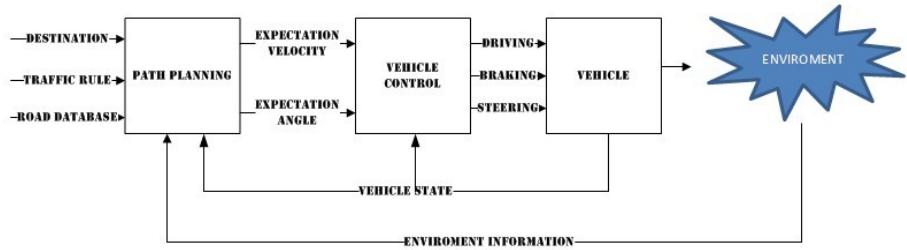


**Fig. 1.** Driver's Thinking Layer

From the view point of controlling, the vehicle is the controlled object, the driver is the sensor and controller, while the road and traffic regulations are the constraints of the system in the process of driving. Driving safety is the basic requirements of the system. From the above analysis, it can be found that the design target of an autonomous

vehicle is to replace the driver by the sensors and controllers, and to meet the requirements of various constraints and conditions.

The ultimate goal of autonomous vehicles is to replace the driver from the closed-loop control system of "driver - vehicle - road". Autonomous vehicle control system showed in Fig.2 consists of four parts in this article: path planning, vehicle control, vehicle and environment. The path planning of this closed-loop control system is studied in this paper.



**Fig. 2.** Closed-Loop Control System of Autonomous Vehicle

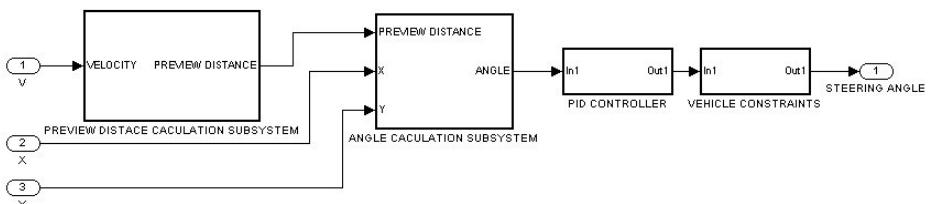
### 3 Design of Path Planning Controller

The behavior-based planning method is applied to manipulate the unmanned vehicles in uncertain environmental information. Behavior decision is made at each sampling time according to real-time sensor data, the results of global and local path planning, the distance and relative velocity between the autonomous vehicle and obstacles. Autonomous vehicle behaviors are classified into road following, turning, obstacle avoidance, emergency braking, ACC, merging into traffic flow, intersection crossing and so on. In this paper, the first four behaviors of road following, turning, obstacle avoidance and emergency braking are discussed.

#### 3.1 Road Following

When there is no obstacle or the distance between autonomous vehicle and obstacle is greater than a safe distance, the road following behavior is triggered. The control algorithm of road following behavior is based on preview following theory [13] combined with PID control. The control model is shown in Fig.3.

In this control model,  $V$  is the velocity of autonomous vehicle,  $(X, Y)$  is the current coordinates of autonomous vehicle, they are the input of controller system. The



**Fig. 3.** Road Following Control Model

preview distance is given by the equation:  $S = K * V$ , while  $K$  is a proportional coefficient and  $V$  is the velocity of autonomous vehicle. Steering angle is calculated according to the lateral deviation between vehicle current coordinates and preview point coordinates. The vehicle constrain module is the vehicle dynamics and kinematics constraints such as the maximum steering angle and steering angular velocity, etc.

### 3.2 Turning Behavior

When the autonomous vehicle encounters an intersection, turning behavior is triggered. Because of the vehicle dynamics and kinematics constraints, the curvature of tracking of autonomous vehicle must be bounded and continuous, and the derivative of curvature is also bounded. Therefore, behavioral strategy of turning is calculated by polar polynomial of Nelson [14] [15].

General Nelson polarity polynomial expression is:

$$r(\varphi) = a_0 + a_1\varphi + a_2\varphi^2 + a_3\varphi^3 + a_4\varphi^4 + a_5\varphi^5 \quad (1)$$

$r$  is the polar radius,  $\varphi$  is the polar angle,  $a_i$  ( $i=1,2,\dots$ ) is polynomial coefficients, they will be given in the following equation (4).

The constraint of location, slope  $r'$  and curvature  $\kappa$  are:

$$\begin{cases} r = R, r' = 0, \kappa = 0 & \varphi = 0 \\ r = R, r' = 0, \kappa = 0 & \varphi = \Phi \end{cases} \quad (2)$$

As the curvature of the path is:

$$\kappa(\varphi) = \frac{r^2 + 2r'^2 - rr''}{(r^2 + r'^2)^{\frac{3}{2}}} \quad (3)$$

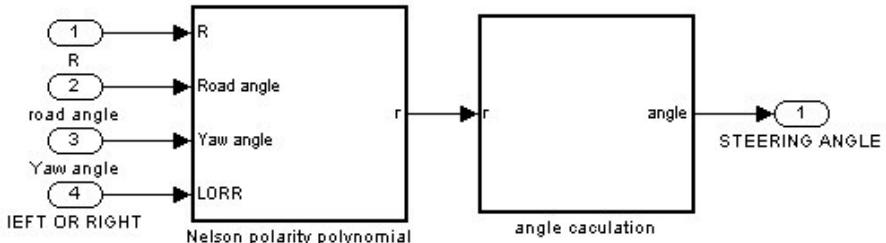
$$r' = \frac{dr}{d\varphi} \quad r'' = \frac{d^2r}{d\varphi^2}$$

From equations of (1) (2) (3), the equation (4) can be obtained.

$$\begin{cases} a_0 = R \\ a_1 = 0 \\ a_2 = \frac{R}{2} \\ a_3 = -\frac{R}{\Phi} \\ a_4 = \frac{R}{2\Phi^2} \\ a_5 = 0 \end{cases} \quad (4)$$

To substitute the parameters of equation (1) by equation (4), the expression will be:

$$r(\varphi) = R \left( 1 + \frac{\varphi^2}{2} - \frac{\varphi^3}{\Phi} + \frac{\varphi^4}{2\Phi^2} \right) \quad (5)$$



**Fig. 4.** Turning Control Model

The control model is shown in Figure4. R is maximum turning radius when turns, road angle (the  $\Phi$  in the formula) is the angle between the current road and the next driving road of autonomous vehicle, LOOR is an sign from global path planning that inform turn right or left. Angle calculation subsystem calculated front wheel angle by equation:  $\text{angle}=\arctan(L/r)$ , L is the length of autonomous vehicle, slip angle is ignored in this equation.

### 3.3 Obstacle Avoidance

When the distance between autonomous vehicle and the obstacle is less than safe distance, the obstacle avoiding behavior is triggered. Obstacle avoidance behavior is achieved by an improved artificial potential field method.

In the structured environment, the road boundary can be considered as a repulsive force field. There is no target point defined in the traditional artificial potential field method. In this paper, the distance between autonomous vehicle and target point is considered in the repulsive potential field [2][3].

Gravitational potential function is follows:

$$U_{at}(X) = \frac{1}{2} k(X - X_{goal})^2 \quad (6)$$

K is the location gain coefficient,  $(X - X_{goal})$  is the distance between autonomous vehicle and target point, the corresponding attractive potential field is the negative gradient of target point.

$$\bar{F}_{at}(X) = -k(X - X_{goal}) \quad (7)$$

Improved repulsive potential field function is as following:

$$U_{re}(X) = \begin{cases} \frac{1}{2}\eta\left(\frac{1}{\rho} - \frac{1}{\rho_o}\right) + \frac{1}{2}\eta\left(\frac{1}{\rho} - \frac{1}{\rho_o}\right)(X - X_{goal})^2 & \rho \leq \rho_o \\ 0 & \rho > \rho_o \end{cases} \quad (8)$$

$\eta$  is the location gain coefficient of repulsion,  $\rho$  is the distance between autonomous vehicle and obstacle,  $\rho_o$  is the safety threshold of autonomous vehicle. When the distance of autonomous vehicle and obstacle is greater than the safety threshold, and the repulsive function is zero. Repulsion function is obtained by following equation:

$$\bar{F}_{re}(X) = \eta\left(\frac{1}{\rho} - \frac{1}{\rho_o}\right)\frac{1}{\rho^2} + \eta\left(\frac{1}{\rho} - \frac{1}{\rho_o}\right)\frac{1}{\rho^2}(X - X_{goal})^2 \quad \rho \leq \rho_o \quad (9)$$

$$0 \quad \rho > \rho_o$$

The force of the autonomous vehicle is:

$$\bar{F}_{sum} = \bar{F}_{at} + \sum \bar{F}_{re} \quad (10)$$

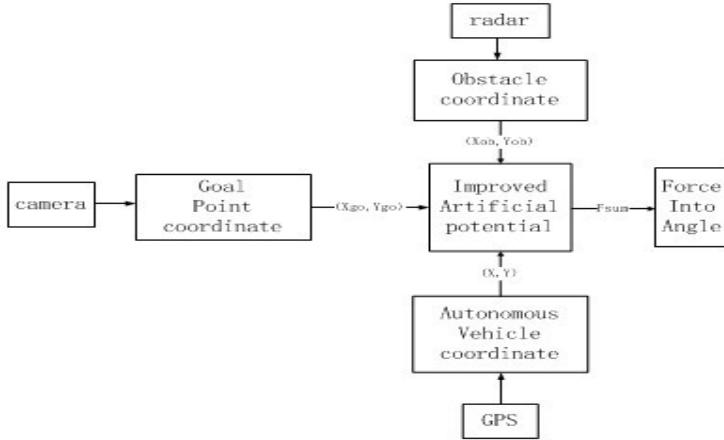
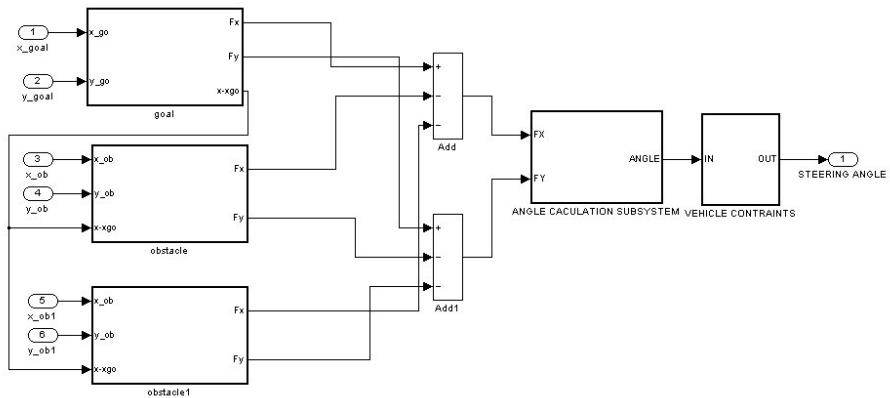


Fig. 5. Sensor Fusion

Figure 5 is the input of APP method,  $(X_{ob}, Y_{ob})$  is the obstacle coordinate that is given by processing of radar data;  $(X_{go}, Y_{go})$  is the goal point coordinate that is given by processing of camera data;  $(X, Y)$  is the autonomous vehicle coordinate that is given by GPS.  $(X - X_{goal})$  and  $\rho$  will be known with the formula of the Length of the line between two points. Then the Fre and Fat are ensured, because the input of autonomous vehicle must be steering angle, according to the orientation of autonomous vehicle, Fre and Fat will be divided into two pitch  $F_x$  and  $F_y$ , and the steering angle is calculated by the following formula:  $\text{angle} = \arctan(\sum F_y / \sum F_x)$ . Obstacle avoidance control Model is shown in Fig.6.

Autonomous vehicle will avoid the obstacles when the front obstacle hinder the travel of autonomous vehicle, and meet road constraint according to this method.



**Fig. 6.** Avoid Obstacle Model

### 3.4 Emergency Braking Behavior

When the autonomous vehicle can not avoid collision by change the steering and deceleration, emergency brake behavior is triggered.



**Fig. 7.** Experiment Platform

## 4 Experiment

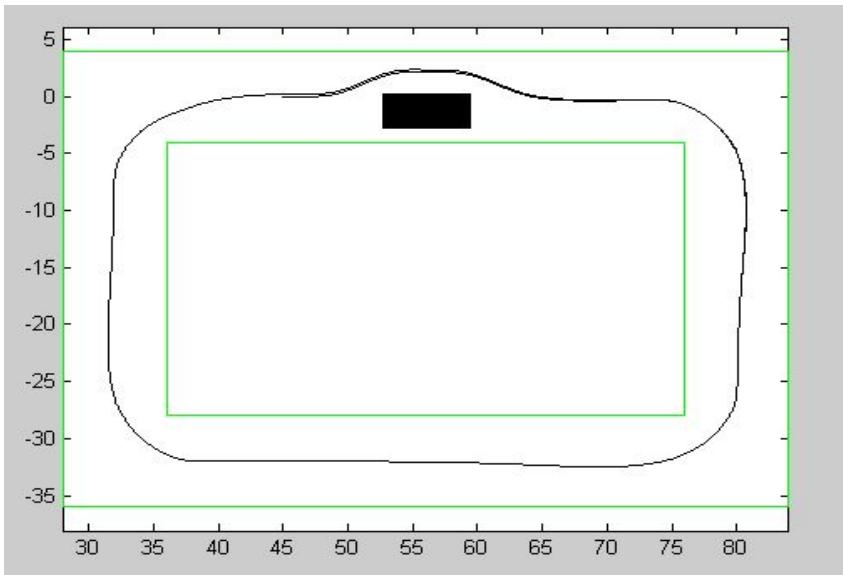
In order to verify the behavior-based path planning algorithm, a real vehicle experiment is performed.

### 4.1 Experiment Condition

The autonomous electric vehicle shown in Fig.7 drives along a rectangular route while a static or moving obstacle will emerge randomly. The vehicle speed is controlled at 10km/h for this experiment. The environment perception sensors are lidar, camera, GPS and so on.

### 4.2 Experiment Result

The trajectory of autonomous vehicle is shown in Figure 8. According to the safety threshold, safety threshold is calculated as  $S = K * V$ , while the coefficient  $K$  is related with the dynamics of vehicle and the communication speed of sensor, and is set up as  $K = 8$ . When obstacles emerge, emergency braking behavior or collision avoiding behavior will be carried out.



**Fig. 8.** Experiment Tracking

## 5 Conclusions

In this paper, a behavior-based path planning method is applied for autonomous electric vehicle. The experiment results show, 1) The applied path planning method is effective for a real-time dynamic environment. 2) The behavior decision can perform a correct

behavior selection only with the information of vehicle status and surrounding environment characteristics. The flexibility and real-time performance of the control system is ensured. 3) Driving behaviors can be successfully decomposed into several elemental behaviors. It shows the applied path planning method could be adapted to an urban environment with complex driving behaviors. The further experiment studies will be performed to verify the adaptability of urban driving condition.

## References

1. Zhang, Y., Zhang, C.: The Study of Mobile Robot Path Planning in Uncertain Environment. Journal of Hangzhou Electronic Science and Technology University (2005)
2. Wang, M., Wang, X.-r.: Study of local path planning of mobile robot based on improved artificial potential field methods. Computer Engineering and Design (2008)
3. Shen, K.-h., Tan, Y.-g.: A Method Solving the Local Minimization Problem of Artificial Potential Field. Journal of Wuhan University of Technology (2005)
4. Conner, D.C.: Sensor fusion,navigation, and control of autonomous vehicles, Virginia Polytechnic Institute and State University, Thesis for the degree of Master
5. Hodge, N.E., Trabia, M.B.: Steering fuzzy logic controller for an autonomous vehicle. In: Proceedings of the 1999 IEEE International Conference on Robotics & Automation (1999)
6. El Hajjaji, A., Oulasine, M.: Modeling human vehicle driving by fuzzy logic for standardized ISO double lane change maneuver. In: Proceedings of the 2001 IEEE Symposium on Robot and Human Interactive Communication (2001)
7. Engedy, I., Horvath, G.: Artificial Neural Network based Mobile Robot Navigation. In: IEEE International Symposium on Intelligent Signal Processing (2009)
8. Doctor, S., Venayagamoorthy, G.K.: Unmanned Vehicle Navigation Using Swarm Intelligence, IEEE (2004)
9. Ramirez-Serrano, A., Pettinaro, G.C.: Navigation of Unmanned Vehicles Using a Swarm of Intelligent Dynamic Landmarks, IEEE (2005)
10. Langer, D., Rosenblatt, J.K., Hebert, M.: A Behavior-Based System For Off-Road Navigation. The Robotics Institute Carnegie Mellon University
11. Gindele, T., Jagszent, D.: Design of the Planer of Team AnnieWAY's Autonomous Vehicle Used in the DARPA Urban Challenge 2007. In: 2008 IEEE Intelligent Vehicles Symposium (2008)
12. He, H.: The Study of the Improvement and Technology of Intelligent Vehicle, Defense Technology university
13. Gao, Z.-h., Guan, H., Guo, K.-h.: Driver directional control model and the application in the research of intelligent vehicle. China Journal of Highway and Transport 13(3) (2000)
14. Nelson, W.: Continuous-Curvature Paths for Autonomous Vehicles. In: IEEE Int. Conf. on Robotics and Automation, Scottsdale, Ariz, vol. 3, pp. 1260–1264 (May 1989)
15. Pinchard, O., Liegeois, A., Pougnet, F.: Generalized polar polynomials for vehicle path generation with dynamic constraints. In: Proc. of the IEEE Int. Conf. on Robotics and Automation, Minneapolis, MN (US), April 1996, vol. 1, pp. 915–920 (April 1996)

# An Uncertainty Approach for Fixture Layout Optimization Using Monte Carlo Method

Xiaoping Zhang, Wenyu Yang, and Miao Li

State Key Lab of Digital Manufacturing Equipment and Technology,  
Huazhong University of Science and Technology, Wuhan, China, 430074  
[shoppinggre@hotmail.com](mailto:shoppinggre@hotmail.com), [mewyang@mail.hust.edu.cn](mailto:mewyang@mail.hust.edu.cn),  
[limiao712@yahoo.cn](mailto:limiao712@yahoo.cn)

**Abstract.** The fixture layout is well known as an important factor that influences the localization accuracy of the workpiece in mass production. Meanwhile, due to the existence of the uncertain errors that originated from assembly and manufacture process, the positional variability which depends on the tolerance allocation has statistical features. In this work, we propose an uncertainty approach of optimizing the fixture layout to improve overall product quality. Firstly, we analyze the deterministic localization model and static equilibrium condition by introducing related uncertainty errors. Then, using the Monte Carlo Method (MCM), we randomly assign the fixture parameters from corresponding probability distributions. Further, we compute the norm of position error of critical point according to the principles of minimum potential energy and the Nonlinear Least Square Method (NLSM). After we sequentially search the minimal mean and variance of the position error in discrete point set domain, the optimal solution of the fixture layout is obtained. Finally, a numerical example is illustrated and compared with the result of FEA simulation.

**Keywords:** Fixture layout, positional variability, uncertainty errors, Monte Carlo Method, mean, variance, Nonlinear Least Square Method.

## 1 Introduction

Obtaining the high product quality has been the target that was pursued by researchers and engineers for decades. The main methods are concentrated on improving the machining accuracy and repeatability so that the geometric error of the workpiece can be eliminated as possible. However, the sources of the errors are not always determinate or predictable. To overcome the difficulties, efficient controllable equipment like sensory feedback suite was introduced for machine tools to actively control the error [1]. But, for mass production, in practice, the cost is too high for factory rectification and implementation. Meanwhile, fixture is another major factor that responsible for the quality of the workpiece. To ensure the geometric features of the critical points within the tolerance zone, carefully designing the fixture layout that the fixture-workpiece system is insensitive to the uncertain errors is a relatively convenient way to lower the total manufacturing cost.

In the previous work, Asada [2] proposed the geometric perturbation techniques of describing the deterministic localization of the workpiece based on classic screw theory. Extending the analysis of perturbation method and form-closure criterion, the design of the fixture layout was formulated as optimization problems [3], [4], [5], [6], in which various quantitative indexes were introduced. In addition, the mathematical criterion of concerning the accessibility and approachability for the complex fixture was proposed in [7]. Then, the compliant information was considered afterwards. Finite Element Method (FEM) was primarily applied in [8], [9], and the contact elasticity was explored by Johnson [10], in which Hertz contact theory was introduced. Based on the works of Johnson, Xiong [11] and Kulankara [12] proposed the constrained nonlinear programming of solving the feasible clamping force.

However, the model of deterministic localization [2] and force closure, which were applied in above studies, considered the variabilities as the determinate ones which did not truly reflect the reality in production. Unfortunately, there were relatively less works that referred the uncertainties in the researching communities of fixture design. Only early works of robotic grasping included the impacts of uncertainty on motion planning [13], contact representation [14] and Jacobian Matrices[15]. For more, assuming the locator errors as statistical variables that follow independent normal distribution, Wang [16] derived the variance of the localization error. Taking into account the uncertainties of contact position and friction, Zheng and Qian [17] evaluated the force closure property of the grasping system. Recently, Sanchez [18] analyzed the product quality in a simple machining fixture by means of introducing the MCM and the cost loss function.

Before the formulation of the uncertainty method, we, firstly, assume that the fixture-workpiece system is pseudo-rigid that only local contact compliance is considered at contact points and the contact surface of the potential point set is continuously smooth that the normal and frictional forces are well defined. Meanwhile, we assume the deformation of contacts is generated locally that the normal direction of the contact force remains unchanged. To concentrate on analyzing the effects of the random errors in fixtures, other uncertainties as vibration of external forces, human experience, etc, will not be considered in this paper.

## 2 Uncertainty Analysis of Deterministic Localization and Static Equilibrium Condition

As shown in Fig. 1, without considering the uncertainty effects, a general workpiece which is constrained by the  $i$ th fixture element<sup>1</sup>, we establish the global frame  $\{\mathbf{G}\}$ , the workpiece frame  $\{\mathbf{W}\}$  and the  $i$ th fixel frame  $\{\mathbf{L}_i\}$ .  $\mathbf{n}_i \in \mathbf{R}^{3\times 1}$  is the inner unit normal directed inward to the workpiece.  $\mathbf{r}_i \in \mathbf{R}^{3\times 1}$  and  $\mathbf{r}_{li} \in \mathbf{R}^{3\times 1}$  are the position vector of the  $i$ th contact point with respect to the workpiece and the  $i$ th fixel, respectively. Once there is a positional variability  $\delta y_i$  at the  $i$ th contact point, a small displacement  $\delta \mathbf{q} = (\delta \mathbf{p}^T, \delta \boldsymbol{\theta}^T)^T \in \mathbf{R}^{6\times 1}$  will be generated as:

---

<sup>1</sup> We refer the fixture element as fixel for the rest of the paper.

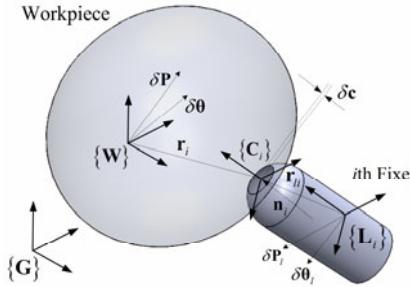
$$\mathbf{G}_i^T \delta \mathbf{q} = \delta \mathbf{y}_i, \quad i = 1, \dots, m+n. \quad (1)$$

$$\delta \mathbf{y}_i = \mathbf{G}_{li}^T \delta \mathbf{q}_{li} + (\delta \mathbf{D}_{wi} - \delta \mathbf{D}_{li}) + \delta \mathbf{c}_i \in \mathbf{R}^{3 \times 1} \quad (2)$$

Where  $\mathbf{G}_i = (\mathbf{I}, (\mathbf{r}_i \times)^T)^T \in \mathbf{R}^{6 \times 3}$  and  $\mathbf{G}_{li} = (\mathbf{I}, (\mathbf{r}_{li} \times)^T)^T \in \mathbf{R}^{6 \times 3}$ <sup>2</sup> are referred as the grasping matrices at the  $i$ th contact point of the workpiece and the  $i$ th fixel, respectively.  $\delta \mathbf{q}_{li} = (\delta \mathbf{p}_{li}^T, \delta \theta_{li}^T)^T \in \mathbf{R}^{6 \times 1}$  is the position and orientation error of the  $i$ th locator.  $\delta \mathbf{D}_{wi} \in \mathbf{R}^{3 \times 1}$  and  $\delta \mathbf{D}_{li} \in \mathbf{R}^{3 \times 1}$  are the datum errors of the  $i$ th contact surface of the workpiece and the  $i$ th fixel, respectively.  $\delta \mathbf{c}_i \in \mathbf{R}^{3 \times 1}$  is the local deformation of the  $i$ th contact. When the workpiece is fully constrained by  $m$  locators and  $n$  clamps, the deterministic location of the workpiece can be described as:

$$\mathbf{G}^T \delta \mathbf{q} = \delta \mathbf{y} \quad (3)$$

Where  $\mathbf{G} = (\mathbf{G}_1, \dots, \mathbf{G}_{m+n}) \in \mathbf{R}^{6 \times 3(m+n)}$ ,  $\delta \mathbf{y} = (\delta \mathbf{y}_1^T, \dots, \delta \mathbf{y}_{m+n}^T)^T \in \mathbf{R}^{3(m+n) \times 1}$ .



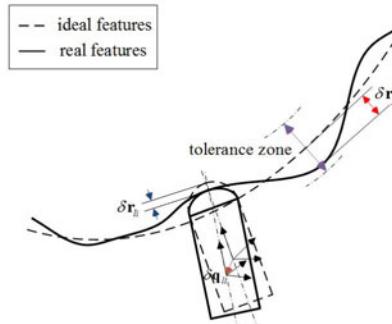
**Fig. 1.** Fixture-workpiece System

However, lacking the information of the uncertainties of the error sources, the localization model stated above is not necessarily enough when confronting with condition of the mass production in practice. Assuming there are  $s$  workpiece need to be machined with the same model, the datum error of the  $k$ th ( $k = 1, \dots, s$ ) workpiece  $\delta \mathbf{D}_{wi}^k$  appears to be randomly allocated in the tolerance zone with a level of confidence due to the uncertainty of pre-manufacturing process. As shown in Fig.2, there exist a random error of the  $i$ th position vector  $\delta \mathbf{r}_i^k = (\delta r_{xi}^k, \delta r_{yi}^k, \delta r_{zi}^k)^T$  which may

---

<sup>2</sup>  $\mathbf{r}_i \times = (r_{ix}, r_{iy}, r_{iz})^T \times = \begin{bmatrix} 0 & -r_{iz} & r_{iy} \\ r_{iz} & 0 & -r_{ix} \\ -r_{iy} & r_{ix} & 0 \end{bmatrix} \in \mathbf{so}(3)$ ,  $\mathbf{so}(3)$  is the Lie algebra of the special orthogonal group  $\mathbf{SO}(3)$ .

follows a 3-dimensional normal distribution<sup>3</sup>  $\delta\mathbf{r}_i^k \sim N((\boldsymbol{\mu}_r^k)^T, (\boldsymbol{\sigma}_r^k)^T, (\rho_r)^T)$  with mean  $\boldsymbol{\mu}_{ri}^k = (\mu_{rix}^k, \mu_{riy}^k, \mu_{riz}^k)^T$  and variance  $\boldsymbol{\sigma}_r^k = (\sigma_{rix}^k, \sigma_{riy}^k, \sigma_{riz}^k)^T$ , which are referred as the accuracy and repeatability [19]. The tolerance range of  $\pm t_r^k$  can be described as  $t_r^k = 3\sigma_r^k$ . Thus, the  $i$ th position vector of the workpiece is actually to be  $\hat{\mathbf{r}}_i = \mathbf{r}_i + \delta\mathbf{r}_i$ . In addition, tear and wear of the  $i$ th datum surface on the fixel will be generated during the continuous loading and unloading manipulation or manufacturing process. The datum error of the  $i$ th fixels  $\delta\mathbf{D}_{li}$  is reflected as the random error  $\delta\mathbf{r}_{li} = (\delta r_{lix}, \delta r_{liy}, \delta r_{liz})^T$ , which may follows a 3-dimensional exponential distribution  $E(\boldsymbol{\lambda}^T)$  with  $\boldsymbol{\lambda} = (\lambda_{lix}, \lambda_{liy}, \lambda_{liz})^T > \mathbf{0}$  according to the theory of reliability [20]. Hence, the real position vector of the  $i$ th fixel is  $\hat{\mathbf{r}}_{li} = \mathbf{r}_{li} - \delta\mathbf{r}_{li}$ . Further, taking into account of the manufacture and assembly errors of the  $i$ th locator, the variability  $\hat{\delta\mathbf{q}}_{li} = (\hat{\delta\mathbf{p}}_{li}^T, \hat{\delta\theta}_{li}^T)^T$  may follows a 6-dimensional normal distribution that  $\hat{\delta\mathbf{q}}_{li} \sim N((\boldsymbol{\mu}_{qli})^T, (\boldsymbol{\sigma}_{qli})^T)$ . Where  $\boldsymbol{\mu}_{qli} = (\mu_{plix}, \mu_{ply}, \mu_{pliz}, \mu_{\thetalix}, \mu_{\thetaly}, \mu_{\thetalz})^T$ , and  $\boldsymbol{\sigma}_{qli} = (\sigma_{plix}, \sigma_{ply}, \sigma_{pliz}, \sigma_{\thetalix}, \sigma_{\thetaly}, \sigma_{\thetalz})^T$ .



**Fig. 2.** Uncertain Errors of the Fixture-Workpiece System

Considering all the uncertainties of error sources stated above, the localization model can be rewritten as:

$$\hat{\mathbf{G}}^T \hat{\delta\mathbf{q}} = \hat{\delta\mathbf{y}} \quad (4)$$

Where  $\hat{\delta\mathbf{y}}_i = \hat{\mathbf{G}}_{li}^T \hat{\delta\mathbf{q}}_{li} + \delta\mathbf{c}_i \in \mathbf{R}^{3 \times 1}$  is the real variability of the  $i$ th contact point.  $\hat{\mathbf{G}}_i = (\mathbf{I}, (\hat{\mathbf{r}}_i \times)^T) \in \mathbf{R}^{6 \times 3}$  and  $\hat{\mathbf{G}}_{li} = (\mathbf{I}, (\hat{\mathbf{r}}_{li} \times)^T) \in \mathbf{R}^{6 \times 3}$  are the real grasping matrix of the  $i$ th contact point.

<sup>3</sup> We assume that all elements of the uncertain errors are mutual independent that correlation coefficient  $\rho = 0$  for the rest of the paper.

Being loaded into fixtures, the workpiece requires to be fully constrained in terms of interacting at  $(m + n)$  contact points. And in order to counterbalance the external disturbances such as cutting force and gravity, the static equilibrium condition of the fixture-workpiece system should be satisfied. As shown in Fig. 1, we establish the  $i$ th contact frame  $\{\mathbf{C}_i\}$  at the  $i$ th contact point.  $\mathbf{n}_i$ ,  $\mathbf{t}_{i1}$  and  $\mathbf{t}_{i2}$  are the unit inner normal and two orthogonal tangential vectors, respectively. The unit contact wrench of the  $i$ th contact can be expressed as:

$$\mathbf{B}_i = \mathbf{G}_i [\mathbf{n}_i, \mathbf{t}_{i1}, \mathbf{t}_{i2}] = \begin{bmatrix} \mathbf{n}_i & \mathbf{t}_{i1} & \mathbf{t}_{i2} \\ \mathbf{r}_i \times \mathbf{n}_i & \mathbf{r}_i \times \mathbf{t}_{i1} & \mathbf{r}_i \times \mathbf{t}_{i2} \end{bmatrix} \in \mathbf{R}^{6 \times 3}, \quad i = 1, \dots, m+n. \quad (5)$$

And the static force equilibrium can be expressed as:

$$\mathbf{BF} = -\mathbf{W}_{ext} \quad (6)$$

Where  $\mathbf{W}_{ext} \in \mathbf{R}^{6 \times 1}$  is the external wrench,  $\mathbf{B} = [\mathbf{B}_1, \dots, \mathbf{B}_{m+n}] \in \mathbf{R}^{6 \times 3(m+n)}$  and  $\mathbf{F} = [\mathbf{f}_1^T, \dots, \mathbf{f}_{m+n}^T]^T \in \mathbf{R}^{3(m+n) \times 1}$ .  $\mathbf{f}_i = (f_{in}, f_{it1}, f_{it2})^T \in \mathbf{R}^{3 \times 1}$  is the  $i$ th contact force.

Considering the uncertainty of the datum errors of the workpiece and reasonably ignoring the small change of the contact force direction, the unit contact wrench can be rewritten as:

$$\hat{\mathbf{B}}_i = \hat{\mathbf{G}}_i [\mathbf{n}_i, \mathbf{t}_{i1}, \mathbf{t}_{i2}] = \begin{bmatrix} \mathbf{n}_i & \mathbf{t}_{i1} & \mathbf{t}_{i2} \\ \hat{\mathbf{r}}_i \times \mathbf{n}_i & \hat{\mathbf{r}}_i \times \mathbf{t}_{i1} & \hat{\mathbf{r}}_i \times \mathbf{t}_{i2} \end{bmatrix} \in \mathbf{R}^{6 \times 3}, \quad i = 1, \dots, m+n. \quad (7)$$

Then, the static force equilibrium with uncertainty analysis can be expressed as:

$$\hat{\mathbf{B}}\mathbf{F} = -\mathbf{W}_{ext} \quad (8)$$

### 3 Optimiztion Scheme

#### 3.1 Monte Carlo Approach

Monte Carlo method (MCM) is a technique for iteratively evaluating a deterministic model using sets of random numbers as inputs. This method is often used when the model is complex, nonlinear, or impossible to compute an exact result with a deterministic algorithm. As we stated in Section 2, in the situation of mass production, it is unfeasible of ignoring the uncertain parameters as  $\delta \mathbf{r}_i$ ,  $\delta \mathbf{r}_h$  and  $\delta \mathbf{q}_h$  when utilizing the model of deterministic localization and static equilibrium condition. Thus, MCM can be introduced to simulate the impacts of the uncertain errors on the positional accuracy of the critical point.

According to the algorithm of the MCM, we describe the simulating process as follow:

- a. Define  $\delta\mathbf{r}_i$ ,  $\delta\mathbf{r}_{li}$  and  $\delta\mathbf{q}_{li}$  as the uncertain inputs;
- b. Randomly generate these inputs, according to their characteristics, from the domain using Normal and Exponential distributions, respectively;
- c. Perform a deterministic computation, which will be elaborated in the next subsection, using these inputs;
- d. Gather the value of the positional error  $\delta\mathbf{u}$  of the critical point from the individual computations into sets of results;
- e. Evaluate the mean and variance of  $\|\delta\mathbf{u}\|$ ;

Once the MCM simulation is done, the accuracy and the repeatability of the product quality under one specified scheme of fixture layout can be obtained by the deterministic computation.

### 3.2 Computation of Positional Error of Critical Point

As we known, solving the displacement  $\delta\mathbf{q}$  of the workpiece with contact compliance is a static indeterminate problem [11], [12]. The classic Hertz contact theory is, then, introduced to describe the nonlinear relationship of local deformation and contact force. The most common one depicts the contact between a sphere and a plane which can be used as a back-of-the-envelope engineering estimation for the contact between two curved surfaces [21]. The contact compliance is referred as follows:

$$\delta d_{in} = (9f_{in}^2 / (16E_e^2 R_e))^{1/3} \quad (9)$$

$$\delta d_{it1} = f_{it1} / (8a_i G_e) \quad (10)$$

$$\delta d_{it2} = f_{it2} / (8a_i G_e) \quad (11)$$

Where  $R_e = (2(R_{wi1}^{-1} + R_{wi2}^{-1} + R_{fi1}^{-1} + R_{fi2}^{-1}))^{-1}$  is the equivalent radius of the  $i$ th contact point.  $R_{wi1}$ ,  $R_{wi2}$  and  $R_{fi1}$ ,  $R_{fi2}$  are the two principal curvature radii of the workpiece and the  $i$ th fixel, respectively.  $E_e = (E_w^{-1}(1 - \nu_w^2) + E_{fi}^{-1}(1 - \nu_{fi}^2))^{-1}$  is the equivalent elastic modulus that  $E_w$ ,  $E_{fi}$  and  $\nu_w$ ,  $\nu_{fi}$  are the elastic modulus and Poisson's ratio of the workpiece and the  $i$ th fixel.  $G_e = (G_w^{-1}(2 - \nu_w) + G_{fi}^{-1}(2 - \nu_{fi}))^{-1}$  is the equivalent shear modulus that  $G_w$ ,  $G_{fi}$  are the shear modulus of the workpiece and the  $i$ th fixel, respectively.  $a_i = (3f_{in}R_e / (4E_e))^{1/3}$  is the radius of the contact region.

Meanwhile, the deformation we obtained from the compliant model is described in the  $i$ th local contact frame  $\{\mathbf{C}_i\}$ , we, then, transformed into the global frame as:

$$\delta\mathbf{c}_i = {}^G_{C_i} \mathbf{R} \delta\mathbf{d}_i, \quad i = 1, \dots, m+n. \quad (12)$$

Where  $\delta\mathbf{d}_i = (\delta d_{it1}, \delta d_{it2}, \delta d_{in})^T \in \mathbf{R}^{3 \times 1}$ .

Once the contact compliance is introduced, the models of deterministic localization and static force equilibrium can be linked together for solving  $\delta\mathbf{q}$ . And according to the theory of Elasticity [22], the real displacement of the workpiece will be generated

under the condition that potential energy of the fixture-workpiece system is minimized. Hence, we can establish a constrained nonlinear optimization problem according to the principle of minimum potential energy. The optimization model is described as below:

$$\begin{aligned} & \text{minimize } \Pi_c = E_p \\ & \text{subject to } C: \begin{cases} h_s(\mathbf{X}) = 0, & s = 1, \dots, 3(m+n)+6 \\ g_t(\mathbf{X}) \geq 0, & t = 1, \dots, (m+n) \end{cases} \end{aligned} \quad (13)$$

Where  $h_s(\mathbf{X}) = 0$  represent the equalities (4) and (8),  $g_t(\mathbf{X}) \geq 0$  are the inequalities (14) of coulomb law:

$$\mu_i f_{in} - \sqrt{f_{it1}^2 + f_{it2}^2} \geq 0 \quad i = 1, \dots, m+n \quad (14)$$

$E_p = \sum_{i=1}^{(m+n)} [\int f_{in} d(\delta d_{in}) + \int f_{it1} d(\delta d_{it1}) + \int f_{it2} d(\delta d_{it2})] - \mathbf{W}_g \cdot \delta \mathbf{q}$  is the potential energy, in which  $\mathbf{W}_g$  is the wrench of gravitational force,  $\delta \mathbf{q}$  is the displacement of the workpiece.  $\mathbf{X} = (\delta \mathbf{q}^T, \delta \mathbf{d}^T)^T \in \mathbf{R}^{3(m+n+2) \times 1}$  is the design variable.

Then, using the Exterior Penalty Method, we can transform this constrained nonlinear programming into an unconstrained one as:

$$\text{minimize } \Pi = E_p + r \left[ \sum_1^{3(m+n)+6} (h_s(\mathbf{X}))^2 + \sum_1^{m+n} (g_t(\mathbf{X}))^2 \right] \quad (15)$$

Where  $r$  is a positive penalty parameter which is designed to be very large that penalize the violation of constrains. Further, we transform (15) into a nonlinear least-square problem:

$$\text{minimize } \Pi = \Psi(\mathbf{X})^T \Psi(\mathbf{X}) \quad (16)$$

Where  $\Psi(\mathbf{X}) = (E_p, (\mathbf{h}(\mathbf{X}))^T, (\mathbf{g}(\mathbf{X}))^T)^T$ .

We can obtain optimal solution of  $\mathbf{X}$  by using some common algorithm as efficient L-M method. Meanwhile, due to the displacement  $\delta \mathbf{q} \in \mathbf{R}^{6 \times 1}$  of the workpiece, the positional error  $\delta \mathbf{u}$  can be described as follow, according to [3],

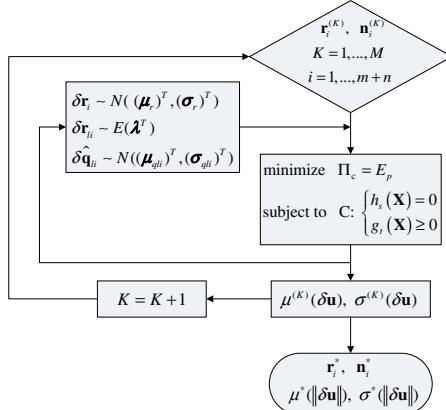
$$\delta \mathbf{u} = \mathbf{G}_u^T \delta \mathbf{q} \quad (17)$$

Where  $\mathbf{G}_u = (\mathbf{I}, (\mathbf{r}_u \times)^T)^T \in \mathbf{R}^{6 \times 3}$ .  $\mathbf{r}_u \in \mathbf{R}^{3 \times 1}$  is the position vector of the critical point.

### 3.3 Sequential Optimization in Discrete Point Set Domain

To optimize the fixture layout, we need to check the available exterior surface on which the set of the possible fixel location appears to be continuous and infinite. However, the due to the characteristics of the fixture and workpiece, it is impossible for the fixel location to be adjusted arbitrarily in practice. The candidates are limited

in a point set domain [16]. Therefore, we can sequentially compute the positional error  $\delta\mathbf{u}$  for all  $M$  candidates of in a point set domain, and obtain the best solution that mean or variance of  $\|\delta\mathbf{u}\|$  is minimal. The flowchart of the optimization process in a point set domain is depicted in Fig.3.



**Fig. 3.** Flowchart of the optimization process in a point set domain

## 4 Numerical Example

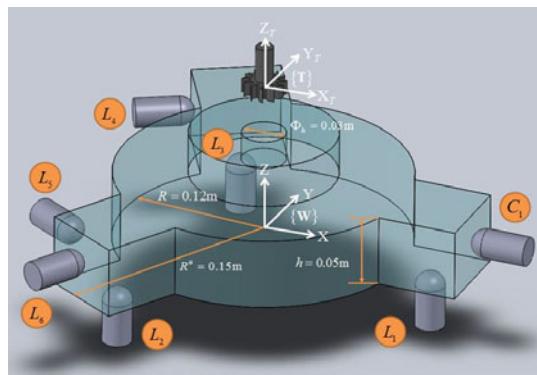
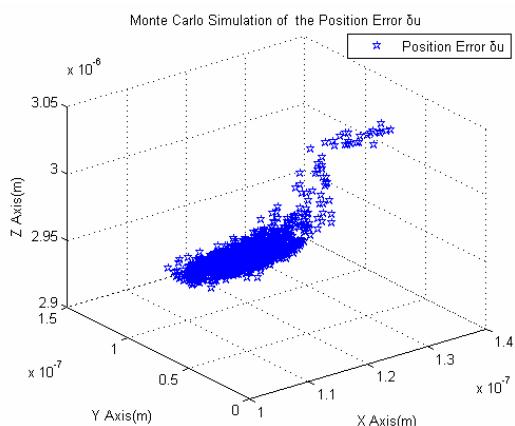
In this section, an example is presented to verify the effectiveness of the uncertainty approach. As shown in Fig.4, the workpiece is a homogeneous flange cover with 18 Kg in weight. The workpiece is made of 304L stainless steel with  $E_w = 203\text{GPa}$  and  $\nu_w = 0.3$ . The mass centroid of the workpiece is at  $[0, 0, 0.0289]^T \text{m}$ . To locate and constrain the workpiece, 6 spheri-cal tipped locating pin and 1 spherical tipped clamp, which are made of 45# steel with  $E_f = 196\text{GPa}$  and  $\nu_f = 0.27$ , are used. The radii of spherical tips are 0.01 m. The clamping force is set as  $f_{cn} = 100\text{N}$ . The operation is to mill a hole on the top of the workpiece, the milling force and torque are given as  $f_{tx} = -40\text{N}$ ,  $f_{ty} = -50\text{N}$ ,  $f_{tz} = -60\text{N}$  and  $m_{tz} = 1\text{N} \cdot \text{m}$ , which are expressed in cutting tool frame  $\{\mathbf{T}\}$ . The hole is located at  $[0, 0, 0.08]^T$  with 0.02m in depth. The static friction coefficient is  $\mu = 0.29$ .

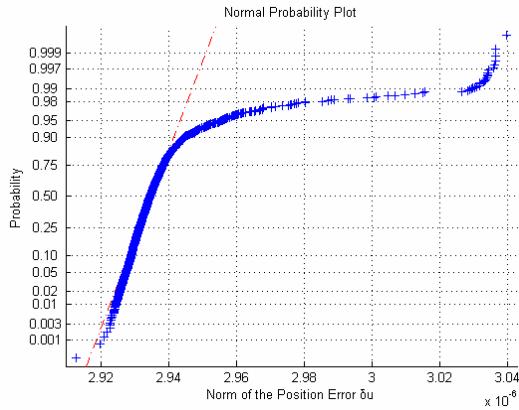
**Table 1.** Assignment of Uncertain Error Source with specified Distribution

Errors (m)	Distribution
$\delta\mathbf{r}_i^k = (\delta r_{xi}^k, \delta r_{yi}^k, \delta r_{zi}^k)^T$	$\delta r_{xi}^k \sim N(0, 1E-8)$ ; $\delta r_{yi}^k \sim N(0, 1E-8)$ ; $\delta r_{zi}^k \sim N(0, 1E-8)$ ;
$\delta\mathbf{r}_{li} = (\delta r_{lx}, \delta r_{ly}, \delta r_{lz})^T$	$\delta r_{lx}^k \sim E(1E+8)$ ; $\delta r_{ly}^k \sim E(1E+8)$ ; $\delta r_{lz}^k \sim E(1E+8)$ ;
$\delta\mathbf{p}_{li} = (\delta x_{li}, \delta y_{li}, \delta z_{li})^T$	$\delta x_{li} \sim N(0, 1E-8)$ ; $\delta y_{li} \sim N(0, 1E-8)$ ; $\delta z_{li} \sim N(0, 1E-8)$ ;
$\delta\boldsymbol{\theta}_{li} = (\delta\theta_{xli}, \delta\theta_{yli}, \delta\theta_{zli})^T$	$\delta\theta_{xli} \sim N(0, 1E-8)$ ; $\delta\theta_{yli} \sim N(0, 1E-8)$ ; $\delta\theta_{zli} \sim N(0, 1E-8)$ ;

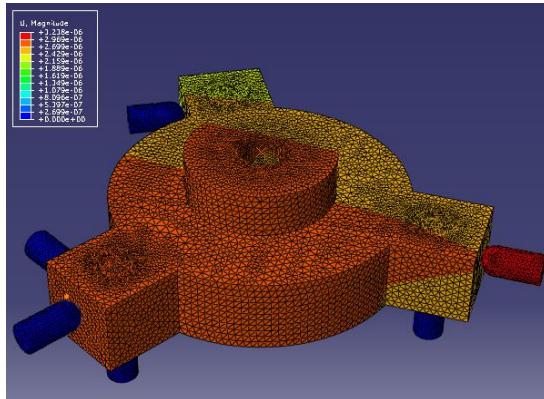
**Table 2.** Scheme of Fixture Layout in Point set domain

Coordinates	
$L_1$	$(0.12 + \alpha, 0, 0)^T$
$L_2$	$(-(0.12 + \alpha) / 2, -\sqrt{3}(0.12 + \alpha) / 2, 0)^T$
$L_3$	$(-(0.12 + \alpha) / 2, \sqrt{3}(0.12 + \alpha) / 2, 0)^T$
$L_4$	$(-0.091 - \beta / 2, 0.0976 + \sqrt{3}\beta / 2, 0.03 + \gamma)^T$
$L_5$	$(-0.091 - \beta / 2, -0.0976 - \sqrt{3}\beta / 2, 0.03 + \gamma)^T$
$L_6$	$(-0.075, -0.13, 0.03 + \gamma)^T$
$C_1$	$(0.15, 0, 0.03 + \gamma)^T$

**Fig. 4.** Workpiece to be Milled



**Fig. 6.** Histogram and Normal Probability plot of the estimated norm of  $\delta\mathbf{u}$



**Fig. 7.** Simulation of Position Error under Optimal Solution of Fixture Layout

The assignment of the uncertain errors with related distribution is shown in Table 1. The location of the fixels which is based on a point set domain is shown in Table 2. Where  $\alpha = -0.02 : 0.01 : 0.02$ ,  $\beta = -0.02 : 0.01 : 0.02$ ,  $\gamma = -0.01 : 0.01 : 0.01$ . In each iteration step, we assign the uncertain parameters with 2000 random samples in terms of MCM function **normrnd** and **exprnd** in MATLAB Statistics Toolbox, respectively. We use a Laptop with 2.66GHz Core Duo2 Processor for computation. The optimal fixture layout is obtained when  $\alpha = 0$ ,  $\beta = 0.02$ ,  $\gamma = 0.01$ . The minimal mean and variance are  $\mu(\|\delta\mathbf{u}\|) = 2.9376E-6$  and  $\sigma(\|\delta\mathbf{u}\|) = 1.4457E-8$ , respectively. The MC simulation of the position error  $\delta\mathbf{u}$  of the optimal fixture layout is shown in Fig.5. The MC simulation of accuracy  $\|\delta\mathbf{u}\|$  for the optimal solution is described in Fig. 6. It shows that the distribution of  $\|\delta\mathbf{u}\|$  could only be approximated as normal due to the fact that the probability of the norm appears to be a curve rather than a straight line which

deviates at about 0.90 that 10% of the  $\|\delta\mathbf{u}\|$  do not follow the normal distribution. The deviation, which is proved in Fig.5 as well, might be resulted from the excessive wear of the fixels, and pertinent fixels should be replaced to ensure the position accuracy. The work of FEA simulation of the optimal solution is conducted in Abaqus 6.8. And the result, as shown in Fig.7, is conformed to the MCM prediction of  $\mu(\|\delta\mathbf{u}\|)$ .

## 5 Summary

In this paper, we provide an uncertainty method of simulating the machine accuracy and repeatability of the workpiece in mass production by means of considering the characteristics of the random errors in fixtures. An optimiztion approach for the fixture layout is also derived to improve the positional precision of the products. Our future work will combine this uncertainty approach into the optimization of the configuration of supports in complex fixture which is based on the theory of [10], the heavy loaded condition will be taken into account as well.

## Acknowledgments

This work was supported by the National Basic Research Development Program of China (No. 2009-CB724306) and the Independent Research Project of State Key Lab of Digital Manufacturing Equipment and Technology. The authors would like to thank Dr. Jin and MS. Ma for their constructive comments.

## References

1. Ramesh, R.M., Mannan, A., Poo, A.N.: Error compensation in machine tools — a review. Part I: Geometric, cutting-force induced and fixture-dependent errors. *International Journal of Machine Tools and Manufacture* 40(9), 1235–1256 (2000)
2. Asada, H., By, A.B.: Kinematic analysis of workpart fixturing for flexible assembly with automatically reconfigurable fixtures. *IEEE Trans. Robot. Autom.* 1(2), 86–93 (1985)
3. Wang, M.Y.: Characterizations of positioning accuracy of deterministic localization of fixtures. *IEEE Trans. Robot. Autom.* 18(6), 976–981 (2002)
4. Liu, Y.H.: Qualitative test and force optimization of 3-D frictional formclosure grasps using linear programming. *IEEE Trans. Robot. Autom.* 15, 163–173 (1999)
5. Zhu, X.Y., Ding, H.: Optimality Criteria for Fixture Layout Design A Comparative Study. *IEEE Trans. Robot. Autom.* 6(4), 658–669 (2009)
6. Yu, Z., Chew, C.M.: Efficient Procedures for Form-Closure Grasp Planning and Fixture Layout Design. *ASME J. Manuf. Sci. Eng.* 131, 041010:1–11 (2009)
7. Xiong, Y.L., Xiong, X.R.: Algebraic structure and geometric interpretation of rigid complex fixture systems. *IEEE Trans. Robot. Autom.* 4(2), 252–264 (2007)
8. Cai, W., Hu, S.J., Yuan, J.X.: Deformable sheet metal fixturing: Principles, algorithms, and simulation. *ASME J. Manuf. Sci. Eng.* 118, 318–324 (1996)
9. Malluck, J.A., Melkote, S.N.: Modeling of deformation of ring shaped workpieces due to chucking and cutting forces. *ASME J. Manuf. Sci. Eng.* 126, 141–147 (2004)
10. Johnson, K.L.: Contact Mechanics. Cambridge Univ. Press, Cambridge (1985)

11. Xiong, C.H., Ding, H., Xiong, Y.L.: On Clamping Planning in Workpiece-Fixture Systems. *IEEE Trans. Robot. Autom.* 5(3), 407–419 (2008)
12. Kulankara, K., Satyanarayana, S., Melkote, S.N.: Iterative fixture layout and clamping force optimization using the Genetic Algorithm. *ASME J. Manuf. Sci. Eng.* 124, 119–125 (2002)
13. Goldberg, K.Y., Mason, M.T., Requicha, A.: Geometric Uncertainty in Motion Planning: Summary Report and Bibliography, Catalina Island, CA, June 15-17 (1992)
14. Xiao, J., Zhang, L.: Towards obtaining all possible contacts — growing a polyhedron by its location uncertainty. *IEEE Transactions on Robotics and Automation* 12(4), 553–565 (1996)
15. Cheah, C.C., Han, H.Y., Sawamura, S., Arimoto, S.: Grasping and position control for multi-fingered robot hands with uncertainty Jacobian Matrices. In: Proceeding of the IEEE International Conference on Robotics and Automation, pp. 2403–2408 (1998)
16. Wang, M.Y.: An optimal design for 3D fixture synthesis in a point-set domain. *IEEE Trans. Robot. Autom.* 16(6), 839–846 (2000)
17. Zheng, Y., Qian, W.H.: Coping with uncertainties in Force-closure analysis. *International Journal of Robotics Research* 24(4), 311–327 (2005)
18. Sanchez, H.T., Estrems, M., Faura, F.: Determination of Key workpiece product characteristics in a machining fixture using uncertainty analysis and loss cost function implementations. *International Journal of Advanced Manufacturing Technology* 41, 452–460 (2009)
19. Slocum, A.H.: Precision Machine Design. Prentice-Hall International Inc., Englewood Cliffs (1992)
20. Barlow, R.E., Proschan, F.: Mathematical Theory of Reliability. Society for Industrial and Applied Mathematics (1987)
21. Tripp, J.: Hertian Contact in two and three Dimension. Paper NASA Tech. 2473 (1985)
22. Timoshenko, S.P., Goodier, J.N.: Theory of Elasticity. The Macgraw-Hill Companies Inc., New York (1970)

# The Small-World Network Model of Mixed Production Line

Ting Yang, Dinghua Zhang, Bing Chen, and Shan Li

Key Laboratory of Contemporary Design & Integrated Manufacturing Technology  
(Northwestern Polytechnical University), Ministry of Education  
127 Youyi Xilu, Xi'an Shaanxi, 710072, P.R. China  
[t.yang@mail.nwpu.edu.cn](mailto:t.yang@mail.nwpu.edu.cn)

**Abstract.** Mixed production line is a complicate system, so the small-world network model is introduced to describe and analyze the planning and design of mixed production line to raise its stability. Manufacturing units undertaking various production duties are regard as vertices, while relationships between vertices are defined as directed weighted edges. Then the mathematic model is built to characterize their basic elements and network architecture with some feature sets and matrixes. Finally an actual production line with 224 manufacturing vertices and 361 products is presented as an example. Moreover some statistical data might uncover its topological properties and dynamic behavior. Through comparing with random network, the mixed production line shows small-world property. Then redundancy of some vertices with high centrality should be enhanced to reduce frangibility and improve robustness and reliability. Therefore the network model is an available method to solve stability in production line planning.

**Keywords:** Complex network, weight, power-law, small-world.

## 1 Introduction

Mixed production line yielding various products with different quantity in a certain period is a common organization mode, which is widely adopted by discrete manufacturing, such as aero-engine, automobile. Because of complexity in structure, dynamic status of elements, multi-step in manufacturing process and flexibility in equipments configuration, doubtlessly mixed production line is a complicated system with all kinds of factors affecting each other. The complex system is a hot topic in 21<sup>st</sup> century, and many theories and methods are brought forward including complex network, cellular automata and multi-agent.

The production line is the basic part of production system, which represents its basic production capability. Owing to diversity in type and quantity of products, all the physical elements of production line are in dynamic environment. For example, manufacturing unit, material-flow facility and workers are variable constantly. The stability of manufacturing unit, e.g. machines, manual operation, has a great effect on the whole production line and other elements through the manufacturing relationships between them. Thereby the relationships between manufacturing units and effect of individual unit upon the whole system are emphasized in design and planning a production line.

Mixed production line is structured in the form of networks linking a large number of natural and man-made systems (large communication systems, transportation infrastructures, biological systems). Specifically, many of these networks show the small-world property [1~4]. Similarly some key vertices and edges will have most important to stability of production line. Anyway their reliability must be endured. Consequently, in this paper network model is built to analyze structure and dynamic behavior of mixed production line. The manufacturing units are regarded as vertices, and relationships between them are defined the directed edge. The status and dynamic behavior of whole line will be analyzed, and key elements might be obtained. Therefore the stability of production line could be raised in planning and design phase.

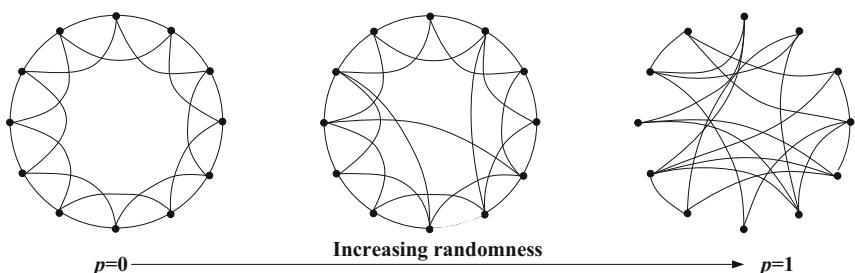
The paper is organized as follows. Small-world network is introduced in section 2. The definition and topological properties of network model are given in section 3. The tools to analyze the network are described in detail. In section 4, an actual production line with 224 vertices and 361 products is studied to exhibit the approach to build a small-world network. The statistical characterization is obtained to reflect the performance status of production line. Then the methods to improve robustness and reliability of production will be discussed. Finally observations and open questions are concluded.

## 2 Small-World Network

Many research results showed that many complex networks are highly clustered, and have small characteristic path lengths, just like the famous conclusion “six degrees of separation”. Watt and Strogatz proposed them as “small-world” networks, by analogy with the small-world phenomenon in social network [1].

To interpolate between regular and random networks, the following random rewiring procedure is represented by Watt and Strogatz in [1]. Starting from a ring lattice with  $n$  vertices and  $k$  edges per vertex, each edge is rewired at random with probability  $p$  (Fig. 1). This construction allows to tune the graph between regularity ( $p=0$ ) and disorder ( $p=1$ ), and thereby to probe the intermediate region  $0 < p < 1$ , about which little is known.

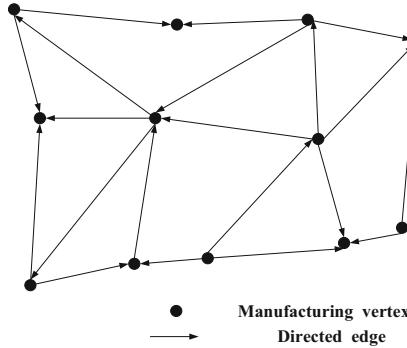
Through simulation and experiment it reveals that for intermediate value of  $p$  a small-world network is highly clustered like a regular graph, yet with small characteristic path length, like a random graph, which is the important character of small world network.



**Fig. 1.** Random rewiring procedure

### 3 Complex Weighted Network Model of Mixed Production Line

On production line level, manufacturing processes act as a topological generator to form network architecture between manufacturing units. Due to diversity in products and manufacturing processes the network model shows complexity. Mixed production line could be defined as a weighted graph  $G(V,E,w)$ , where  $V$  runs over all manufacturing units,  $E$  is directed relationship from one unit to another, as we can see in Fig. 2. Weight  $w$  is assigned to represent the intensity of relationship between vertices.



**Fig. 2.** The architecture of directed network

There are some tools and definitions to describe the basic elements and topological properties of network model as follow. At the same time the mathematical model of mixed production line is built.

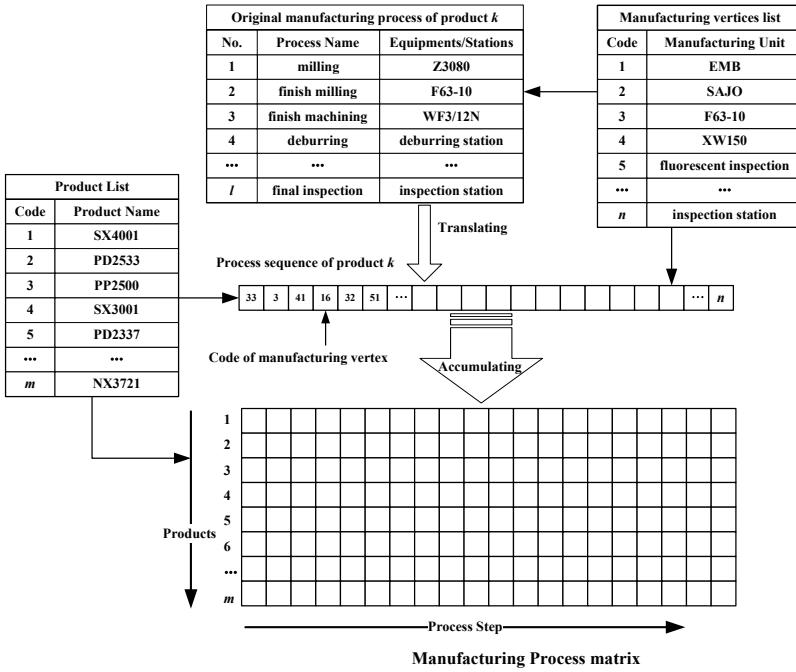
#### 3.1 Basic Elements of Mixed Production Line

Manufacturing resources, products and manufacturing process are the main elements of a production line. In network model, these elements are described as some sets and matrixes. Then a mixed production line with  $m$  products and  $n$  vertices are defined as follow.

*Manufacturing vertices:*  $V\{ v_i \mid v_i=i \}, (i=1,\dots,n)$ . Manufacturing resources which can undertake manufacturing duties such as machines and work stations are simplified to manufacturing units. Then these manufacturing units are coded using numbers from 1 to  $n$  (see Fig.3.) to become vertices in network. Accordingly all the manufacturing vertices constitute a code list.

*Directed edges:*  $E\{ E_{ij} \mid E_{ij}=v_i \rightarrow v_j \}, (i,j=1,\dots,n)$ . The relationships between vertices are described as directed edges. The route from vertex  $i$  to vertex  $j$  is defined as the directed edge  $E_{ij}$  which takes the value 1 if an edge exists and 0 otherwise. It should be paid more attentions that  $E_{ij}$  is unequal to  $E_{ji}$  in directed network.

*Products:*  $P\{ P \mid P_i=i \}, (i=1,\dots,m)$ . Various products in production line are coded using number list from 1 to  $m$ .



**Fig. 3.** Translating procedure of manufacturing process matrix

**Manufacturing process:** Manufacturing process is the cause for the evolution of production system. Moreover process is the driving force to handle material between manufacturing resources. Therefore in mixed production line, manufacturing process determines the basic network structure. In this paper process matrix  $Pr\{ Pr_{ij}|Pr_{ij} \in V \}$ , ( $i=1,\dots,m$ ) is defined to describe all the manufacturing process of the mixed production line. The element  $Pr_{kj}$  is equal to the code of the manufacturing vertex on  $j$ -step of product encoded  $k$ , such that finishing milling in F63-10 on step 2 is shown in Fig.3.

According to manufacturing vertices list, manufacturing process of the product  $k$  has been translated to a code sequence. For example, the process sequence  $P_{rk}=[33 3 41 16 32 51 \dots n]$  means the manufacturing route is  $33 \rightarrow 3 \rightarrow 41 \rightarrow 16 \rightarrow 32 \rightarrow 51 \rightarrow \dots \rightarrow n$ . Through traversal along all the products in the production line, the process sequence is accumulated into the process matrix in accordance with product list as figured in Fig.3.

### 3.2 Tools for the Topological Property of Small-World Network

The performance status of whole production line could be described through some topological properties by statistical method as follow.

**Weight matrix:**  $w\{w_{ij}\}$ , ( $i,j=1,\dots,n$ ) .Weight matrix represents the intensity of edges to exhibit the relationship between vertices. The element  $w_{ij}$  is calculated as

$$\begin{cases} \text{If } Pr_{ij} = k, Pr_{ij+1} = l, \text{ then } w_{kl} = 1 \\ \text{If } Pr_{ij} = k, Pr_{ij+1} = l, \text{ then } w_{kl} = w_{kl} + 1 \ (l > i) \\ \text{If there is no relationship from vertex } k \text{ to vertex } l, \text{ then } w_{kl} = 0 \\ w_{kk} = 0 \\ w_{kl} \neq w_{lk} \end{cases} \quad (1)$$

The multiple relationship should be accumulate ( $w_{kl} = w_{kl} + 1$ ), as shown in equation 1. The probability distribution of any given edge with weight  $w$  could be expressed as  $P(w)$ .

*Adjacent matrix* [5]:  $A\{a_{ij}\}$ , ( $i,j=1,\dots,n$ ). Adjacent matrix describes the connectivity between vertices whose element  $a_{ij}$  takes the value 1 if an edge connects the vertex  $i$  to vertex  $j$  and 0 otherwise. Therefore  $a_{ij}$  is obtained through Boolean operation on weight.

$$a_{ij} = \text{Boolean}(w_{ij}) \quad (2)$$

*Degree of vertices* [6]:  $k\{k_i\}$ , ( $i=1,\dots,n$ ). Degree  $k_i$  is defined as the number of other vertices linked to vertex  $i$ . In directed network, degree is divided into out-degree  $k_{out}$  and in-degree  $k_{in}$ . They are defined as

$$k_{out}: \ k_i = \sum_{j=1}^n a_{ij} \quad k_{in}: \ k_j = \sum_{i=1}^n a_{ij} \quad (3)$$

*Strength of vertices* [7]:  $s\{s_i\}$ , ( $i=1,\dots,n$ ). The strength measures the total connectivity of vertices. Similarly the strength is divided into out-strength and in-strength, which is defined as

$$s_{out}: \ s_i = \sum_{j=1}^n w_{ij} \quad s_{in}: \ s_j = \sum_{i=1}^n w_{ij} \quad (4)$$

The strength represents the total quantity of manufacturing task to be accomplished in the given vertex and reflects the importance in the whole system. The average of degree and strength is defined as  $\langle k \rangle$  and  $\langle s \rangle$ .

*Clustering coefficient of vertices*:  $c\{c_i\}$ , ( $i=1,\dots,n$ ). The clustering coefficient measures the local cohesiveness in the neighborhood of the vertex. Indeed, it yields the fraction of interconnected neighbors of a given vertex. The clustering coefficient is obtained through validity of connected triple [6] [8].

$$c_i = \frac{1}{k_i(k_i - 1)} \sum_{j,h} a_{ij} a_{ih} a_{jh} \quad (5)$$

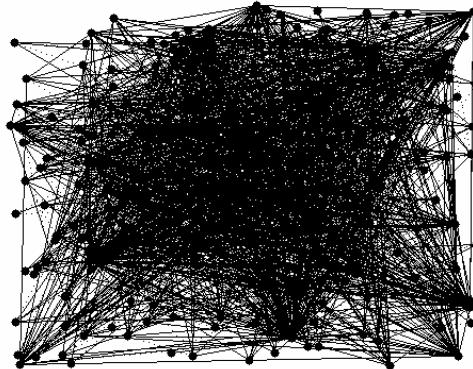
The clustering coefficient of the whole network is equal to the average of all vertices, which is  $C = \sum_{i=1}^n c_i / n$ .

*Shorted path length*:  $d\{d_{ij}\}$ , ( $i,j=1,\dots,n$ ). The distance between the vertex  $i$  and vertex  $j$  is defined as total length of the shorted path. The shorted path can be obtained by Dijkstra algorithm. The average path length of the whole network is defined as

$$L = \frac{1}{\frac{1}{2}n(n-1)} \sum_{i,j} d_{ij} \quad (6)$$

## 4 Analysis on an Actual Mixed Production Line

A production line for aero-engine case is proposed as an example, which has 224 manufacturing vertices and yields 361 kinds of products. These products are in various quantities and have great diversity in manufacturing process. And it is a typical mixed production line with complicated property. Failure of one manufacturing unit might cause a terrible chaos in the whole line. The stability analysis is an important issue in production controlling. Therefore this production line is built to a network model and its topology is shown in Fig.4. It could be seen that there are some vertices having a very large number of connections compared with the most vertices. These vertices with high connectivity are called hubs.



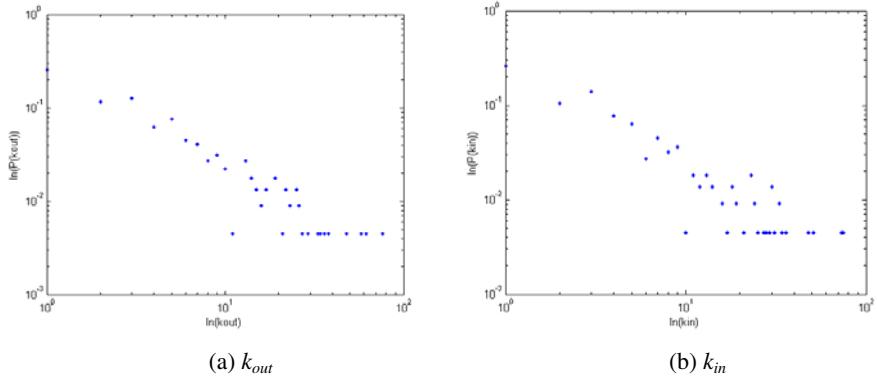
**Fig. 4.** The topology of mixed production line with 224 vertices and 1668 edges

Using methods and algorithms mentioned in section 3, the process matrix ( $224 \times 81$ ) is generated and coded according to products list and vertices list. Then the weight matrix and adjacent matrix could be obtained. Consequently by means of equation (3) ~ (6), the topological properties are given in Table 1. It could be seen that the maximum is much larger than the average. Therefore distribution of these properties should be paid more attention to uncover the load of vertices and edges.

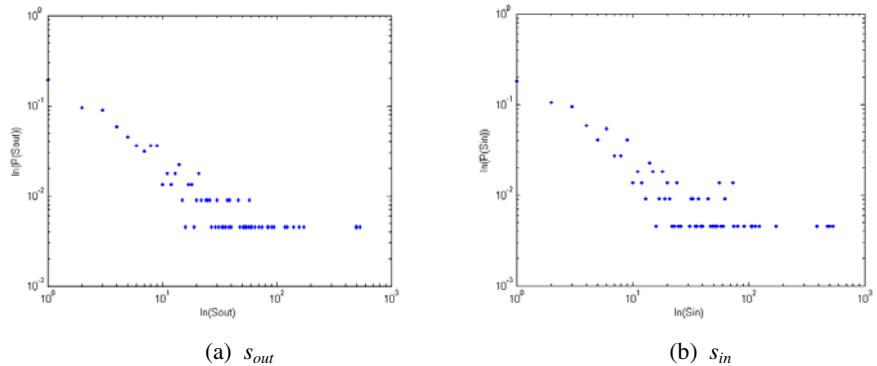
**Table 1.** Topological property of mixed production line

Property	Average	Max.(Out/In)
Degree	7.446	76/75
Strength	23.531	537/536
Weight	3.160	190
Clustering Coefficient	0.278	1
Path Length	2.720	4.516

The probability distribution of degree, strength of vertices is drawn in double logarithm coordinate (Fig.5 Fig.6).



**Fig. 5.** Degree distribution



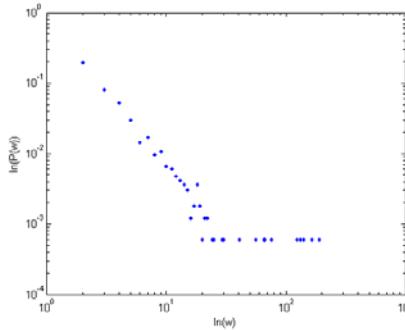
**Fig. 6.** Strength distribution

It can be seen that these figures are approximate to a linear distribution in double logarithm coordinate. It means that the increasing relationship follows power-law behavior, as shown in equation (7) (8).

$$P(k) \sim k^{-\alpha}, \quad \ln P(k) \sim -\alpha \ln k \quad (7)$$

$$P(s) \sim s^{-\beta}, \quad \ln P(s) \sim -\beta \ln s \quad (8)$$

Meanwhile the distribution of weight is shown in Fig.7.



**Fig. 7.** Weight distribution

Similarly, the probability distribution takes form

$$P(w) \sim w^{-\gamma}, \quad \ln P(w) \sim -\gamma \ln w \quad (9)$$

The networks with power-law distribution are called as scale-free network.

On overall view, the whole property of mixed production line is compared with the ER random network with the same quantity and average degree of vertices [6].

The clustering coefficient

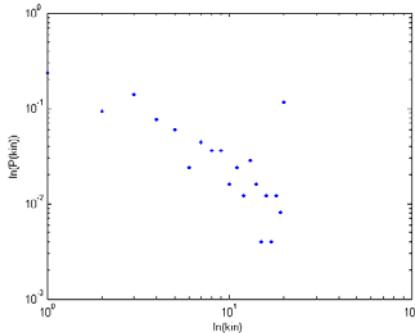
$$C = 0.278 \gg C_{ER} \approx \frac{\langle k \rangle}{n} = 0.033$$

And the average path length

$$L = 2.720 \approx L_{ER} \approx \frac{\ln n}{\ln \langle k \rangle} = 2.695$$

Obviously the mixed production line has greatly higher clustered than ER random network, yet small average path length like random network. The network model shows small-world property. However power-law distribution and small-world property mean that system have higher tendency to failure. If one critical vertex was disabled in the accident, it would lead to a catastrophic failure of the whole system [9][10]. The stability of production line is threatened.

Therefore the configuration of production line must be adjusted to eliminate centrality. On the one hand the redundancy of some vertices should be enhanced to avoid impact from partial failure, which could improve robustness and reliability of production line. So a threshold of in-degree might be set to limit the working-load of manufacturing on certain vertex, and spare duty would be completed by enhancive vertices. For aero-engine case production line in this paper, the in-degree threshold [ $k_{in}$ ] is given as 20. So vertices encoded (10 12 14 16 23 51 68 70 82 92 99 101 110 121 122 139 145 155 160 175 179 200 221) are confirmed as key vertices. These vertices will be increased and the amount of vertices is increased from 224 to 253. Then the probability distribution of  $k_{in}$  is shown in Fig.8.



**Fig. 8.** The distribution of  $k_{in}$  after adding vertices

The figure indicates that the power-law is broken. The maximum of degree is restricted in a reasonable area, and the production line is in safety.

Moreover the centrality on some vertices sometimes results from process planner's predilection for some equipment. In fact different equipments in a same production line could interexchange in manufacturing capability. Thereby the centrality could be decreased through balancing production duty within same manufacturing capacity. And the operation capacity of production line will be ensured through replacement when one disablement.

Furthermore for edges with high weight, the redundancy should increase to ensure the connectivity between vertices, for example adding the amount of handling facilities.

## 5 Conclusions

Small-world network is introduced to model mixed production line and analyze the stability of the whole network. The mathematical model is brought forward to describe product, process and manufacturing resource in this paper. An actual production line with 224 vertices and 361 products is presented to research the whole performance and dynamic status. The degree, strength and weight are approximate to follow power-law distribution through statistical analysis. The whole system show small-world property compared with random network, which is dangerous to stability. Then through increasing redundancy of vertices, excessive centralities on some key vertices are demolished. Hereby the small-world network is a forceful method to study the stability of mixed production line through uncovering its whole architecture and statistic character. Moreover the method and algorithm proposed in this paper provide a new approach and theory to analyze production system with complicated characterizations.

## Acknowledgements

The project is supported by the National High Technology Research and Development Program of China (863). (No. 2009AA04Z122).

## References

1. Watts, D.J., Strogatz, S.H.: Collective dynamics of “small-world” networks. *Nature* 393, 440–442 (1998)
2. Shen, H.C., Chau, H.L., Wong, K.K.: An Extended Cellular Automaton Model for Flexible Manufacturing system. *International journal of advanced manufacturing technology* 11, 258–266 (1996)
3. Freeman, L.C.: *Sociometry* 40, 35–41
4. Barrat, A., Barthélemy, M., Pastor-Satorras, R., et al.: The architecture of complex weighted networks. *Applied physical sciences* 101, 3747–3752 (2004)
5. Wang, Y.L.: Theory method and application of system engineering. Higher Education Press, Beijing (1992)
6. Wang, X.F., Li, X., Chen, G.R.: The theory and application of complex network. Tsinghua University Press, Beijing (2006)
7. Barthélemy, M., Barrat, A., Pastor-Satorras, R., et al.: Characterization and modeling of weighted networks. *Physical A* 346, 34–43 (2005)
8. Barrat, A., Barthélemy, M., Vespignani, A.: Modeling the evolution of weighted networks. *Physical Review E* 70, 066149 (2004)
9. Wang, H.L., Shan, Y., Wang, X.F.: The small-world network model of combinational logic circuits. *Electric Machines and Control* 10, 370–374 (2006)
10. Yan, L.M., Jin, H.Z., Rong, P.X., et al.: Brittleness and the brittle source of system. *Journal of Harbin Engineering University* 27, 223–226 (2006)

# Distributed Cooperative Control of Industrial Robotic Systems Using Petri Net Based Multitask Processing

Gen'ichi Yasuda

Nagasaki Institute of Applied Science  
536 Aba-machi, Nagasaki 851-0193, Japan  
[YASUDA\\_Genichi@NiAS.ac.jp](mailto:YASUDA_Genichi@NiAS.ac.jp)

**Abstract.** The methods of modeling and control of discrete event robotic manufacturing systems using Petri nets are considered, and a methodology of decomposition and coordination is presented for hierarchical and distributed control. First, a task specification is defined as a Petri net model at the conceptual level, and then transformed to the detailed Petri net representation of the sub-tasks. Finally, the overall Petri net is decomposed and the constituent subnets are assigned to local Petri net based controllers. The controllers are coordinated so that the decomposed transitions fire at the same time by the coordinator. System coordination algorithm through communication between the coordinator and the controllers, is presented. By the proposed method, modeling, simulation and control of large and complex manufacturing systems can be performed consistently using Petri nets.

**Keywords:** Distributed control, robotic manufacturing systems, Petri nets, multitask processing.

## 1 Introduction

The rapid advancement of industrial techniques makes robotic application systems larger-scale and more complicated. In the field of flexible manufacturing systems that are typical examples of the discrete event driven system, demands for automatic control have diversified and the control logic has become extremely complicated. To realize model based control systems for complex robotic systems such as flexible manufacturing systems, it is necessary to provide effective tools for describing process specifications and developing control algorithms in a clear and consistent manner. Conventionally, process modeling at the conceptual level is based on time charts, state transition graphs, etc., and the control algorithms at the execution level are generally developed with relay ladder diagrams, or procedural languages for programmable logic controllers or common microprocessors. However, in the area of real-time control of discrete event manufacturing systems the main problems that the system designer has to deal with are concurrency, synchronization, and resource sharing problems. For this class of problems, Petri nets have intrinsic favorable qualities and it is very easy to model sequences, choices between alternatives, rendezvous and concurrent activities including nondeterministic, asynchronous behaviors [1]. The network model can describe the execution order of

sequential and parallel tasks directly without ambiguity [2]. Moreover, the formalism allowing a validation of the main properties of the Petri net control structure (liveness, boundedness, etc.) guarantees that the control system will not fall immediately in a deadlocked situation which sometimes leads a system dangerous condition. The last aspect is essential because the sequences of control are complex and change very often in flexible manufacturing systems,. Furthermore, a real-time implementation of the Petri net specification by software called the token game player can avoids implementation errors, because the specification is directly executed by the token game player and the implementation of these control sequences preserves the properties of the model. In this approach, the Petri net model is stored in a database and the token game player updates the state of the database according to the operation rules of the model. For control purposes, this solution is very well suited to the need of flexibility, because, when the control sequences change, only the database needs to be changed [3].

However, in the field of flexible manufacturing systems, the network model becomes complicated and it lacks for the readability and comprehensibility. Therefore, the flexibility and expandability are not satisfactory in order to deal with the specification change of the manufacturing system. Despite the advantages offered by Petri nets, the synthesis, correction, updating, etc. of the system model and programming of the controllers are not simple tasks [4]. Furthermore, the increasing use of robots and other automated machines has generated control software written in different levels and it is impossible for one person to understand all the control specifications in a large and complex manufacturing system. To deal with the complexity, programming techniques that make it possible to realize systematic and high-level description of system specification are necessary.

The author proposes a Petri net based specification and real-time control method for large and complex discrete event manufacturing systems. Based on the hierarchical and distributed structure of the manufacturing system, the specification procedure is a top-down approach from the conceptual level to the detailed level of the discrete event manufacturing systems. The macro representation of the system is broken down to generate the detailed Petri nets at the local machine control level. Then the Petri nets are decomposed and assigned to the machine controllers to perform distributed control using Petri net based multitask processing. An algorithm is proposed for coordination of machine controllers. By the proposed method, modeling, simulation and control of large and complex manufacturing systems can be performed consistently using Petri nets.

## 2 Modeling with Petri Nets

A Petri net has two types of nodes, called places and transitions. A place is represented by a circle and a transition by a bar. The places and transitions are connected by arcs. The number of places and transitions are finite and not zero. An arc is connected directly from one place to a transition or a transition to a place. In other words a Petri net is a bipartite graph, i.e. places and transitions alternate on a path made up of consecutive arcs. An ordinary Petri net [5] is a bipartite graph represented by the 4-tuple  $G = \{P, T, I, O, M\}$  such that:

$P = \{p_1, p_2, \dots, p_n\}$  is a finite, not empty, set of places;

$T = \{t_1, t_2, \dots, t_m\}$  is a finite, not empty, set of transitions;

$P \cap T = \emptyset$ , i.e. the sets  $P$  and  $T$  are disjointed;

$I : P \times T \rightarrow \{0,1\}$  is the input incidence function;

$O : P \times T \rightarrow \{0,1\}$  is the output incidence function;

$I(p_i, t_j)$  is the weight of the arc  $p_i \rightarrow t_j$ .  $O(p_i, t_j)$  is the weight of the arc  $t_j \rightarrow p_i$ . The pre-incidence matrix of a Petri net is  $C^- = [c_{ij}^-]$  where  $c_{ij}^- = I(p_i, t_j)$ ; the post-incidence matrix is  $C^+ = [c_{ij}^+]$  where  $c_{ij}^+ = O(p_i, t_j)$ , then the incidence matrix of the Petri net  $C = C^+ - C^-$ .

Each place contains integer (positive or zero) marks or tokens. The number of tokens in each place is defined by the marked vector or marking  $M(m_1, m_2, \dots, m_n)^T$ . The number of tokens in one place  $p_i$  is called  $M(p_i)$ . An initial marking  $M_0$  is defined as the state such that  $m_i = 0$  for each place  $p_i$ . The marking at a certain moment defines the state of the Petri net, or the state of the system described by the Petri net. The evolution of the state therefore corresponds to an evolution of the marking, caused by the firing of transitions.

In an ordinary Petri net, where the current marking is  $M_k$ , a transition  $t_j$  is enabled if  $\forall p_i \in P, M_k(p_i) \geq I(p_i, t_j)$ .  $V(v_1, v_2, \dots, v_m)^T$  is the transition vector, where  $v_j = 1$  if transition  $t_j$  fires and 0 if not. An enabled transition can be fired reaching a new marking  $M_{k+1}$  which can be computed as  $M_{k+1} = M_k + C \cdot V$ ; this equation is called state equation of the Petri net.

Reachability is a fundamental basis for analyzing the dynamic properties of any system. The firing of an enabled transition will change the token distribution (marking) in a net according to the transition rule. A sequence of firings will result in a sequence of markings. A marking  $M_n$  is said to be reachable from a marking  $M_0$  if there exists a sequence of firings that transforms  $M_0$  to  $M_n$ . The set of all possible markings reachable from  $M_0$  is denoted  $R(M_0)$ . A Petri net is said to be  $k$ -bounded or simply bounded if the number of tokens in each place does not exceed a finite number for any marking reachable from  $M_0$ , i.e.,  $M(p_i) \leq k$  for every place  $p_i$  and every marking  $M \in R(M_0)$ . A Petri net is said to be safe if it is 1-bounded. Bumping occurs when despite the holding of a condition, the preceding event occurs. This can result in the multiple holding of that condition. From the viewpoint of discrete event process control, bumping phenomena should be excluded. The Petri net was modified so that the system is free of this phenomenon [6]. Because the modified Petri net is safe, for each

place  $p_i$ ,  $m_i = 0$  or 1, and the weight of the input and output incidence functions  $I, O$  is 1 if the arc exists and 0 if not. Thus a transition  $t_j$  is enabled if the two following conditions are met:

- (1) For each place  $p_k$ , such that  $I(p_k, t_j) = 1$ ,  $m_k = 1$
- (2) For each place  $p_l$  such that  $O(p_l, t_j) = 1$ ,  $m_l = 0$

When an enabled transition  $t_j$  fires, the marking  $M$  is changed to  $M'$ , where

- (1) For each place  $p_k$ , such that  $I(p_k, t_j) = 1$ ,  $m'_k = m_k - 1$
- (2) For each place  $p_l$  such that  $O(p_l, t_j) = 1$ ,  $m'_l = m_l + 1$

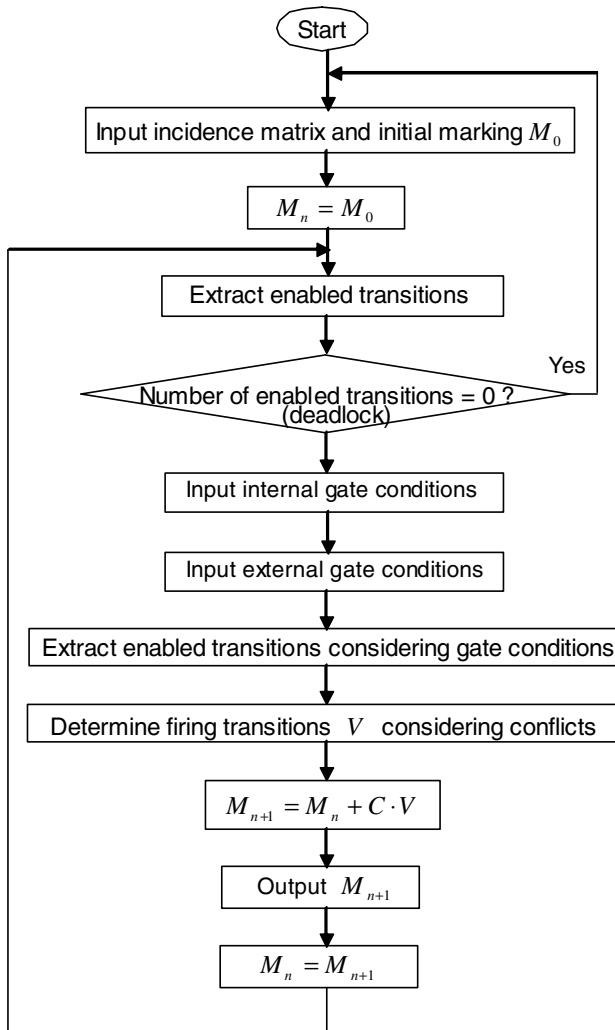
The firing of a transition removes tokens from all its input places and put a token in each output place connected to it. In any initial marking, there must not exist more than one token in a place. According to these rules, the number of tokens in a place never exceeds one, thus, the Petri net is essentially a safe graph.

Besides the guarantee of safeness, considering not only the modeling of the systems but also the actual manufacturing system control, the additional capability of input/output signals from/to the machines are required. The extended Petri net consists of the following six elements: (1) Place, (2) Transition, (3) Directed arc, (4) Token, (5) Gate arc, (6) Output signal arc. A place represents a condition of a system element or action. A transition represents an event of the system. A token is placed in a place to indicate that the condition corresponding to the place is holding. When a transition is enabled, this does not imply that it will be immediately fired; this only remains a possibility. The firing of a transition is indivisible; the firing of a transition has duration of zero.

A gate arc connects a transition with a signal source, and depending on the signal, it either permits or inhibits the occurrence of the event which corresponds to the connected transition. Gate arcs are classified as permissive or inhibitive, and internal or external. An output signal arc sends the signal from a place to an external machine. To summarize the above, a transition is enabled if and only if it satisfies all the following conditions:

- (1) It does not have any output place filled with a token.
- (2) It does not have any empty input place.
- (3) It does not have any internal permissive arc signaling 0.
- (4) It does not have any internal inhibitive arc signaling 1.

An enabled transition may fire when it does not have any external permissive arc signaling 0 nor any external inhibitive arc signaling 1. If a place has two or more input transitions or output transitions, these transitions may be in conflict for firing. When two or more transitions are firable only one transition should fire using some arbitration rule. Fig. 1 shows the flow chart of the extended Petri net simulation.



**Fig. 1.** Flow chart of extended Petri net simulation

A Petri net is said to be live if no matter what marking has been reached from an initial marking, it is possible to ultimately fire any transition of the net by progressing through some further firing sequence, and the initial marking is said to be a live marking [7]. If the above is not satisfied, the marking is said to be dead. In case of the dead marking, some transition can not fire at all after that time. If there is no marking which is live, the net is said to be structurally dead. If the marking is live or dead according to the initial distribution of tokens, the net is said to be conditionally dead. For the modified Petri net, the following properties have already been found and are important with respect to deadlock phenomena.

- (1) By reversing the direction of each arc from/to a place and the presence of a token in the place, there is no change in the firing condition of any transition in the petri net.
- (2) A Petri net which has no conflict place can be transformed to a Petri net which has no loop.
- (3) If there is a loop which has no conflict place, the number of tokens in the loop is invariable.
- (4) There is a marking, which is reachable from a live initial marking and returns the same marking through some further firing sequence, and the firing sequence can contain all the transitions.
- (5) In case that a marking returns to the same marking through some nonzero firing sequence, the sum of firing of input transitions is equal to the sum of firing of output transitions, for each place in the net.
- (6) If the rank of the incidence matrix of the net is equal to the number of the transitions, the net is structurally dead. Practically, to reduce the amount of computation, the Petri net is decomposed into a set of simple nets with respect to conflict places.

For the actual control, the operations of each machine are broken down into a series of unit motions, which is represented by mutual connection between places and transitions. A place means a concrete unit motion of a machine. From these places, output signal arcs are connected to the machines, and external gate arcs from the machines are connected to the transitions of the Petri net when needed, for example, to synchronize and coordinate operations. When a token enters a place that represents a subtask, the machine defined by the machine code is informed to execute a determined subtask with a determined data; these data are defined as the place parameters. Fig.2 shows the Petri net model of the general robotic task, and Fig. 3 shows the Petri net representation of real-time control of a robotic unit action.

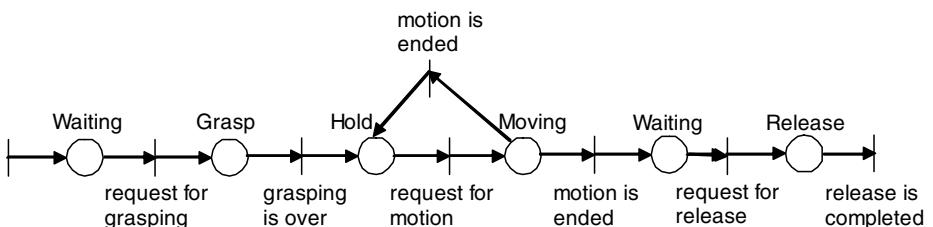


Fig. 2. Petri net model of robotic task

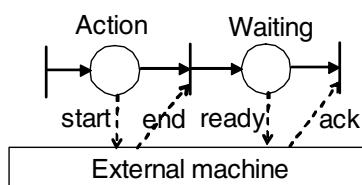


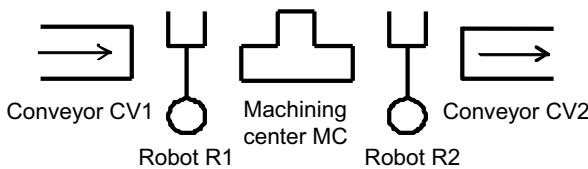
Fig. 3. Petri net representation of real-time control of robotic task

### 3 Petri Net Model Based Manufacturing Process Specification

A specification procedure for discrete event manufacturing systems based on Petri nets is as follows. First, the conceptual level activities of the discrete event manufacturing systems are defined through a Petri net model considering the task specification corresponding to the manufacturing process. A conceptual Petri net model describes the aggregate manufacturing process. At the level, a place which defines each subtask composing the task specification is represented in parallel with a place which defines each equipment used in each subtask is also represented as a place.

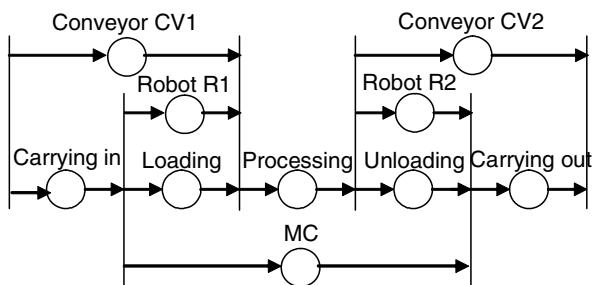
The example manufacturing system has two robots, one machining center, and two conveyors, where one is for carrying in and the other is for carrying out, as shown in Fig. 4. The main execution of the system is indicated as the following task specification:

- (1) A workpiece is carried in by the conveyor CV1.
- (2) The robot R1 loads the workpiece to the machining center MC.
- (3) The machining center MC processes the workpiece.
- (4) The robot R2 unloads the workpiece from the machining center and places it on the conveyor CV2.
- (5) The workpiece is carried out by the conveyor CV2.



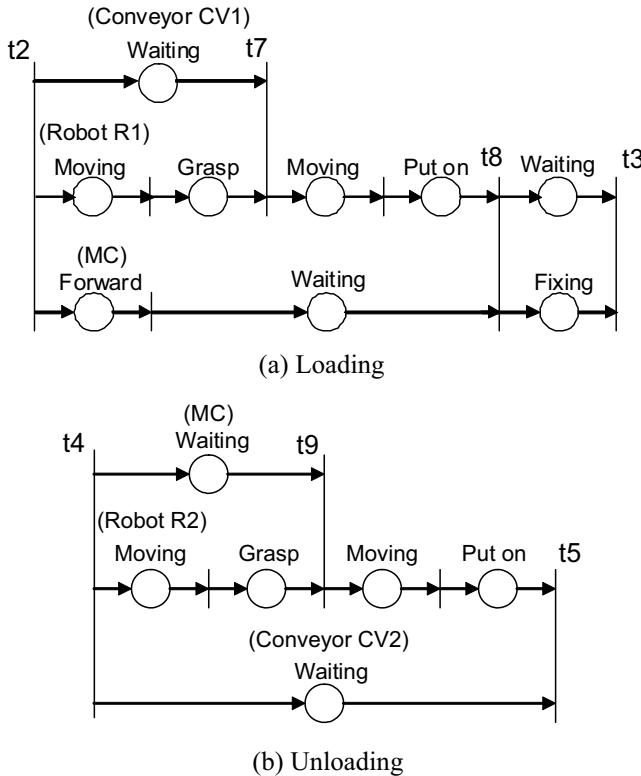
**Fig. 4.** Example of robotic manufacturing system

At the conceptual level, the places which represent the subtasks indicated as the task specification are connected by arcs via transitions in the specified order corresponding to the flow of subtasks and a workpiece. The places representing the existence of machines are also added to connect transitions which correspond to the beginning and ending of their subtasks as shown in Fig. 5.



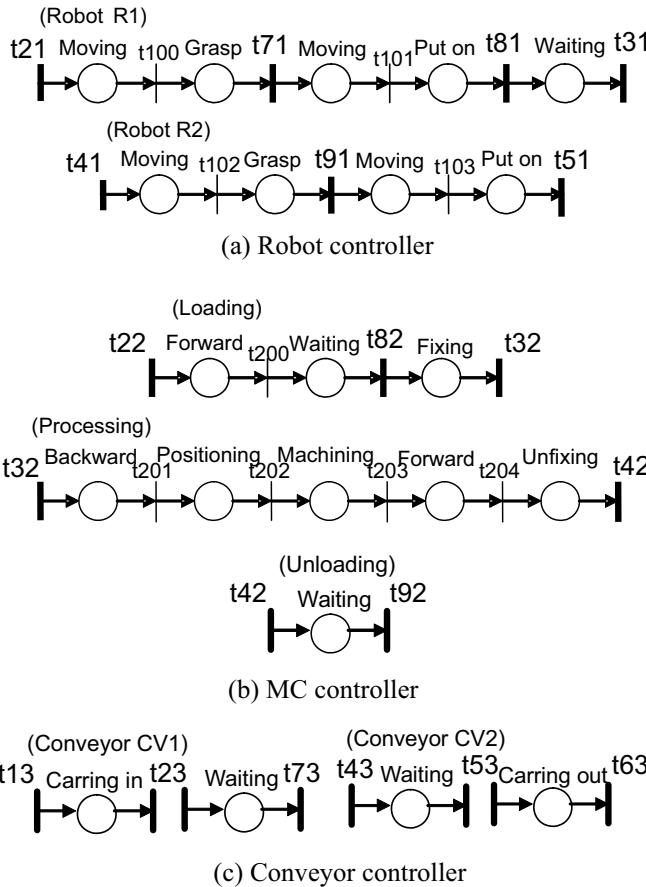
**Fig. 5.** Petri net representation of the example system at the conceptual level

Next, each place representing a subtask at the conceptual level is translated into a detailed subnet considering machine activity specifications and required control strategies. Fig. 6 shows the detailed Petri net representation of loading, processing and unloading in Fig. 5. Based on the hierarchical approach, the Petri net is translated into the detailed Petri net by stepwise refinements from the highest conceptual level to the lowest machine control level. At each step of detailed specification, a place of the Petri net is substituted by a subnet in a manner which maintains the structural properties [8].



**Fig. 6.** Detailed Petri net representation of subtasks

It is natural to implement a hierarchical and distributed control system, where one controller is allocated to each control layer or block. For the robotic manufacturing systems composed of robots, machine tools, and conveyors, an example structure of hierarchical and distributed control is composed of one station controller and three machine controllers, although each robot may be controlled by one robot controller. The detailed Petri net is decomposed into subnets, which are assigned to machine controllers. For the distributed control system, the Petri net representations assigned to the machine controllers are shown in Fig. 7.



**Fig. 7.** Petri net representation of machine controllers

( [ ] global transition, | : local transition)

In the decomposition procedure, a transition may be divided and distributed into different machine controllers. The machine controllers should be coordinated so that these transitions fire simultaneously so that the aggregate behavior of decomposed subnets should be the same as that of the original Petri net. Decomposed transitions are called global transitions, and other transitions are called local transitions.

By the Petri net model, the state of the discrete event system is represented as the marking of tokens, and firing of any transition brings about change to the next state. So the firing condition and state (marking) change before decomposition should be the same as those after decomposition. From the logical calculus, it is proved that the firability condition of the original transition is equal to AND operation of firability conditions of decomposed transitions. If and only if all of the decomposed transitions are firable, then the global transitions are firable. To exploit the above results, the coordinator program has been introduced to coordinate the decomposed subnets so that the aggregate behavior of decomposed subnets is the same as that of the original Petri net.

In case that a transition in conflict with other transitions is decomposed, these transitions should be arbitrated together as a group by the station controller, because if arbitration of the transitions are performed independently in separate subnets, the results may be inconsistent with the original rule of arbitration. On the other hand, arbitration of local transitions in conflict is performed by local machine controllers.

## 4 Implementation of Distributed Cooperative Control

The hierarchical and distributed control system composed of one station controller and several machine controllers has been implemented. The conceptual Petri net model is allocated to the Petri net based controller for management of the overall system. The detailed Petri net models are allocated to the Petri net based controllers in the machine controllers. Each machine controller directly monitors and controls the sensors and actuators of its machine. The control of the overall system is achieved by coordinating these Petri net based controllers. System coordination is performed through communication between the coordinator in the station controller and the Petri net based controllers in the machine controllers.

From the coordination algorithm, the actions of the coordinator and the Petri net based controller are mutually exclusive in time. Therefore hardware executing the coordinator program can afford to install the Petri net based controller algorithm. The station controller is composed of the Petri net based controller and the coordinator, and installs the global Petri net model, the Petri net engine, and the coordinator together. The overall system control is performed by the coordinator, which coordinates the Petri net models at the station level and the machine control level. The details of coordination in a two-level control system composed of a global controller and several local controllers have been implemented.

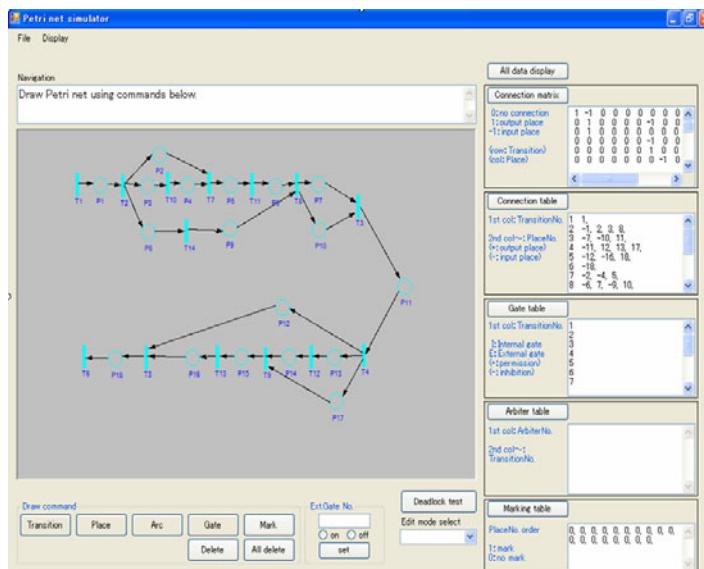
The overall control structure of an example robotic manufacturing system was implemented on a local area network of microcomputers. Each Petri net based local controller is implemented on a dedicated microcomputer (Renesas H8/3069) under the real-time OS ITRON 4.0. The local controller in charge of robot control executes robot motion control through the transmission of command and the reception of status report with serial interface to the real real robot controller. The upper level coordinator is implemented on another microcomputer. Communications among the controllers are performed using TCP/IP protocol. The coordinator sends the commands based on the conceptual, global Petri net model and coordinates the global transitions, which are accessed by the controllers as a shared file, such that decomposed transitions fire simultaneously.

The names of global transitions and their conflict relations are loaded into the coordinator in the station controller. The connection structure of a decomposed Petri net model and conflict relations among local transitions are loaded into the Petri net based controller in a machine controller. In the connection structure, a transition of a Petri net model is defined using the names of its input places and output places. Using the names of transitions in the subsystems, global transitions are defined; for example, G2: t0-2, t1-21, t2-22, t3-23 indicates that the global transition G2 is composed of the transition

no.2 of Station controller (subsystem no.0), the transition no.21 of Robot controller, the transition no.22 of MC controller (subsystem no.2), and the transition no.23 of Conveyor controller (subsystem no.3). Then, the coordinator information for the example distributed control system is as follows.

- G1: t0-1, t3-13
  - G2: t0-2, t1-21, t2-22, t3-23
  - G3: t1-71, t3-73
  - G4: t1-81, t2-82
  - G5: t0-3, t1-31, t2-32
  - G6: t0-4, t1-41, t2-42, t3-43
  - G7: t1-91, t2-92
  - G8: t0-5, t1-51, t3-53
  - G9: t0-6, t3-63

By executing the coordinator and Petri net based controllers algorithms based on loaded information, simulation experiments have been performed. The Petri net simulator initiates the execution of the subtasks attached to the fired transitions through the serial interface to the robot or other external machine. When a task ends its activity, it informs the simulator to proceed with the next activations by the external permissive gate arc. Experimental results show that the decomposed transitions fire at the same time as the original transition of the detailed Petri net of the whole system task. Firing transitions and marking of tokens can be directly observed on the display at each time sequence using the simulator [9] as shown in Fig.8.



**Fig. 8.** Example display of Petri net simulator at the conceptual level

## 5 Conclusions

A methodology to construct hierarchical and distributed control systems, which correspond to the hardware structure of manufacturing systems, has been proposed. By introduction of the coordinator, the Petri net based controllers are arranged according to the hierarchical and distributed nature of the manufacturing system. The coordination mechanism between the upper level and the lower level controllers is based on firability test of global transitions, and its software can be easily implemented using token game player. The overall control structure of an example robotic manufacturing system was implemented on a local area network of microcomputers using a real-time OS.

For large and complex manufacturing systems, the overall system is consistently controlled, such that a coordinator in a layer coordinates one-level lower Petri net based controllers and is coordinated by the one-level upper coordinator. The coordination mechanism is implemented in each layer repeatedly. The Petri net model in each Petri net based machine controller is not so large and easily manageable. The lower level controller provides a common interface for programming robots and other intelligent machines made by different manufacturers. The control method can be expanded to large and complex, discrete event manufacturing systems. Thus, modeling, simulation and control of large and complex manufacturing systems can be performed consistently using Petri nets.

## References

1. Thuriot, E., Valette, R., Courvoisier, M.: Implementation of a Centralized Synchronization Concept for Production Systems. In: Proceedings of IEEE Real-time Systems Symposium, pp. 163–171 (1983)
2. Jockovic, M., Vukobratovic, M., Ognjanovic, Z.: An Approach to the Modeling of the Highest Control Level of Flexible Manufacturing Cell. *Robotica* 8, 125–130 (1990)
3. Silva, M.: Petri Nets and Flexible Manufacturing. In: Rozenberg, G. (ed.) APN 1989. LNCS, vol. 424, pp. 374–417. Springer, Heidelberg (1990)
4. Taubner, D.: On the Implementation of Petri Nets. In: Rozenberg, G. (ed.) APN 1988. LNCS, vol. 340, pp. 418–439. Springer, Heidelberg (1988)
5. Reisig, W.: Petri Nets. Springer, Berlin (1985)
6. Hasegawa, K., Takahashi, K., Masuda, R., Ohno, H.: Proposal of Mark Flow Graph for Discrete System Control. *Trans. of SICE* 20, 122–129 (1984)
7. Murata, T.: Petri Nets: Properties, analysis and applications. *Proceedings of the IEEE* 77, 541–580 (1989)
8. Miyagi, P.E., Hasegawa, K., Takahashi, K.: A Programming Language for Discrete Event Production Systems Based on Production Flow Schema and Mark Flow Graph. *Trans. of SICE* 24, 183–190 (1988)
9. Yasuda, G.: Implementation of Distributed Cooperative Control for Industrial Robot Systems using Petri Nets. In: Preprints of the 9th IFAC Symposium on Robot Control (SYROCO 2009), pp. 433–438 (2009)

# A Graph Based Hybrid Approach of Offline Pre-planning and Online Re-planning for Efficient Assembly under Realtime Constraints

Daniel Ewert<sup>1</sup>, Sebastian Thelen<sup>1</sup>, Ralph Kunze<sup>1</sup>, Marcel Mayer<sup>2</sup>,  
Daniel Schilberg<sup>1</sup>, and Sabina Jeschke<sup>1</sup>

<sup>1</sup> RWTH Aachen University

Center for Learning and Knowledge Management and  
Institute of Information Management in Mechanical Engineering  
Dennewartstrasse 27, 52068 Aachen, Germany  
[jeschke.office@zlw-imam.rwth-aachen.de](mailto:jeschke.office@zlw-imam.rwth-aachen.de)

<sup>2</sup> RWTH Aachen University

Chair and Institute of Industrial Engineering and Ergonomics  
Bergdriesch 27, 52056 Aachen, Germany  
[info@ciaw.rwth-aachen.de](mailto:info@ciaw.rwth-aachen.de)

**Abstract.** Assembly tasks, e.g. the assembly of an automobile headlight, are a big challenge for nowadays planning systems. Depending on the problem domain, a planner has to deal with a huge number of objects which can be combined in several ways. Uncertainty about the outcome of actions and the availability of parts to be assembled even worsens the problem. As a result, classic approaches have shown to be of little use for reactive (online) planning during assembly, due to the huge computational complexity. The approach proposed in this paper bypasses this problem by calculating the complex planning problems, e.g. which parts must be mounted in which sequence, prior to the actual assembly. During assembly the precalculated solutions are then used to provide fast decisions allowing an efficient execution of the assembly. Within this paper this online planning combined with offline planning and the assessment of realtime constraints during assembly could be executed in the future will be described.

## 1 Introduction

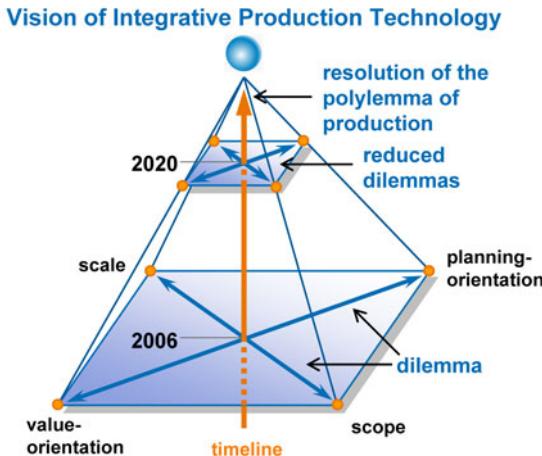
### 1.1 The Polylemma of Production Technology

In the last years, production in low-wage countries became popular with many companies by reason of low production costs. Hence, competitive production engineering is particularly important for high-wage countries such like Germany. Competition between manufacturing companies in high-wage and low-wage countries typically occurs within two dimensions: the production-oriented economy and the planning-oriented economy [1].

In the dimension of the production-oriented economy low-wage countries production compensates possible economic disadvantages within process times, factor consumption and process mastering by low productive factor cost. Companies in high-wage countries try to maximize the economies-of-scale by the usage of relatively expensive productivity factors. These disadvantages of relatively high-unit cost high-wage countries try to compensate through customizing and fast adaptations to market needs - the economies-of-scope. In addition the share of production within the value chain decreases more and more. Thus, the realizable economies-of-scale decreases, too. The escape into sophisticated niche markets is not promising, either.

Within the second dimension - the planning-oriented economy - companies in high-wages countries try to optimize processes with sophisticated, investment-intensive planning and production systems. Since processes and production systems do not reach their limit of optimal operating points, additional competitive disadvantages for high-wage countries emerge. In contrast, companies in low-wage countries implement simple, robust value-stream-oriented process chains [2].

The production industry in high-wage countries is confronted with these two dichotomies mentioned above. On the one hand there is the dilemma value orientation vs. planning orientation and on the other hand the dilemma scale vs. scope. These two dilemmas span the so called polylemma of production technology as shown in figure [1] [2].



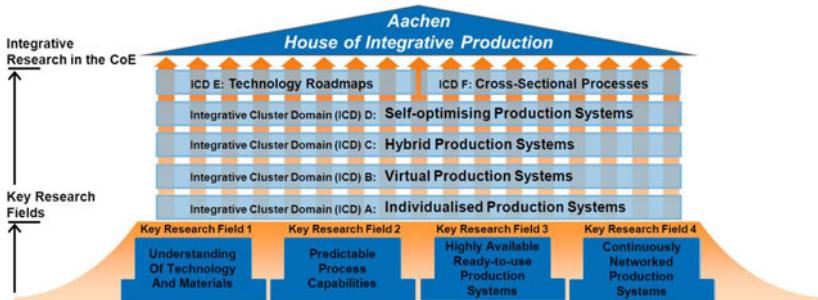
**Fig. 1.** The polylemma of production

## 1.2 Cluster of Excellence “Integrative Production Technology for High-Wage Countries”

The Cluster of Excellence “Integrative production technology for high-wage countries” unites several institutes of the RWTH Aachen University, which are dealing with production engineering. 19 professors of RWTH Aachen University,

who work in materials and production technology, as well as affiliated institutes such as the Fraunhofer Institutes, work in this cluster with the objective to contribute to the maintenance of production in high-wage countries which is relevant for the labor market. With regard to economy, focus is on products that address both niche markets and volume markets. The solution for the problems addressed sometimes require a fundamental new understanding of product and production technological contexts.

The Cluster of Excellence aims for the development of basics for sustainable production-scientific strategy and theory, as well as for the necessary technological approaches. Based on the scientific analysis of the polylemma of production technology, the individualisation, virtualisation and hybridisation, as well as self-optimisation of production technology were identified as areas in need for action. Therefore, four major fields of research have been defined: Individualised Production Systems, Virtual Production Systems, Hybrid Production Systems and Self-optimising Production Systems. Institutions and participants of the same field of research become strategically consolidated in so called Integrative Cluster Domains (ICD) related to their competences in production technology at RWTH Aachen University. These ICD form the Aachen House of Integrative Production (see figure 2).



**Fig. 2.** The Integrative House of Production

This paper deals with cognitive planning and control systems of production which are addressed in a sub-project of the ICD-D “Self-optimising Production Systems”.

## 2 The Subproject “Cognitive Planning and Control System for Production”

### 2.1 Objective

An advanced rationalization of production systems and processes is typical for high-wage countries. A significant challenge is to involve the implementation of

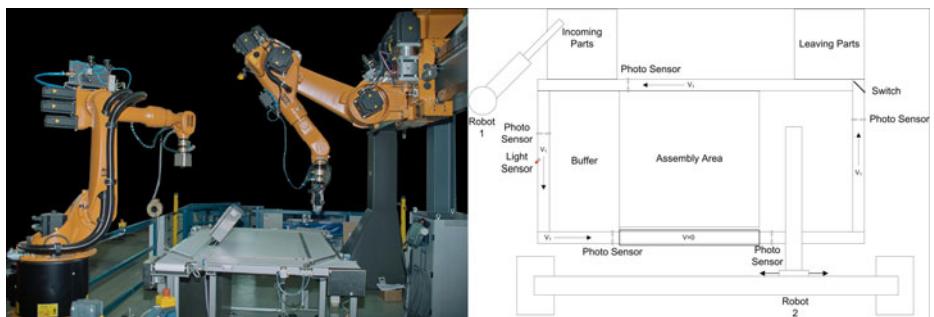
value stream-oriented approaches along with simultaneously increasing planning efficiency. A promising approach for the reduction of previous planning efforts consists of the development of a production system that is able to autonomously plan the production during a running process. Furthermore, such a system could autonomously react to changes in customer demands.

Within this sub-project a cognitive planning and control unit (CCU) is developed that aims to automate the assembly planning process, so that only a CAD description of the product which has to be fabricated is required as input to the system [3]. By means of this unit it will be possible to increase the flexibility of manufacturing and assembly systems and to decrease the planning effort in advance.

## 2.2 Approach

The task of the CCU is to plan and control the assembly of a product which is described solely by its CAD data. The system plans and executes the assembly autonomously, after receiving further descriptions, which are entered by a human operator. The CCU cooperates with human operators during the entire assembly process. While most of the assembly actions are executed by the assembly robot, certain tasks can only be accomplished by the operator.

The presented approach is evaluated within a scenario which is depicted in figure 3. The setup contains a robot cell with two robots, in which one robot (robot 2) is only controlled by the CCU [4], [5]. The other robot (robot 1) delivers separate parts for the final product in random sequence to a circulating conveyor belt. The CCU decides whether to pick up the delivered parts or to refuse them. If a part got picked up, the parts will be put into a buffer area or into the assembly area for immediate use.



**Fig. 3.** Photo and diagram of the robot cell

Due to the random block delivery future states of the system cannot be predicted. The CCU is therefore facing a non-deterministic planning problem, requiring either an online re-planning during the assembly, whenever a not

expected event occurs, or a plan in advance for all possible delivery sequences. Each of these strategies results in extensive computations, which lead either to slow responses during the assembly, or an unacceptable amount of pre-planning. The approach proposed in this paper therefore follows a hybrid approach, based on state graphs as described in the remainder of the paper.

### 3 Graph Based Planning under Realtime Constraints

#### 3.1 Related Work

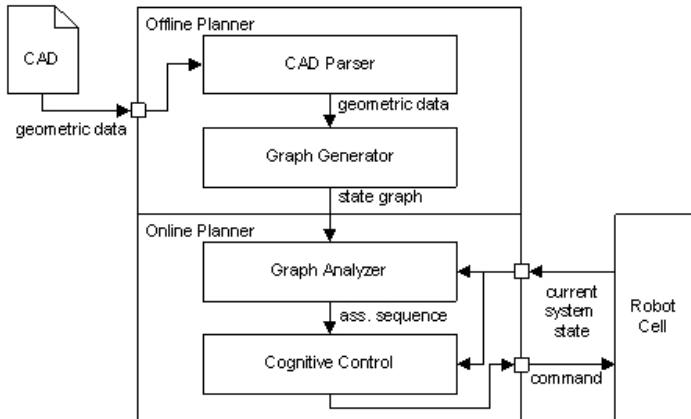
In the field of artificial intelligence, planning is of great interest. Several different approaches do exist, which are suitable for the application on planning problems. Hoffmann et al. developed the so called Fast-Forward Planning System, which is suitable to derive action sequences for given problems in deterministic domains [6]. Other planners are capable to deal with uncertainty [7], [8]. However, all these planners rely on a symbolic representation based on logic. Thus, the corresponding representations of geometric relations between objects and their transformations, which are needed for assembly planning, become very complex even for small tasks. As a result, these generic planners fail to compute any solution within acceptable periods of time.

Other planners have been designed especially for assembly planning. They directly work on geometric data to derive action sequences. A widely used approach is the Archimedes System [9] that uses AND/OR Graphs and an "Assembly by Disassembly" strategy to find optimal plans. U. Thomas follows the same strategy [10]. But where the Archimedes System relies on additional operator-provided data to find feasible sub-assemblies in his solution, Thomas uses geometric information of the final product as input only. However, both approaches are not capable of dealing with uncertainty.

#### 3.2 Basic Idea

To allow fast reaction times during assembly the planning process is separated into an offline part(the Offline Planner, described in [3.3]), executed prior to the assembly, and an online part(the Online Planner, described in [3.4]), executed in a loop during the assembly. The resulting system is drafted in Figure 4. While the Offline Planner is allowed computation times of up to several hours, the Online Planner's computation time must not exceed several seconds. The task of the Offline Planner is the precalculation of all feasible assembly sequences leading from the single parts to the desired product. The output is a graph representing these sequences, which is transmitted to the Online Planner. During assembly the Online Planner then repeatedly maps the current system state to a state contained in the graph. Following, it extracts an assembly sequence that transforms this state into a goal state containing the finished product.

This assembly sequence is then handed to the Cognitive Control component which triggers the accordant robot commands, interacts with the human operator



**Fig. 4.** System Overview of the CCU

and reacts to unforeseen changes in the state. Thereby, it is not bound to follow the given assembly, for example it can decide to invoke actions that move blocks from the conveyor belt to the buffer and vice versa instead of continuing an assembly. The details of the described process are discussed in the following sections in particular.

### 3.3 Offline Planning

#### Graph generation

During the offline planning phase, a solution space for the assembly sequence planning problem is derived from a product description. The current approach relies on a description of the assembled product's geometry and its constituting parts, possibly enriched with additional mating directions or mating operation specifications. Based on this information, an “assembly by disassembly” strategy is applied to ultimately generate a state graph that can be efficiently interpreted during online planning [11]. The main concept of this strategy is a recursive analysis of all possibilities of an assembly or subassembly, in which any assembly or subassembly is being separated into two further subassemblies until only nonseparable parts remain.

The geometric data is read by the CCU from a CAD file and stored in a tree data structure that is used to represent the product's assembly related properties. This assembly model tracks each atomic part that can not be disassembled further with a separate identity. These objects are reused in the assembly graph and reference all properties including the geometry of the represented part. To describe functional aspects like a thread or face of such a part, assembly feature instances can be used. This is relevant, if additional data beside the part geometries shall be taken into account by the assembly planner.

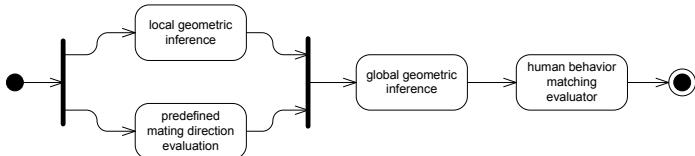
## Evaluation of assembly separations

While the CCU logically disassembles the complete assembly, multiple calculations have to be performed on each separation:

- The feasibility of the separation has to be estimated.
- The assembly plan for the resulting assembly operation has to be calculated.
- A static cost for the assembly operation has to be assigned, which later will be utilized by the online planner or the graph pruning algorithms.

These operations are performed by an evaluation pipeline built from multiple *separation evaluators*. This enables later additions of evaluation operations and encapsulates all evaluation done on a single separation. In [10] an algorithm for pure geometrical inference on assembly parts is depicted. This algorithm is used for evaluation, if a separation can be performed at all. Furthermore, it delivers a set of assembly directions from which the assembly plan for an operation can be generated. Following Kaufman et al. [9], the results from the geometric inference are additionally enriched with possible userdefined overridings that have shown to be necessary for complicated geometries like screws or threads. They provide a way to drastically enhance the resulting assembly plans.

To enable this mixing of geometric inference and predefined mating directions, the geometric analysis of a mating operation is separated into two steps. In the first step, only the geometries of assembly parts which change contact state when the operation is performed are taken into account. Even if this processing yields no viable mating directions, but there is at least one that is predefined by the user, the processing is carried on without directly discarding the separation in graph building. The second, global geometric inference takes the complete subassemblies of the separation into account. If the geometric inference from the first step labeled user defined mating directions as blocked the responsible geometric features are masked from each other to suppress their lokal blocking behavior.

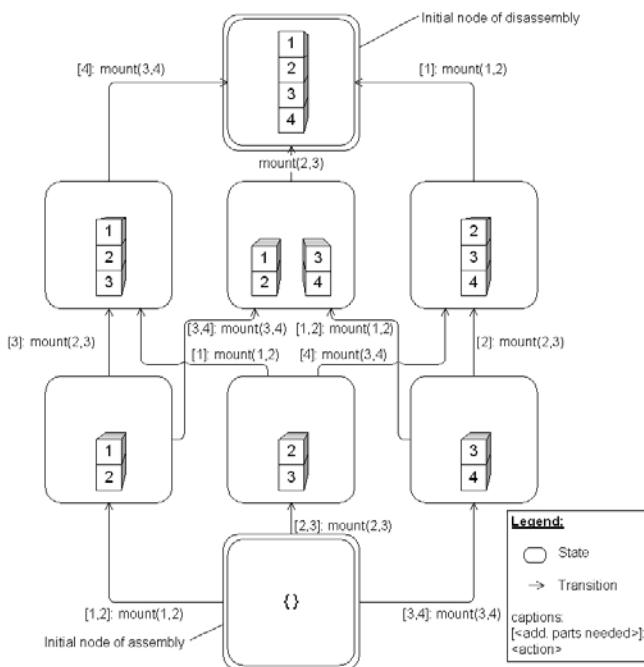


**Fig. 5.** The current computations performed to evaluate and rate a separation

The resulting evaluation pipeline is shown in figure 5. A separation is only excluded from the assembly graph if it is evaluated as being impossible to perform. This contrasts to other approaches like [10] and [12] where graph pruning heuristics are used to reduce computational demands and the utilized solution space. To enable the online planner to react upon unforeseen assembly states or events the complete solution space has to be processed. Though, high costs

might be assigned to separations that should be avoided, as is discussed later in this section.

Where actual online planning relies on a state graph representing the solution space, the graph builder initially constructs an AND/OR Graph representing all viable assembly operations and states, that might appear [11]. A state graph, as shown in figure 6, is then derived from that graph in a second step. This conversion is necessary, because the planning algorithm which is used during the online phase is not capable to handle the AND/OR Graph, which is a directed hypergraph.



**Fig. 6.** Example of a state graph representation

All data, which is generated by the separation evaluators for a separation, is stored in the AND/OR Graph's respective hyperarc. Currently, this information contains static operation costs and mating operation descriptions for each assembly, that might be active (e.g. moved) during the assembly step represented by the edge. The state graph is created in such a way, that its transitions/edges are directed in the opposite direction of the AND/OR Graph's hyperarc they originate from. Each state only contains the passive assembly to which other parts may be added. The availability of the active parts is encoded as condition for a transition which contains the operations static cost and mating operation description.

## Increasing machine transparency by mimicking human behaviour

Mayer et al. found that assemblies executed by human individuals follow - in most cases - simple rules [13]:

- Parts are placed next to each other. In other words, two separate subassemblies are not constructed in parallel.
- Structures are built layer by layer from bottom to top.
- Assembly begins at the boundary of the desired product.

During graph generation, each state is evaluated regarding its consistency with these (or similar) rules and receives an accordant rating. In the later phases of the online plan generation these ratings are taken into account and states with better ratings are preferred to other states. This strategy allows the CCU to mimic human behavior to a certain level.

### 3.4 Online planning

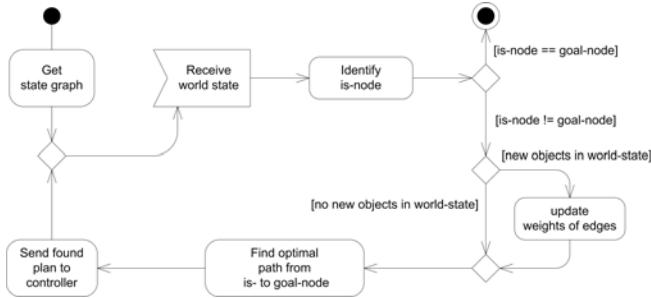
#### Responsibilities and Workflow of the Online Planner

The Online Planner is responsible for deriving plans during the assembly. Therefore it uses the state graph provided by the Offline Planner and current information about the robot cell's situation. This approach is similar to a system developed by [14], which reactively generates assembly task instructions for a human worker. The general process of the plan derivation from the state graph is depicted in figure 7.

The Graph Analyzer receives the generated state graph from the offline planning phase. It then has to pause until the Cognitive Control sends the actual world state describing, which contains the current situation of the assembly. The Graph Analyzer maps this world state onto the matching state contained in the state graph, now the “is-node”. If the is-node is identical to the goal-node, the node containing the completed assembly, the assembly has been completed and the process ends. Otherwise, the current world state is compared to the previously received world state. If new objects have occurred since the last run, the graph is being updated. This process is detailed in the section “update phase”. After the graph has been updated, the optimal path from the is- to the goal-node is calculated as described in section “sequence detection”. The found path represents the optimal assembly plan for the given current situation. It is sent to the Cognitive Control component, which then executes the assembly.

#### Update phase

In this phase, dynamic costs are assigned to the state graphs edges. The cost of an edge depends on the static costs of the assigned action and costs depending on the availability of required parts. Here, robotic actions allocate lower costs than actions that must be executed by the human operator. Actions that rely on parts which are currently not available receive penalty costs. Due to this weighting of actions the Online Planner prefers currently realizable assemblies before assemblies that rely on parts, which are currently not present. However,



**Fig. 7.** Workflow of the Online Part

reducing the penalty costs so that it becomes possible for a sum of realizable actions to outweigh the cost of one unrealizable action introduces speculation in the planning process: The algorithm could now choose to prefer a short but yet not realizable sequence over a realizable that needs much more actions to fabricate the final product. This decision could turn out to be better, if the missing part is delivered to the assembly before the previously not realizable action is to be executed. This behaviour can be facilitated even more by assigning penalty costs depending on the distance of the action in question to the current state. Considering time: The further away the action to be executed, the lower are the penalty costs it retrieves.

### Sequence detection

The algorithm identifies the node of the updated graph which is matching the current state of the assembly. This node becomes the initial node from which an optimal path to the final node has to be calculated. This is achieved using the A\* search algorithm [15]. A\* chooses nodes for expansion with a minimal value for  $f(x) = g(x) + h(x)$ , where  $h(x)$  is a heuristic for the distance from node  $x$  to the goal node, and  $g(x)$  represents the present path costs. Here,  $h(x)$  takes into account the number of not correctly mounted parts as well as the “performance” of the node during its evaluation as described in 3.3. The path costs  $g$  are calculated as the sum of the costs of the traversed edges plus the costs for every necessary tool change along that path. The alteration between actions executed by the human operator and actions done by a robot also counts as a tool change, resulting in a preference for assembly sequences where human involvement is concentrated in single time slots.

### 3.5 Realisation of the assembly

The actual assembly is realised by the Cognitive Control component. It is based on Soar, a cognitive framework for decision finding that aims on modelling the human decision process [16]. It is responsible to invoke accordant actions based on the current world state received from the robot cell and the assembly sequence

sent by the Graph Analyzer. The default behavior of the Cognitive Control is to realize the given assembly sequence by invoking the accordant robot command or instructing the operator. In case of unforeseen changes to the world state, due to a delivery of an object, it decides whether to discard or store the new object, depending on if the object can be used in later assembly steps. In such a situation the Cognitive Control can request a new plan from the Graph Analyzer which regards the new object. The component is also responsible for dealing with failure, e.g. failed robot actions or similar.

## 4 Outlook

Empiric studies need to show the capability of the described approach to accomplish feasible assembly sequences for different products of varying complexity. It is to be determined, which weightings of realizable and not realizable actions lead to optimal results. Furthermore, research on how much graph pruning influences the average computation time is required, and if the gain in time is worth the risk of missing an optimal assembly. Additionally, the scalability and applicability of the approach to industrial use cases has to be examined.

## 5 Conclusion

The polylemma of production can be alleviated by incorporating automatic planning systems. Depending on the assembly setup, these systems have to face non-deterministic behaviour during the assembly, which increases the computational complexity by orders of magnitude. The presented approach follows a hybrid approach of offline pre-planning and online re-planning to deal with this problem under realtime constraints. The actual planning problem is solved prior to the assembly by generating a state graph that describes all possible assemblies of the intended product. This graph is updated during the assembly and at this time optimal assembly sequence is extracted. This process is by far less time consuming and allows a fluent assembly. By taking into account the human behaviour during assembly it is possible to speed up the graph generation. Situations during an assembly that would not occur during an assembly executed by humans are pruned from the graph. This results in reduced computation times and increased machine transparency by meeting human expectancy.

## Acknowledgement

The authors would like to thank the German Research Foundation DFG for supporting the research on human-robot cooperation within the Cluster of Excellence "Integrative Production Technology for High-Wage Countries".

## References

1. Brecher, C., et al.: Excellence in Production. Apprimus Verlag (2007)
2. Klocke, F.: Production technology in high-wage countries – From ideas of today to products of tomorrow. *Industrial Engineering and Ergonomics*, 13–30 (2009)
3. Hauck, E., Gramatke, A., Isenhardt, I.: Einsatz Kognitiver Technischer Systeme im Produktionsumfeld. In: AUTOMATION 2008, Lösungen für die Zukunft. VDI Berichte, Band 2032, 305–308. VDI Verlag GmbH (2008)
4. Kempf, T., Herfs, W., Brecher, C.: Soar-based sequence control for a flexible assembly cell. In: ETFA 2009: Proceedings of the 14th IEEE international conference on Emerging technologies & factory automation, Piscataway, NJ, USA, pp. 488–496. IEEE Press, Los Alamitos (2009)
5. Hauck, E., Gramatke, A., Henning, K.: A Software Architecture for Cognitive Technical Systems Suitable for an Assembly Task in a Production Environment. In: AUTOMATION CONTROL – Theory and Practic, pp. 13–28. INTECH (2009)
6. Hoffmann, J.: FF: the fast-forward planning system. *The AI Magazine* 22 (2001)
7. Hoffmann, J., Brafman, R.: Contingent planning via heuristic forward search with implicit belief states. In: Proceedings of ICAPS 2005, pp. 71–80 (2005)
8. Castellini, C., Giunchiglia, E., Tacchella, A., Tacchella, O.: Improvements to sat-based conformant planning. In: Proc. of 6th European Conference on Planning, ECP 2001 (2001)
9. Kaufman, S.G., Wilson, R.H., Calton, R.E.J.T.L., Ames, A.L.: LDRD final report: Automated planning and programming of assembly of fully 3d mechanisms. Technical Report SAND96-0433, Sandia National Laboratorie (1996)
10. Thomas, U.: Automatisierte Programmierung von Robotern für Montageaufgaben. Fortschritte in der Robotik, vol. 13. Shaker Verlag, Aachen (2008)
11. Homem de Mello, L.S., Sanderson, A.C.: Representations of mechanical assembly sequences. *IEEE Transactions on Robotics and Automation* 7(2), 211–227 (1991)
12. Deshmukh, A., Yung, J., Wang, H.P.: Automated generation of assembly sequence based on geometric andfunctional reasoning. *Journal of Intelligent Manufacturing* 4(4), 269–284 (1993)
13. Mayer, M., Odenthal, B., Faber, M., Kabuš, W., Kausch, B., Schlick, C.: Simulation of human cognition in self-optimizing assembly systems. In: Proceedings of the IEA 2009 – 17th World Congress on Ergonomics (2009)
14. Zaeh, M.F., Wiesbeck, M.: A Model for Adaptively Generating Assembly Instructions Using State-based Graph. In: Manufacturing Systems and Technologies for the New Frontier, pp. 195–198. Springer, Heidelberg (2008)
15. Hart, P.E., Nilsson, N.J., Raphael, B.: A formal basis for the heuristic determination of minimum cost paths. *Systems Science and Cybernetics. IEEE Transactions on Systems Science and Cybernetics* 4(2), 100–107 (2007)
16. Laird, J.E., Newell, A., Rosenbloom, P.S.: Soar: An architecture for general intelligence. *Artif. Intell.* 33(1), 1–64 (1987)

# Improved Inertial Pose Estimation Algorithms for On-Site Hydraulic Turbine Blade Repairing Robot

Xiande Ma, Qiang Chen, Zhenguo Sun, and Wenzeng Zhang

Key Laboratory for Advanced Materials Processing Technology, Ministry of Education;  
Dept. of Mechanical Engineering, Tsinghua University, Beijing 100084, China  
mxdmxd@gmail.com

**Abstract.** For pose estimation of the mobile robots repairing the hydraulic turbine blades, due to almost full pose range, there is the problem of degenerated pose matrix and singular solution of pose angles. A better approach to measure the pose errors of mobile robots using quaternion is suggested, called One-Rotational Pose Error (ORPE). Based on ORPE, the accumulated errors resulted from long-term use of the low-precision inertial system are studied, and the inclinometers are used to adjust the pose estimation. Theoretically, only inclinometers cannot adjust all the pose angles or the pose matrix. However, the Frobenius norm of the matrix or ORPE can be used to give the best adjustment. This paper deduces an interesting result, and simulation results not only testify the validity of the approaches mentioned above, but also show that the combination of low-precision inertial system and low-precision inclinometers can give rise to relatively high precision pose estimation.

**Keywords:** Mobile robot, pose estimation, pose adjustment, turbine blade, inclinometer, quaternion.

## 1 Introduction

Hydro energy is a very important part of Chinese energy system. However, serious ablation of the giant turbine blades requires periodic repairing jobs. It is becoming more and more realistic in recent years to do the repairing jobs using mobile robots to replace human beings. This technology can not only save a lot of human resource, but also make it possible that the entire repairing job is done on-site, eliminating the great difficulty and cost in traditional methods to disassemble the whole giant turbines and move them out of the foundation pit [1].

In the robot-based on-site repairing system for giant hydraulic turbine blades, welding, polishing and gouging are the main tasks of the mobile robots. Pose estimation for the mobile robot is a key technology in this system, since it deeply affects the welding, polishing and gouging parameters during the repairing job, and thus greatly affects the control strategy and parameters eventually. The robot's autonomous control is possible only when it has the knowledge of its own pose.

Schemes for pose estimation include dead reckoning, inertial navigating, visual odometry, and so forth. Inertial system has robust performance and strong adaptability in industrial environment, immune to electromagnetic disturb, mechanical vibration,

wheel slippage, recession of the ground surface, et al. Thus it is used in this paper for pose estimation of the mobile robots.

Pose estimation for these mobile robots is under several restrictions.

A. Wide pose range. The turbine blades are in a complex shape, and the locations on the blade surface to be repaired are uncertain, thus the robots' pose can cover a wide range.

B. Low-precision inertial sensors are used in this paper, and the inclinometers are used in order to get better estimation.

By inertial algorithm, the pose matrix is firstly obtained; however, when the Euler pose angles are computed from the pose matrix, sometimes the singular state emerges. Better measurement for pose errors is necessary.

Euler angles and the quaternion are two important frameworks of pose description for rigid bodies, not restricted in mobile robots. Euler angles are usually used to describe the pose of mobile robots [2-9], especially in the field of pose control. However, there is the singular problem in the Euler angle description. The quaternion is an indirect description of the robotic pose, defined to describe spatial rotations between two poses [4-9]. In the inertial navigation algorithm, the quaternion is an important tool to solve the differential equations of the angular motion. The quaternion is not a direct description of the pose, but it does not have the singular problem. When the pose error is the focus, the quaternion becomes very useful.

Literature [4] introduces the quaternion and gives a detail comparison between spatial transforms and the quaternion. Literature [5] discusses in detail the characteristic of the quaternion and how quaternion transformations represent robotic motion and states. Literature [6] uses quaternion to measure "difference" between rotations. Similar to [6], this paper uses the quaternion as a better measurement of pose errors.

To diminish the accumulated errors of the inertial measurement, the inclinometers are used to get adjustments. However, only inclinometers themselves cannot give all the three Euler angles and the pose matrix. Approaches for adjustment of the pose estimation using both the inclinometers and the pose obtained by inertial sensors should be established.

This paper is organized as follows. Section 2 briefly reviews the definition of pose angles and pose matrix, and introduces the simulation platform for the algorithms discussed below. Section 3 discusses the singular state of Euler pose angle calculation and proposes a better method to measure pose errors. Section 4 focuses on the method to merge the information from inertial sensors and inclinometers, in order to achieve more precise pose estimation. Section 5 shows the simulation results and corresponding analysis. And section 6 concludes the whole paper.

## 2 Pose Descriptions and Simulation Model

**Definition of Pose Angles and Pose Matrix.** Local geography coordinate system is built with its x-axis pointing east, y-axis pointing north, and z-axis pointing up. And robot coordinate system is usually built with its x-axis pointing right (relative to the robot itself), y-axis pointing forward, and z-axis pointing up. The relationship between these two coordinate systems decides the pose of the robot, usually in the form of pose matrix or Euler angles. [2] [3]

The pose of the mobile robot is indicated as Euler angles, i.e. pitch ( $\theta$ ), roll ( $\gamma$ ) and yaw ( $\psi$ ), whose domain defined by

$$\psi \in (-\pi, \pi], \theta \in (-\pi/2, \pi/2], \gamma \in (-\pi, \pi].$$

The transform matrix from local geography coordinate system to robot coordinate system is shown as Formula (1), called the pose matrix[2].

$$\mathbf{C}_n^b = \begin{bmatrix} \cos \gamma \cos \psi + \sin \gamma \sin \psi \sin \theta & -\cos \gamma \sin \psi + \sin \gamma \cos \psi \sin \theta & -\sin \gamma \cos \theta \\ \sin \psi \cos \theta & \cos \psi \cos \theta & \sin \theta \\ \sin \gamma \cos \psi - \cos \gamma \sin \psi \sin \theta & -\sin \gamma \sin \psi - \cos \gamma \cos \psi \sin \theta & \cos \gamma \cos \theta \end{bmatrix}. \quad (1)$$

For any given point  $\mathbf{p}$ , suppose its coordinates in local geography coordinate system are designated as vector  $\mathbf{p}_n$ , whose counterpart in robot coordinate system is  $\mathbf{p}_b$ , then

$$\mathbf{p}_b = \mathbf{C}_n^b \mathbf{p}_n. \quad (2)$$

In the process of pose estimation, the pose matrix is obtained as formula (3), by inertial navigation algorithm.

$$\mathbf{C}_n^b = \begin{bmatrix} T_{11} & T_{21} & T_{31} \\ T_{12} & T_{22} & T_{32} \\ T_{13} & T_{23} & T_{33} \end{bmatrix}. \quad (3)$$

The simplest solution to the Euler pose angle is shown in formula (4)-(6).

$$\theta = \arcsin(T_{32}). \quad (4)$$

$$\gamma = \arctan2(-T_{31}, T_{33}). \quad (5)$$

$$\psi = \arctan2(T_{12}, T_{22}). \quad (6)$$

Where  $\arctan2(y, x)$  function is defined as Tab.1.

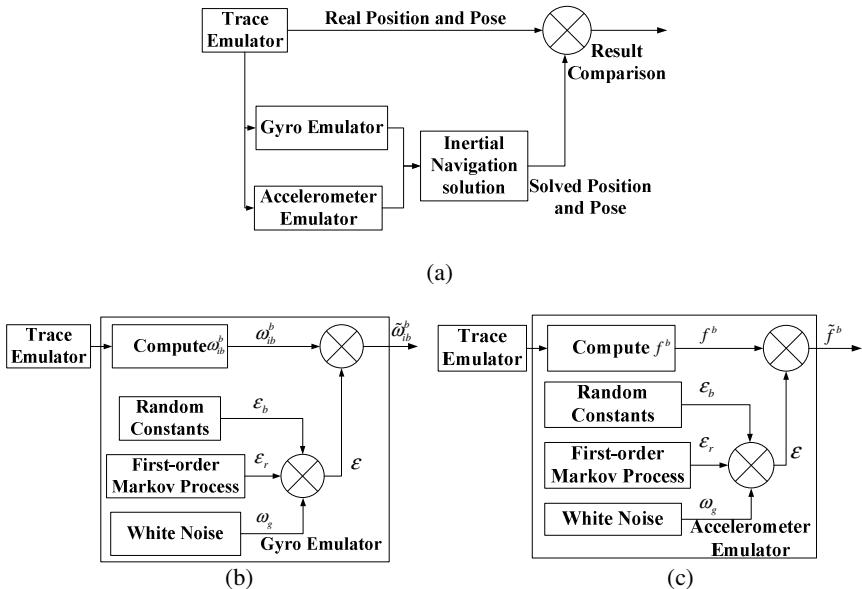
**Table 1.** Definition of  $\arctan2(y, x)$  function

condition	$x > 0$	$x < 0, y \geq 0$	$x < 0, y < 0$	$x = 0, y > 0$	$x = 0, y < 0$	$x = 0, y = 0$
$\arctan2(y, x)$	$\arctan(y/x)$	$\arctan(y/x) + \pi$	$\arctan(y/x) - \pi$	$\pi/2$	$-\pi/2$	0

**Simulation Model of Mobile Robot Motion and Inertial System.** A simulation model is built, as shown in Fig. 1. The simulation conditions are:

- 1) The motion of the robot centroid is independent of the robotic angular motion.
- 2) The motion of the robot centroid is random, and its three components in the east, the north and the up direction are independent to each other, each of which has a normal distribution with zero-mean and a standard deviation of 0.02m/s (the maximal velocity of the robot's wheels is approximately 0.03m/s).

- 3) The robotic angular motion is also random, and its three components of pitch, roll and yaw angular velocities are independent to each other, each of which has a normal distribution with zero-mean and a standard deviation of 0.2rad/s (the maximal angular velocity is approximately 0.25rad/s).
- 4) The drift of the gyro is 0.17deg/hr, and the drift of the accelerometer is  $3.5 \times 10^{-3} \text{m/s}^2$  (ADIS16355, MEMS IMU, Analog Device. Inc).
- 5) The error of the inclinometer conforms to normal distribution with the standard deviation of 1degree (ADIS16203, MEMS inclinometer, Analog Device. Inc).



**Fig. 1.** Principle of the simulation for strap down inertial navigation algorithm. (a) Principle of the strap down inertial algorithm. (b) Principle of the gyro emulator. (c) Principle of the accelerometer emulator.

### 3 Instability of Pose Solution and the Criterion of Measuring Pose Errors

#### 3.1 Instability of Pose Solution

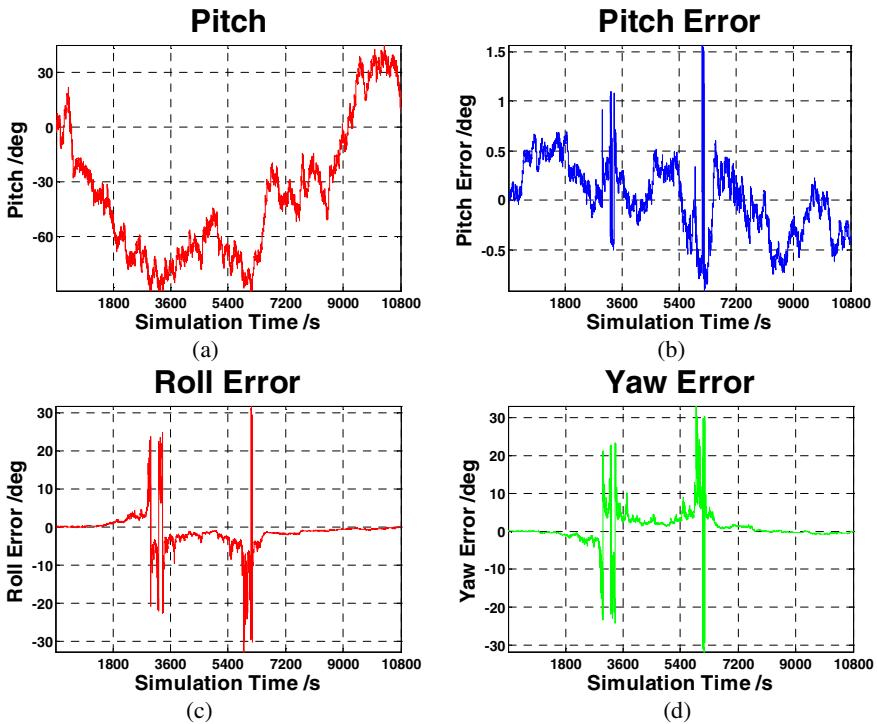
The solution shown in formula (4)-(6) works quite well in the normal cases, but sometimes the singular state emerges, as shown in Fig.2.

This is caused by the Euler angle definition. When  $\theta \approx \pm\pi/2$  the first rotation and the third rotation are around nearly the same axis, thus roll and yaw angle are inseparable. In fact, we have

$$\theta \approx \frac{\pi}{2} \Rightarrow \mathbf{C}_n^b \approx \begin{bmatrix} \cos(\gamma-\psi) & \sin(\gamma-\psi) & 0 \\ 0 & 0 & 1 \\ \sin(\gamma-\psi) & -\cos(\gamma-\psi) & 0 \end{bmatrix}, \theta \approx -\frac{\pi}{2} \Rightarrow \mathbf{C}_n^b \approx \begin{bmatrix} \cos(\gamma+\psi) & -\sin(\gamma+\psi) & 0 \\ 0 & 0 & 1 \\ \sin(\gamma+\psi) & \cos(\gamma+\psi) & 0 \end{bmatrix}$$

Here the matrix is obviously degenerated, and only  $\gamma - \psi$  or  $\gamma + \psi$  can be obtained. Another explanation is that in this case, formula (5) and (6) use trivial numbers to compute roll and yaw angles, small absolute errors of  $T_{31}$ ,  $T_{33}$ ,  $T_{12}$ ,  $T_{22}$  can cause very big relative errors of roll or yaw angle.

The singular problem is a kind of “physical” singularity. Thus it needs to find better ways to demonstrate the pose and pose error of the robot.



**Fig. 2.** Pose errors when the pitch angle is abnormal. (a) true pitch value. (b) pitch error. (c) roll error. (d) yaw error.

### 3.2 Use Quaternion to Measure Pose Error

It can be proved that the pose of a coordinate system relative to another coordinate system is equivalent to a single rotation around an equivalent axis [2]. Suppose the unit vector of the equivalent axis is  $\mathbf{u}$ , and the rotation angle is  $\alpha$ ,  $0 \leq \alpha \leq \pi$ , with the direction of rotation and the direction of  $\mathbf{u}$  conforming to right-hand rule. Then the relative pose can be uniquely determined by  $\mathbf{u}$  and  $\alpha$ , except that  $\alpha = \pi$  corresponding to both  $\mathbf{u}$  and  $-\mathbf{u}$ . A quaternion  $\mathbf{Q}$  can be constructed as a result.

$$\mathbf{Q} = \cos \frac{\alpha}{2} + \mathbf{u} \sin \frac{\alpha}{2} = q_0 + q_1 i + q_2 j + q_3 k. \quad (7)$$

With this definition, we easily have

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1. \quad (8)$$

$$q_0 = \cos(\alpha/2).$$

The relationship between the quaternion and the relative pose matrix is demonstrated as formula (9).

$$\mathbf{C} = \begin{bmatrix} 1 - 2(q_2^2 + q_3^2) & 2(q_1 q_2 + q_0 q_3) & 2(q_1 q_3 - q_0 q_2) \\ 2(q_1 q_2 - q_0 q_3) & 1 - 2(q_1^2 + q_3^2) & 2(q_2 q_3 + q_0 q_1) \\ 2(q_1 q_3 + q_0 q_2) & 2(q_2 q_3 - q_0 q_1) & 1 - 2(q_1^2 + q_2^2) \end{bmatrix}. \quad (9)$$

If  $\mathbf{C}$  is already known, then

$$q_0^2 = \frac{1}{4}(1 + \text{tr}(\mathbf{C})).$$

Where  $\text{tr}(\mathbf{C})$  is the trace of the matrix  $\mathbf{C}$ . So we can get

$$\cos \alpha = \frac{1}{2}(\text{tr}(\mathbf{C}) - 1). \quad (10)$$

Meanwhile  $0 \leq \alpha \leq \pi$ , thus

$$\alpha = \arccos\left(\frac{1}{2}(\text{tr}(\mathbf{C}) - 1)\right). \quad (11)$$

The definition of quaternion immediately inspires that the rotation angle  $\alpha$  can be used to measure pose errors. Suppose  $\mathbf{C}_0$  is the true transform matrix, and  $\mathbf{C}_1$  is the obtained transform matrix corresponding to  $\mathbf{C}_0$  by pose estimation algorithm. Let

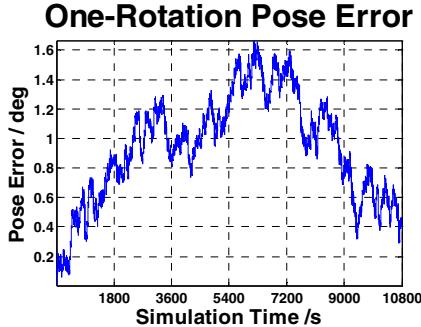
$$\mathbf{C} = \mathbf{C}_1 \mathbf{C}_0^{-1} = \mathbf{C}_1 \mathbf{C}_0^T.$$

This means that  $\mathbf{C}$  is the transform matrix between  $\mathbf{C}_0$  and  $\mathbf{C}_1$ . Then angle

$$\alpha = \arccos\left(\frac{1}{2}(\text{tr}(\mathbf{C}_1 \mathbf{C}_0^T) - 1)\right).$$

Thus the rotation angle  $\alpha$  of the quaternion corresponding to  $\mathbf{C}$  indicates the pose error of  $\mathbf{C}_1$  compared to  $\mathbf{C}_0$ . In the text below, angle  $\alpha$  is also called one-rotation pose error, or ORPE. Fig.3 shows ORPE of the same pose matrix as Fig.2.

One obvious advantage of ORPE compared to Euler angle errors is that ORPE has a reasonable and simple physical significance. The other advantage comes from the property of the quaternion: there is no singular state for this ORPE.



**Fig. 3.** ORPE when the pitch angle is abnormal (derived from the same pose matrix as Fig.2)

## 4 Inclinometer Adjustment to the Pose of Mobile Robot

**Inclinometer.** An inclinometer can measure the angle between its axis and the gravity direction. This angle is an absolute value, which does not have cumulative errors. Inclinometer does not have good dynamic responds, but under its maximal frequency allowed, it is an ideal sensor to adjust the pose estimation made by the inertial system.

Assume that the adjusted Euler angles are  $\tilde{\psi}$ ,  $\tilde{\theta}$ ,  $\tilde{\gamma}$ , with

$$\tilde{\psi} \in (-\pi, \pi], \quad \tilde{\theta} \in [-\pi/2, \pi/2], \quad \tilde{\gamma} \in (-\pi, \pi].$$

And the adjusted pose matrix is shown in Formula (12).

$$\tilde{C}_n^b = \begin{bmatrix} \cos \tilde{\gamma} \cos \tilde{\psi} + \sin \tilde{\gamma} \sin \tilde{\psi} \sin \tilde{\theta} & -\cos \tilde{\gamma} \sin \tilde{\psi} + \sin \tilde{\gamma} \cos \tilde{\psi} \sin \tilde{\theta} & -\sin \tilde{\gamma} \cos \tilde{\theta} \\ \sin \tilde{\psi} \cos \tilde{\theta} & \cos \tilde{\psi} \cos \tilde{\theta} & \sin \tilde{\theta} \\ \sin \tilde{\gamma} \cos \tilde{\psi} - \cos \tilde{\gamma} \sin \tilde{\psi} \sin \tilde{\theta} & -\sin \tilde{\gamma} \sin \tilde{\psi} - \cos \tilde{\gamma} \cos \tilde{\psi} \sin \tilde{\theta} & \cos \tilde{\gamma} \cos \tilde{\theta} \end{bmatrix} \quad (12)$$

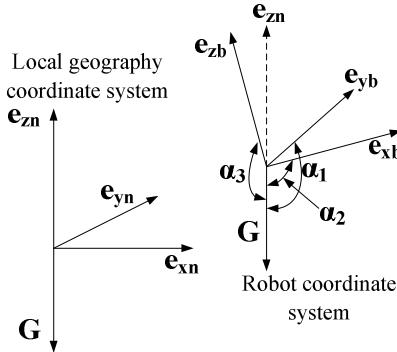
**Inclinometer Adjustment for Pitch and Roll Angle.** Install an inclinometer on the roll axis of the robot coordinate system, so that the angle between this axis and the gravity direction  $\alpha_1$ ,  $0 \leq \alpha_1 \leq \pi$ , can be obtained, as shown in Fig. 4.

By Fig. 4 we can easily get

$$\cos \alpha_1 = \mathbf{e}_{yb}^n \bullet (-\mathbf{e}_{zn}^n) = -(\tilde{C}_b^n \mathbf{e}_{yb}^b) \bullet \mathbf{e}_{zn}^n = -(\mathbf{e}_{zn}^n)^T \tilde{C}_b^n \mathbf{e}_{yb}^b.$$

Where the sign  $\bullet$  indicates inner product,  $\mathbf{e}_{zn}^n$  is the indication of vector  $\mathbf{e}_{zn}$  in local geography coordinate system, and  $\mathbf{e}_{yb}^b$  is the indication of vector  $\mathbf{e}_{yb}$  in robot coordinate system. Apparently,  $\mathbf{e}_{zn}^n = [0 \ 0 \ 1]^T$ ,  $\mathbf{e}_{yb}^b = [0 \ 1 \ 0]^T$ , thus

$$\cos \alpha_1 = -[0 \ 0 \ 1] \tilde{C}_b^n [0 \ 1 \ 0]^T = -\sin \tilde{\theta}. \quad (13)$$



**Fig. 4.** Measurement of inclinometers (all the vectors are unit vectors)

This yields

$$\tilde{\theta} = \alpha_1 - \pi/2. \quad (14)$$

Similarly, we have

$$\cos \alpha_2 = \sin \tilde{\gamma} \cos \tilde{\theta}. \quad (15)$$

$$\cos \alpha_3 = -\cos \tilde{\gamma} \cos \tilde{\theta}. \quad (16)$$

When  $\tilde{\theta} \in (-\pi/2, \pi/2)$ ,

$$\tilde{\gamma} = \arctan 2(\cos \alpha_2, -\cos \alpha_3). \quad (17)$$

**Inclinometer Adjustment for Yaw Angle Using ORPE.** Obviously, yaw angle cannot be obtained in a similar way. However, the robot moves quite slowly, and the adjust frequency can be raise to a reasonable level so that the difference between the adjusted pose matrix and the original pose matrix is trivial. Under such condition, there are two reasonable approaches to get the best adjustment of yaw angle and the pose matrix.

With the knowledge of the quaternion, another approach to adjust the yaw angle and the pose matrix using ORPE is natural.

Suppose  $C_\Delta$  to be the transform matrix between the two poses.

$$C_\Delta = \tilde{C}_n^b (C_n^b)^{-1} = \tilde{C}_n^b (C_n^b)^T. \quad (18)$$

$$\alpha_\Delta = \arccos \left( \frac{1}{2} (tr(C_\Delta) - 1) \right). \quad (19)$$

And let

$$\Delta\theta = \tilde{\theta} - \theta, \Delta\gamma = \tilde{\gamma} - \gamma, \Delta\psi = \tilde{\psi} - \psi$$

And all these three angles are normalized to  $(-\pi, \pi]$ .

Naturally, smaller  $\alpha_\Delta$  means that the two poses are closer. So the best  $\Delta\psi$  is considered to be the one minimizing  $\alpha_\Delta$ . Meanwhile, formula (19) indicates that  $\alpha_\Delta$  is a monotonically decreasing function of  $tr(\mathbf{C}_\Delta)$ . Thus, minimizing  $\alpha_\Delta$  is equivalent to maximizing  $tr(\mathbf{C}_\Delta)$ .

It can be deduced that

$$tr(\mathbf{C}_\Delta) = a \cos(\Delta\psi) + b \sin(\Delta\psi) + d. \quad (20)$$

Where

$$a = \cos \tilde{\theta} \cos \theta + \cos(\Delta\gamma) (1 + \sin \tilde{\theta} \sin \theta). \quad (21)$$

$$b = \sin(\Delta\gamma) (\sin \tilde{\theta} + \sin \theta). \quad (22)$$

$$d = \cos(\Delta\gamma) \cos \tilde{\theta} \cos \theta + \sin \tilde{\theta} \sin \theta. \quad (23)$$

Maximizing  $tr(\mathbf{C}_\Delta)$  is now equivalent to satisfying

$$\frac{\partial (tr(\mathbf{C}_\Delta))}{\partial (\Delta\psi)} = -a \sin(\Delta\psi) + b \cos(\Delta\psi) = 0. \quad (24)$$

$$\frac{\partial^2 (tr(\mathbf{C}_\Delta))}{(\partial (\Delta\psi))^2} = -a \cos(\Delta\psi) - b \sin(\Delta\psi) \leq 0. \quad (25)$$

Formula (21) and (22) yield

$$\Delta\psi = \arctan2(b, a). \quad (26)$$

It can be proved that as long as  $|\Delta\theta| \approx 0$  and  $|\Delta\gamma| \approx 0$ ,  $\Delta\psi$  gained by Formula (21)-(23) is stable, with a maximal absolute value similar to  $|\Delta\gamma|$ .

**Inclinometer Adjustment for Yaw Angle Using Frobenius Norm.** Another criterion to measure the similarity of two matrices is the Frobenius norm of the difference of these two matrices.

Let

$$\Delta\mathbf{C}_n^b = \tilde{\mathbf{C}}_n^b - \mathbf{C}_n^b.$$

The best  $\Delta\psi$  is now considered to be the one minimizing  $\|\Delta\mathbf{C}_n^b\|_F$ , the Frobenius norm of the matrix  $\Delta\mathbf{C}_n^b$ .

The Frobenius norm of a  $n \times n$  real matrix  $\mathbf{A}$  is defined by

$$F(\mathbf{A}) = \|\mathbf{A}\|_F = \left( \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \right)^{1/2}. \quad (27)$$

Where  $a_{ij}$  is the element in the  $i^{\text{th}}$  row and the  $j^{\text{th}}$  column of matrix  $\mathbf{A}$ .

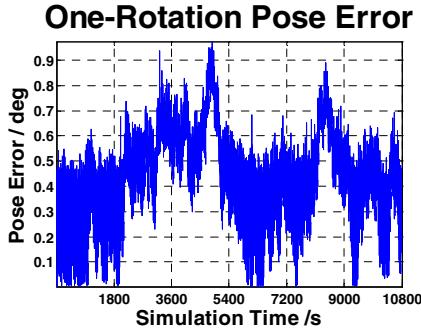
Obviously  $\|\Delta\mathbf{C}_n^b\|_F \geq 0$  is always true, minimizing  $\|\Delta\mathbf{C}_n^b\|_F$  means minimizing  $\|\Delta\mathbf{C}_n^b\|_F^2$ . It is interesting that, by careful deduction, we have

$$\left\| \Delta \mathbf{C}_n^b \right\|_F^2 = 6 - 2 \operatorname{tr}(\mathbf{C}_{\Delta}). \quad (28)$$

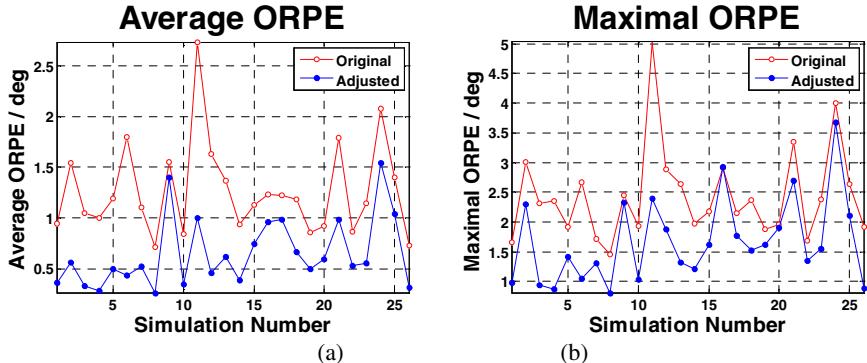
As a result, this approach based on Frobenius norm is equivalent to the above approach using the quaternion.

## 5 Simulation Results

Under the conditions mentioned in section.2, the simulation last for three hours with 200Hz sampling frequency. Fig.5 shows the simulation result with the same pose matrix as Fig.2. For each simulation, the average and maximal ORPE over simulation time seems to be good measurements of the effectiveness of different pose estimation algorithms. And Fig.6 shows the ORPE comparison of different data.



**Fig. 5.** ORPE after inclinometer adjustment (derived from the same pose matrix as Fig.2)



**Fig. 6.** ORPE comparison between the original algorithm and the one using inclinometer adjustment. (a) average ORPE comparison. (b) maximal ORPE comparison.

Fig.6 shows that most of the average pose errors of the adjustment algorithm are 0.5 deg less than their counterparts of the original algorithm. These simulation results show that the proposed approach to adjust the pose estimation is valid. It helps to improve the precision of the pose estimation with low-precision inertial sensors and inclinometers.

The inertial sensors and inclinometers used here are all of low-precision (see section 2), however, relatively high-precision pose estimation are now possible with proposed algorithms.

## 6 Conclusions

This paper focuses on the algorithms to get more precise pose estimation with long-term use of the low-precision inertial system and inclinometers for mobile robots. The method to solve the instability problem of Euler pose angle estimation is discussed, and a better criterion to measure pose errors using quaternion is proposed. Then the algorithms to adjust the pose estimation with inclinometers are proposed. A simulation model is established and the simulation results show that better precision can be achieved using these algorithms. This is meaningful for the autonomous control of the mobile robots in the robot-based on-site repairing system for giant hydraulic turbine blades.

## Acknowledgements

The authors gratefully acknowledge the financial support from the National High Technology Research and Development Program of China (No. 2007AA04Z258), and the National Natural Science Foundation of China (No. 50875147).

## References

1. Gui, Z.: Study on Wall Climbing Robot for Hydraulic Turbine Blade On-site Repair. Tsinghua University, Beijing (2006)
2. Titterton, D.H., Weston, J.L.: Strapdown Inertial Navigation Technology, 2nd edn., pp. 17–58. The Institution of Electrical Engineers, Stevenage (2004)
3. Wertz, J.R.: Spacecraft attitude determination and control. Reidel, Dordrecht (1978)
4. Funda, J., Paul, R.P.: A Comparison of Transforms and Quaternions in Robotics. In: Proc. IEEE Inter. Conf. on Robotics and Automation, vol. 2, pp. 886–891 (1988)
5. Tan, Q., Balchen, J.G.: General Quaternion Transformation Representation for Robotic Application. In: Proc. IEEE Inter. Conf. on Systems, Man and Cybernetics, Touquet, France, vol. 3, pp. 319–324 (1993)
6. Roberts, K.S.: Coordinating a Robot Arm and Multi-finger Hand Using the Quaternion Representation. In: Proc. IEEE Inter. Conf. on Robotics and Automation, Cincinnati, OH, vol. 2, pp. 1252–1257 (May 1990)
7. Isenberg, D.R., Kakad, Y.P.: Input and Output Transformations for a Space Robot Modeled with Quaternions. In: Proc. 41st Annual Southeastern Symposium on System Theory (SSST), University of Tennessee Space Institute Tullahoma, TN, USA, pp. 255–260 (March 2009)
8. Aydm, Y., Kucuk, S.: Quaternion Based Inverse Kinematics for Industrial Robot Manipulators with Euler Wrist. In: IEEE Inter. Conf. on Mechatronics (ICM), Budapest, Hungary, pp. 581–586 (July 2006)
9. Xian, B., de Queiroz, M.S., Dawson, D., et al.: Task-Space Tracking Control of Robot Manipulators via Quaternion Feedback. IEEE Transactions on Robotics and Automation 20(1), 160–167 (2004)

# Tool Path Generation Based on BCELTP for Maximizing Machining Strip Width

Hu Gong<sup>1,2</sup>

<sup>1</sup> Tianjin MicroNano Manufacturing Tech. Co., Ltd, TEDA, 300457, China

<sup>2</sup> State Key Laboratory of Precision Measuring Technology & Instruments

Centre of MicroNano Manufacturing Technology, Tianjin University, 300072, China

gonghu\_2005@yahoo.com

**Abstract.** In this paper, a new method is proposed to optimize the tool path of 5-axis machining free-form surface based on Basic Curvature Equations of Locally Tool Positioning (BCELTP) [Hu Gong, Li-Xin Cao, Jian Liu. Second order approximation of tool envelope surface for 5-axis machining with single point contact. Computer-Aided Design, 2008; 40: 604-15]. This method can be used to choose local optimum cutting direction with maximum machining strip width easily. Since BCELTP are accurate analytical expressions, the proposed method is accurate and effective in computation.

**Keywords:** Basic Curvature Equations of Locally Tool Positioning (BCELTP); 5-axis machining; free-form surface; Tool path; Optimization.

## 1 Introduction

As we know, tool path generation technology of 5-axis machining for non-ball end cutters are widely studied in academic and industrial fields, such as fillet end cutter, cylindrical cutter, flat-end cutter, and so on. It has been proved that they have many advantages compared to ball-end cutters. A very important advantage of them is to increase the machining strip width by using the asymmetric feature of non-ball end cutters. In other words, it is to reduce the machining errors for the given machining strip width.

A lot of papers have been published on optimizing the tool positions to increase machining strip width. According to the characteristics of the distribution of errors, these methods can be classified into two categories: local optimization and global optimization [1]. In this paper, we will focus on local optimization. C.G. Jensen et al. [2, 3] proposed “curvature match” method for finish surface machining using flat-end cutter. Further, based on this technique, an automatic cutter selection method for fillet-end cutter was proposed. N. Rao et al. [4, 5] proposed a Principal Axis Method (PAM) to machining concave surfaces with an alignment between the minimum principle curvature of the cutter and the maximum principle curvature of the designed

surface, with the feed in the principle direction corresponding to the minimum principle curvature of the designed surface. Y-S Lee presented a method to find the admissible tool orientation by considering both local and global surfaces for fillet end cutter [6]. Further, a non-isoparametric tool path generation method was proposed by evaluating machining strip width [7]. Chuang-Jang Chiou etc. presented the machining potential field (MPF) to find largest feasible machining strip width and the optimal cutting direction [8]. By using fillet-end cutter, Andrew Warkentin et al. [9-11] let fillet-end cutter be tangent to designed surface at more than one point. This method is called multi-point machining (MPM) method and only suitable for machining concave surface. By using Dupin indicatrix of conformity of the designed surface and the cutter surface, S. P. Radzevich proposed a tool path generation for all kinds of cutters [12-14].

A curvature catering method was proposed by X.C. Wang et al. [15] to machine free-form surface. A disc cutter with a concave is positioned to make the tool envelope surface has the same derivative as designed surface in the plane orthogonal to the cutting direction. A. Rao et al.[16] presented a mathematic expression of curvature of the tool envelope surface to determine local interference and determine optimal tool orientation angles. The two papers all consider tool envelope surface, which is the true machined surface. But, no matter using flat-end cutter or disc cutter, tool envelope surfaces are formed by trace of tool nose, which is a circle instead of the real surface. These methods can not be extended to other types of cutters. Joung-Hahn Yoon et al. [17] use second order approximation of the machining strip to optimize tool positions by using fillet-end cutter. An interesting point of view is that they point out that the Dupin indicatrix of tool envelope surface have double tangencies with the Dupin indicatrices of the designed surface and the cutter surface. But they don't give the calculation method of Dupin indicatrix of tool envelope surface. By using differential geometry theory, we presented Basic Curvature Equations of Locally Tool Positioning (BCELTP) [18] to describe the intrinsic relationship between second order approximation of tool envelope surface and many parameters at cutter contact point, which include cutting direction (tool path), tangential direction of grazing curve, principle curvatures and principle directions of cutter surface and designed surface. In addition, BCELTP was also used to prove the effectiveness of the existing methods. L. M. Zhu et al. [19] presented a method to generate 5-axis tool path for non-ball end cutter using third order point contact approach. More recently, we proposed a new method to optimize the tool positions for the given tool path based on BCELTP for maximizing the machining strip width [1]. In this paper, we will further explore the potential of BCELTP to discuss the tool path generation method for maximizing the machining strip width. The Principal Axis Method [4, 5] will be discussed from a different angle of tool envelope surface based on BCELTP.

## 2 Local Optimum Cutting Direction Based on BCELTP

In our previous paper [1], we have proposed two basic problems to minimize the relative normal curvature between the tool envelope surface and the designed surface based on BCELTP, which are expressed as follows:

Local optimization problem A:

$$\begin{cases} \min & k_{ortho}^{(01)}(\delta, \alpha, k_1^{(2)}, k_2^{(2)}, k_1^{(1)}, k_2^{(1)}) = \frac{K^{(21)}}{k_{c1}^{(21)}} \\ s.t. & K^{(21)} > 0; \quad H^{(21)} > 0 \quad \text{for } \delta \in [0, 2\pi] \end{cases} \quad (1)$$

Local optimization problem B:

$$\begin{cases} \min & k_{ortho}^{(01)}(\delta, \alpha, k_1^{(2)}, k_2^{(2)}, k_1^{(1)}, k_2^{(1)}) = K^{(21)} = 0 \\ s.t. & H^{(21)} \geq 0 \quad \text{for } \delta \in [0, 2\pi] \end{cases} \quad (2)$$

In this paper, in order to reduce the redundant content, we will use the same symbolic system as the paper [1]. Please refer to it for the detailed explanation. For convenience, we will use [1]-Eq.(X) to describe the equations in the paper [1].

Without loss of generality, assume that  $k_1^{(2)}, k_2^{(2)}, k_1^{(1)}, k_2^{(1)}$  are known for a tool position. There are two variables  $\delta$  and  $\alpha$  left. We can know that  $\theta = \delta + \alpha$  just denotes the cutting direction. Referring to [1]-Eq.(2), [1]-Eq.(3), [1]-Eq.(13) and [1]-Eq.(14), Eq.(1) can be simplified as follows:

$$\begin{cases} \min & k_{ortho}^{(01)}(\delta, \alpha) = \frac{K^{(21)}}{k_{c1}^{(21)}} \\ s.t. & k_{min}^{(2)} > k_{max}^{(1)} \end{cases} \quad (3)$$

Where

$$K^{(21)} = (k_1^{(2)} - k_1^{(1)})(k_2^{(2)} - k_2^{(1)}) + (k_2^{(1)} - k_1^{(1)})(k_1^{(2)} - k_2^{(2)}) \sin^2 \delta \quad (4)$$

$$\begin{aligned} k_{c1}^{(21)} &= k_1^{(2)} - k_1^{(1)} \\ &= k_1^{(2)} \cos^2 \alpha + k_2^{(2)} \sin^2 \alpha - k_1^{(1)} \cos^2(\delta + \alpha) - k_2^{(1)} \sin^2(\delta + \alpha) \end{aligned} \quad (5)$$

Obviously, Eq.(3) is a problem on the extremum of binary function. If we use the extremum theory to solve this problem, we need to differentiate Eq.(3) with respect to variables  $\delta$  and  $\alpha$ . It will become very complicated. In fact, we can do it in a simple way. Since  $K^{(21)} > 0$  and  $k_{c1}^{(21)} > 0$ , if we can get the minimum of  $K^{(21)}$  and the maximum of  $k_{c1}^{(21)}$  under the same conditions,  $k_{ortho}^{(01)}$  will reach minimum.

First, let's observe Eq.(4). Without loss of generality, we can assume

$$k_2^{(2)} > k_1^{(2)} > k_2^{(1)} > k_1^{(1)} \quad (6)$$

Eq.(6) leads to  $(k_1^{(2)} - k_1^{(1)})(k_2^{(2)} - k_2^{(1)}) > 0$  and  $(k_2^{(1)} - k_1^{(1)})(k_1^{(2)} - k_2^{(2)}) < 0$ . Therefore, only if  $\sin^2 \delta = 1$ ,  $K^{(21)}$  reach minimum.  $\sin^2 \delta = 1$  means that  $\delta = \pi/2$  or  $3\pi/2$  within  $[0, 2\pi]$ .

According to Eq.(3), when  $\alpha=\pi/2$  or  $3\pi/2$ ,  $k_{c1}^{(2)}$  reach maximum  $k_2^{(2)}$ . When  $\alpha=\pi/2$ (or  $3\pi/2$ ) and  $\delta=\pi/2$ (or  $3\pi/2$ ),  $k_{c1}^{(1)}$  reach minimum  $k_1^{(1)}$ . Therefore,  $k_{c1}^{(2)}$  reach the maximum  $k_2^{(2)} - k_1^{(1)}$ . According to above, we can get the following results:

When  $\alpha=\pi/2$ (or  $3\pi/2$ ) and  $\delta=\pi/2$ (or  $3\pi/2$ ),  $k_{ortho}^{(01)}$  reach the minimum and can be expressed as:

$$k_{ortho}^{(01)} \Big|_{min} = \frac{(k_1^{(2)} - k_1^{(1)})(k_2^{(2)} - k_2^{(1)}) + (k_2^{(1)} - k_1^{(1)})(k_1^{(2)} - k_2^{(2)})}{k_2^{(2)} - k_1^{(1)}} \quad (7)$$

Therefore, we can get optimum cutting  $\theta=\delta+\alpha=0$  (or  $\pi$ ) within  $[0, 2\pi]$ . Actually, the optimum cutting direction should be in the direction of  $e_1^{(1)}$  (or  $-e_1^{(1)}$ ).

Similarly, if  $k_1^{(2)} > k_2^{(2)} > k_2^{(1)} > k_1^{(1)}$  or  $k_1^{(2)} > k_2^{(2)} > k_1^{(1)} > k_2^{(1)}$ , we can get similar results. Referring to the paper [18], we can get the following results about optimum cutting direction:

For local optimization problem **A**, suppose  $k_{max}^{(2)} > k_{min}^{(2)} \geq k_{max}^{(1)} > k_{min}^{(1)}$ , where  $k_{max}^{(2)} = \max(k_1^{(2)}, k_2^{(2)})$ ,  $k_{min}^{(2)} = \min(k_1^{(2)}, k_2^{(2)})$ ,  $k_{max}^{(1)} = \max(k_1^{(1)}, k_2^{(1)})$  and  $k_{min}^{(1)} = \min(k_1^{(1)}, k_2^{(1)})$ . The four unit vectors  $e_{max}^{(2)}$ ,  $e_{min}^{(2)}$ ,  $e_{max}^{(1)}$  and  $e_{min}^{(1)}$  correspond to the four principal curvatures  $k_{max}^{(2)}$ ,  $k_{min}^{(2)}$ ,  $k_{max}^{(1)}$  and  $k_{min}^{(1)}$  respectively. The optimum cutting direction is in the direction of  $\pm e_{min}^{(1)}$ . In this situation, the tool position is determined by making  $e_{min}^{(2)}$  parallel to  $e_{max}^{(1)}$ . In fact, this conclusion includes the Principal Axis Method [4, 5]. But more clear understanding based on tool envelope surface is presented with limiting condition in this section.

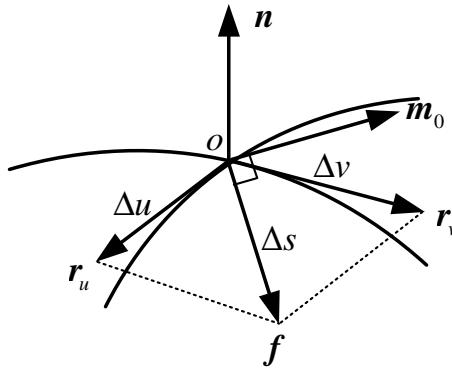
### 3 Tool Path Generation Based on the Optimum Cutting Direction

In this section, we will discuss how to get the optimized tool path using the local optimum cutting direction. As shown in Fig. 1,  $\mathbf{n}$  denotes the unit normal of the designed surface,  $\mathbf{f}$  denotes the optimum cutting direction,  $\mathbf{m}_0$  is perpendicular to  $\mathbf{f}$  and  $\mathbf{n}$ .  $\mathbf{f}$  and  $\mathbf{n}$  are all unit vectors. Suppose the designed surface can be written as  $\mathbf{r} = \mathbf{r}(u, v)$ , given as cutter contact point  $\mathbf{o}$ ,  $\mathbf{n}$  and  $\mathbf{f}$  can be calculated easily.  $\mathbf{m}_0$  also can written as follows:

$$\mathbf{f}(u, v) = \pm \mathbf{m}_0(u, v) \times \mathbf{n}(u, v) \quad (8)$$

Please note that  $\mathbf{f}$ ,  $\mathbf{n}$  and  $\mathbf{m}_0$  are functions with respect to the parameters  $u$  and  $v$  for the whole surface. Since the optimum tool path is a curve on the designed surface, it can be written as

$$\mathbf{r} = \mathbf{r}(u(s), v(s)) \quad (9)$$



**Fig. 1.** Local optimum cutting direction

$s$  denotes the arc length. Differentiating Eq.(9) with respect to  $s$ , we can get unit vector as follows:

$$\mathbf{r}_s = \mathbf{r}_u \frac{du}{ds} + \mathbf{r}_v \frac{dv}{ds} \quad (10)$$

$\mathbf{r}_s$  should be parallel to  $\mathbf{f}$ . Therefore, we have:

$$\begin{cases} \mathbf{r}_s \bullet \mathbf{m}_0 = 0 \\ \mathbf{r}_s \bullet \mathbf{f} = 1 \end{cases} \quad (11)$$

Substituting Eq.(10) into Eq.(11) and simplifying it, we have:

$$\begin{cases} (\mathbf{r}_u \bullet \mathbf{m}_0)du + (\mathbf{r}_v \bullet \mathbf{m}_0)dv = 0 \\ (\mathbf{r}_u \bullet \mathbf{f})du + (\mathbf{r}_v \bullet \mathbf{f})dv = ds \end{cases} \quad (12)$$

Solving Eq.(12) with respect to  $du$  and  $dv$

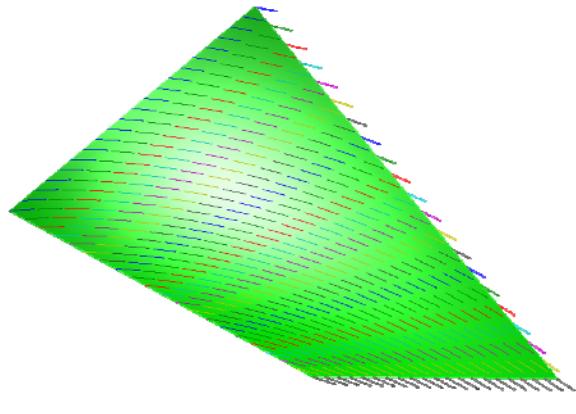
$$\begin{cases} du = \frac{\mathbf{r}_v \bullet \mathbf{m}_0}{(\mathbf{r}_u \times \mathbf{r}_v) \bullet \mathbf{n}} ds \\ dv = -\frac{\mathbf{r}_u \bullet \mathbf{m}_0}{(\mathbf{r}_u \times \mathbf{r}_v) \bullet \mathbf{n}} ds \end{cases} \quad (13)$$

Therefore, we can get the following calculating process of an optimum tool path:

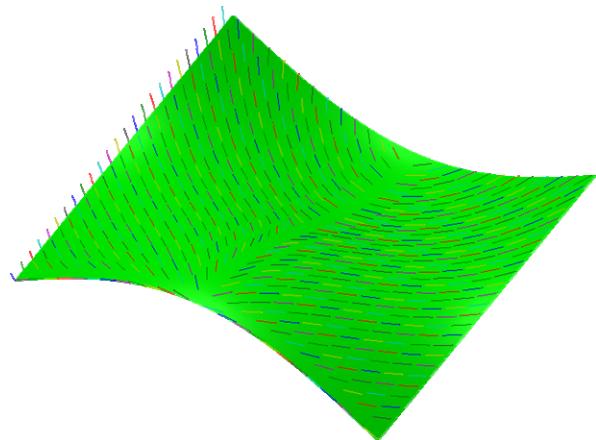
1. Given a designed surface  $\mathbf{r} = \mathbf{r}(u, v)$  and a border point  $\mathbf{r}_0 = \mathbf{r}(u_0, v_0)$  on it;
2. Calculate  $\mathbf{n}$ ,  $\mathbf{f}$  and  $\mathbf{m}_0$ ;
3. Given the small increment  $\Delta s$ , substitute it into Eq.(13), we can calculate two small increments  $\Delta u$ ,  $\Delta v$  as follows:

$$\begin{cases} \Delta u = \frac{\mathbf{r}_v \bullet \mathbf{m}_0}{(\mathbf{r}_u \times \mathbf{r}_v) \bullet \mathbf{n}} \Delta s \\ \Delta v = -\frac{\mathbf{r}_u \bullet \mathbf{m}_0}{(\mathbf{r}_u \times \mathbf{r}_v) \bullet \mathbf{n}} \Delta s \end{cases} \quad (14)$$

4. Calculate the next point of the tool path on the designed surface by using  $\mathbf{r} = \mathbf{r}(u + \Delta u, v + \Delta v)$ . Repeat step 1~4 until to the border of the designed surface.



(a) Case 1



(b) Case 2

**Fig. 2.** The optimized tool path based on local optimal cutting orientation

## 4 Case Study

Two examples are calculated using the methods above. In the first case, the designed surface is written as:

$$\begin{cases} x = 10(u+v) \\ y = 10(u-v) \\ z = -10uv \end{cases} \quad (u, v) \in [0, 1]^2$$

The cutter is a fillet-end cutter. Please refer to the paper [1] for the definition of the fillet-end cutter.  $r_1 = 1$   $r_2 = 2$ . The cutter contact point on the cutter surface is determined by using  $\phi = 30^\circ$ . The optimized tool path is shown in Fig.2(a). In the second case, the designed surface is written as

$$z = 0.2 \cos(y\pi) \sin(x\pi) \quad (x, y) \in [0, 1]^2$$

A flat end cutter is used, which is a special case of fillet-end cutter.  $r_1 = 0$   $r_2 = 0.01$  and  $\phi = 40^\circ$ . The optimized tool path is as shown in Fig.2(b). In the two cases, the initial points are determined using equal parameter interval. Once the initial point is determined, the corresponding tool path can be obtained uniquely. Therefore, we can not get the tool path with equal remained height using this method. The adjacent tool path should be determined using maximum remained height, which is restricted within the tolerance.

## 5 Conclusion

A new tool path generation method is proposed based on BCELTP for maximizing the machining strip width in this paper. By minimizing the relative normal curvature between the tool envelope surface and the designed surface, the optimum cutting direction is derived for the local optimization problem A [1]. From the angle of tool envelope surface, the effectiveness of the Principal Axis Method is also proved. This paper shows that BCELTP is good tool to analyze the relationship between the machining strip width and the cutting direction analytically.

**Acknowledgments.** This work was funded by Post Doctor Science Foundation of China (No. 20090460693). This support is gratefully acknowledged.

## References

1. Gong, H., Fang, F.Z., Hu, X.T., Cao, L.-X., Liu, J.: Optimization of tool positions locally based on the BCELTP for 5-axis machining of free-form surfaces. Computer-Aided Design 42, 558–570 (2010)
2. Jensen, C.G., Anderson, D.C.: Accurate tool placement and orientation for finished surface. Machining Journal Design and Manufacture 3, 251–261 (1993)
3. Jensen, C.G., Red, W.E., Pi, J.: Tool selection for five-axis curvature matched machining. Computer-Aided Design 34, 251–266 (2002)

4. Rao, N., Bedi, S., Buchal, R.: Implementation of the principal-axis method for machining of complex surfaces. *International Journal of Advanced Manufacturing Technology* 11, 249–257 (1996)
5. Rao, N., Ismail, F., Bedi, S.: Tool path planning for five-axis machining using the principal axis method. *International Journal of machine Tools and Manufacture* 37, 1025–1040 (1997)
6. Lee, Y.-S.: Admissible tool orientation control of gouging avoidance for 5-axis complex surface machining. *Computer-Aided Design* 29, 507–521 (1997)
7. Lee, Y.-S.: Non-isoparametric tool path planning by machining strip evaluation for 5-axis sculptured surface machining. *Computer-Aided Design* 30, 559–570 (1998)
8. Chiou, C.-J., Lee, Y.-S.: A machining potential field approach to tool path generation for multi-axis sculptured surface machining. *Computer-Aided Design* 34, 357–371 (2002)
9. Andrew Warkentin, A., Dedi, S., Ismail, F.: 5-axis milling of spherical surfaces. *International Journal of machine Tools and Manufacture* 36, 229–243 (1995)
10. Warkentin, A., Ismail, F., Bedi, S.: Intersection approach to multi-point machining of sculptured surfaces. *Computer Aided Geometric Design* 14, 567–584 (1998)
11. Warkentin, A., Ismail, F., Bedi, S.: Multi-point tool positioning strategy for 5-axis machining of sculptured surfaces. *Computer Aided Geometric Design* 17, 83–100 (2000)
12. Radzevich, S.P.: A closed-form solution to the problem of optimal tool-path generation for sculptured surface machining on multi-axis NC machine. *Mathematical and Computer Modelling* 43, 222–243 (2006)
13. Radzevich, S.P.: CAD/CAM of Sculptured surfaces on a multi-axis NC machine. Morgan & Claypool Publishers (2008)
14. Radzevich, S.P.: Kinematic geometry of surface machining. CPC Press (2008)
15. Wang, X.C., Ghosh, S.K., Li, Y.B., Wu, X.T.: Curvature catering-a new approach in manufacture of sculptured surfaces (part 1. theorem). *Journal of Materials Processing Technology* 38, 159–175 (1993)
16. Rao, A., Sarma, R.: On local gouging in five-axis sculptured surface machining using flat-end tools. *Computer-Aided Design* 32, 409–420 (2000)
17. Yoon, J.-H., Pottmann, H., Lee, Y.-S.: Locally optimal cutting positions for 5-axis sculptured surface machining. *Computer-Aided Design* 35, 69–81 (2003)
18. Gong, H., Cao, L.-X., Liu, J.: Second order approximation of tool envelope surface for 5-axis machining with single point contact. *Computer-Aided Design* 40, 604–615 (2008)
19. Zhu, L.-M., Zhu, H., Xiong, Y.L.: Third Order Point Contact Approach for Five-Axis Sculptured Surface Machining Using Non-ball End Tools -Part I: Third Order Approximation of Tool Envelope Surface. *Science in China-Ser. E* 53, 1904–1912 (2010)

# Dual-Quaternion-Based Variable Structure Control: A New Approach and Application

Hui Zhang<sup>1</sup>, Xiangke Wang<sup>1,2</sup>, and Dapeng Han<sup>3</sup>

<sup>1</sup> College of Mechatronics and Automation, National University of Defense Technology, Changsha 410073, China

<sup>2</sup> Research School of Information Sciences and Engineering, The Australian National University, Canberra ACT 2601, Australia

<sup>3</sup> College of Aerospace and Material Engineering, National University of Defense Technology, Changsha 410073, China  
xiangke.wang@anu.edu.au

**Abstract.** Dual quaternion is a more efficient tool without singularity to describe rigid-body transformation than other tools, such as homogeneous transformation matrix. In this paper, we propose a new variable structure control (VSC) based on unit dual quaternion Lie-group for the transformation of rigid-bodies in configuration space. We set up the dynamic model of the rigid-body motion represented by unit dual quaternion after the geometric structure of dual quaternion is revealed, and then propose a new unit-dual-quaternion-based VSC controller, including a new switching function and a new reaching law. At last, the proposed controller is specified into the configuration space of planar robot and applied with the dynamic model of the NuBot omni-directional mobile robot in simulation to show its effectiveness.

## 1 Introduction

Control in configuration space is a fundamental work of robotics and it is far more complex than pure translational control or rotational control. The conventional approaches are to control translational problem and rotational problem respectively [1,2], or using the Plücker coordinates [3], or to utilize the geometric structure of SE(3) [4]. Our approach in this paper is somewhat different. We make use of another mathematical tool, viz. unit dual quaternion, to represent rigid-body transformation and to design variable structure control (VSC) in configuration space.

Unit dual quaternion is a natural extension of unit quaternion. As we know, unit quaternion provides an efficient global representation for rotations without singularities, and unit dual quaternion inherits this property. In addition, unit dual quaternion owns obvious advantages to represent transformation (rotation and translation simultaneously) than other mathematical tools, such as homogeneous transformation matrix, quaternion/vector pairs, Lie-algebra and alike, because of its compactness and computational efficiency [5,6]. Dual quaternion has been applied in various fields such as Hand-Eye Calibration [7], Pose and

Motion Estimation [8], Robotics [9], Inertial Navigation [10]. The use of dual quaternion in rigid-body control has been reported in [11]. In this paper, deduction parallel (but not equivalent) to [4] has been discussed, resulting in geometric controller in kinematic designed with unit dual quaternion. We study further by relating unit dual quaternion to the dynamic control problems in this study. A generalized dynamic model represented by unit dual quaternion is proposed before VSC controller is adopted to design a new control law. Because unit dual quaternion group is not commutative, a new stable reaching law is proposed instead of the traditional reaching law of VSC. At last, we specify the proposed controller into the configuration space of planar robot, which is homeomorphic to  $\text{SE}(2)$ , and validate its effectiveness with NuBot omni-directional robot.

## 2 Mathematic Preliminaries

*Quaternion* is an extension of the complex number to  $\mathbb{R}^4$ . Formally, quaternion  $q$  can be defined as

$$q = a + bi + cj + dk \quad (1)$$

where  $i^2 = j^2 = k^2 = -1$ ,  $ij = k$ ,  $jk = i$ ,  $ki = j$ . A convenient shorthand notation is  $q = [s, \mathbf{v}]$ , where  $s$  is a scalar (called *scalar part*), and  $\mathbf{v}$  is a three-dimensional vector (called *vector part*). Obviously, a vector can be treated equivalently as a quaternion with vanishing real part, which is called *vector quaternion*. If  $a^2 + b^2 + c^2 + d^2 = 1$ , then quaternion  $q$  is called *unit quaternion*.

*Dual number* is defined as

$$\hat{a} = a + \epsilon b \quad \text{with} \quad \epsilon^2 = 0, \text{ but} \quad \epsilon \neq 0 \quad (2)$$

where  $a$  and  $b$  are real numbers, called *real part* and *dual part*, respectively, and  $\epsilon$  is nilpotent such as  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ . Dual numbers form an abelian group under addition [8]. The conjugate of the dual number  $\hat{a}$  is

$$\hat{a}^* = a - \epsilon b \quad (3)$$

*Dual vectors* are a generalization of dual numbers whose real and dual parts are both three-dimensional vectors. The *dot-product* between dual vectors, which is useful in our controller design, is defined as follows.

**Definition 1.** Let  $\hat{\mathbf{v}} = \mathbf{v}_r + \epsilon \mathbf{v}_d$  and  $\hat{\mathbf{k}} = \mathbf{k}_r + \epsilon \mathbf{k}_d = (k_{r1}, k_{r2}, k_{r3})^T + \epsilon(k_{d1}, k_{d2}, k_{d3})^T$  be dual vectors, then

$$\hat{\mathbf{k}} \cdot \hat{\mathbf{v}} = K_r \mathbf{v}_r + \epsilon K_d \mathbf{v}_d \quad (4)$$

where  $K_r$  and  $K_d$  are both  $3 \times 3$  diagonal matrices with diagonal entries  $k_{r1}, k_{r2}, k_{r3}$  or  $k_{d1}, k_{d2}, k_{d3}$ , namely  $K_r = \text{diag}(k_{r1}, k_{r2}, k_{r3})$  and  $K_d = \text{diag}(k_{d1}, k_{d2}, k_{d3})$ .

*Dual quaternion* is the quaternion with dual number components, i.e.  $\hat{q} = [\hat{s}, \hat{\mathbf{v}}]$ , where  $\hat{s}$  is a dual number and  $\hat{\mathbf{v}}$  is a dual vector. Obviously, a dual vector can

be treated equivalently as a dual quaternion with vanishing real part, which is called *dual vector quaternion*. A dual quaternion can be rewritten into  $\hat{q} = q_r + \epsilon q_d$ , where  $q_r$  and  $q_d$  are both quaternions. The multiplication between dual quaternion  $\hat{q}_1$  and  $\hat{q}_2$  is

$$\hat{q}_1 \circ \hat{q}_2 = q_{r1} \circ q_{r2} + \epsilon(q_{r1} \circ q_{d2} + q_{d1} \circ q_{r2}) \quad (5)$$

Note  $\hat{q}_1 \circ \hat{q}_2 \neq \hat{q}_2 \circ \hat{q}_1$  generically.

The *conjugation* of dual quaternion  $\hat{q}$  is defined as

$$\hat{q}^* = [\hat{s}, -\hat{\mathbf{v}}] \quad (6)$$

The *norm* of dual quaternion  $\hat{q}$  is defined as

$$\|\hat{q}\|^2 = \hat{q} \circ \hat{q}^* \quad (7)$$

The *multiplicative inverse* element of dual quaternion  $\hat{q}$  is

$$\hat{q}^{-1} = (1/\|\hat{q}\|^2) \circ \hat{q}^* \quad (8)$$

If  $\|\hat{q}\|^2 = 1 + \epsilon 0$ , then dual quaternion is called *unit dual quaternion*. For unit dual quaternion,  $\hat{q}^{-1} = \hat{q}^*$ .

Unit quaternions can be used to describe rotation. For the frame rotation about a unit axis  $\mathbf{n}$  with an angle  $|\theta|$ , there is a unit quaternion

$$q = [\cos(\frac{|\theta|}{2}), \sin(\frac{|\theta|}{2})\mathbf{n}] \quad (9)$$

relating a fixed vector expressed in the original frame  $r^o$  with the same vector expressed in the new frame  $r^n$  by

$$r^n = q^* \circ r^o \circ q \quad (10)$$

$r^o$  and  $r^n$  are two vector quaternions. Similarly, unit dual quaternion can be used to represent transformation (rotation and translation simultaneously). Suppose that there is a rotation  $q$  succeeded by a translation  $p^b$  (or a translation  $p^s$  succeeded by a rotation  $q$ ), the whole transformation can be represented using the dual quaternion as [10].

$$\hat{q} = q + \frac{\epsilon}{2} \cdot p^s \circ q \quad (11)$$

$$\hat{q} = q + \frac{\epsilon}{2} \cdot q \circ p^b \quad (12)$$

Unit quaternion  $q$  serves as a rotation taking coordinates of a point from one frame to another. On the other hand, every attitude of a rigid-body that is free to rotate relative to a fixed frame can be identified with a unique unit quaternion  $q$ . Analogous to the rotational case, an unit dual quaternion  $\hat{q}$  serves as both a specification of the configuration (consisting of attitude and position) of a rigid-body and a transformation taking the coordinates of a point from one frame to another.

Unit quaternions form a Lie-group over multiplication with the conjugate being the inverse, denoted by  $Q_u$ . The same goes for unit dual quaternions, which form a Lie-group over dual quaternion multiplication, denoted by  $DQ_u$ . The logarithm of unit quaternion, whose space denoted by  $v$ , is defined as

$$\ln q = \frac{|\theta|}{2} \mathbf{n} \quad (13)$$

where  $\mathbf{n}$  is the unit rotational axis and  $|\theta|$  is the rotational angle related to  $\mathbf{n}$ . As the same, the logarithmical mapping of unit dual quaternion is defined as follows.

$$\ln \hat{q} = \frac{1}{2}(\theta + \epsilon p^b) = \frac{1}{2}(\theta + \epsilon Ad_{q^*} p^s) \quad (14)$$

where  $\theta = |\theta| \mathbf{n}$  relates to rotation and  $p^b$ ,  $p^s$  represent translation. The space consisting of all unit dual quaternion logarithm is denoted by  $\hat{v}$ .  $\hat{v}$  is the Lie-algebra of  $DQ_u$  as well as  $v$  is the Lie-algebra of  $Q_u$ .

For unit quaternion, *Adjoint transformation* is defined as

$$Ad_q V = q \circ v \circ q^{-1} = q \circ v \circ q^* \quad (15)$$

where  $v$  is a vector quaternion and  $q$  is a unit quaternion. The same goes for the *Adjoint transformation* of unit dual quaternion,

$$Ad_{\hat{q}} \hat{v} = \hat{q} \circ \hat{v} \circ \hat{q}^{-1} = \hat{q} \circ \hat{v} \circ \hat{q}^* \quad (16)$$

where  $\hat{v}$  is a dual vector quaternion and  $\hat{q}$  is a unit dual quaternion.

Kinematic equations of the rigid-body expressed with unit dual quaternion are

$$\dot{\hat{q}} = \frac{1}{2} \xi^s \circ \hat{q} \quad (17)$$

$$\dot{\hat{q}} = \frac{1}{2} \hat{q} \circ \xi^b \quad (18)$$

where

$$\xi^s = \omega^s + \epsilon \cdot (\dot{p}^s + p^s \times \omega^s) \quad (19)$$

$$\xi^b = \omega^b + \epsilon \cdot (\dot{p}^b + \omega^b \times p^b) \quad (20)$$

$\xi^s$  and  $\xi^b$  are called twists, especially  $\xi^s$  is the twist in spatial-frame and  $\xi^b$  is the twist in body-frame.

According to (17) and (18), we obtain the relationships between  $\xi^s$  and  $\xi^b$ , i.e.

$$\xi^s = Ad_{\hat{q}} \xi^b \quad (21)$$

$$\xi^b = Ad_{\hat{q}^*} \xi^s \quad (22)$$

### 3 Dynamic Model Represented by Unit Dual Quaternion

No bi-invariant distance metric (i.e., one that is invariant under both left and right translations) exists in SE(3). Physically the left-invariance reflects the invariance of the metric with respect to choice of the body-frame and the right-invariance reflects the invariance with respect to the spatial-frame [12]. The relationship between  $\xi^s$  and  $\xi^b$  lies in (21) and (22), so we just pay attention to the spatial-frame in this study. Little work has been done on dynamics problems with dual quaternions. [13] gave the formulation of a general dynamics problem using dual quaternion components. But the equations of motion are quite complicated and the physical significance of the variables is not intuitively apparent. In this section, we will propose a new dynamic model with the dual quaternion and the traditional Newton Mechanics.

Differentiate the twist in spatial-frame in (19), then

$$\begin{aligned}\dot{\xi}^s &= \dot{\omega}^s + \epsilon(\ddot{p}^s + \dot{p}^s \times \omega^s + p^s \times \dot{\omega}^s) \\ &= \epsilon(\dot{p}^s \times \omega^s) + \dot{\omega}^s + \epsilon(\ddot{p}^s + p^s \times \dot{\omega}^s) \\ &= f + \hat{U}\end{aligned}\quad (23)$$

where

$$f = \epsilon(\dot{p}^s \times \omega^s) \quad (24)$$

$$\hat{U} = U_r + \epsilon U_d = \dot{\omega}^s + \epsilon(\ddot{p}^s + p^s \times \dot{\omega}^s) \quad (25)$$

As  $\dot{\omega}^s$  and  $\ddot{p}^s$  are caused by torques and forces respectively, the  $U_r$  responds to the acting torques on a moving rigid-body, while  $U_d$  corresponds to both forces and torques. Thus, we can take (23) as the rough dynamic model. And now we will build a bridge between this model and the traditional Newton Mechanics. The Newton Mechanics are modeled as follows.

$$\begin{cases} \ddot{p}_i = f_i/m \\ \ddot{\theta}_i = \tau_i/J \end{cases} \quad (26)$$

where  $f_i$  and  $\tau_i$  are the forces and torques respectively,  $p_i = (x, y, z)$  are the positions of rigid-body and  $\theta_i = (\phi, \theta, \psi)$  are the Euler angles to describe the rotation.

From [14], we have

$$\hat{\omega}^s = \dot{R} \cdot R^{-1} \quad (27)$$

thus

$$\dot{\omega}^s = \ddot{R} \cdot R^{-1} + \dot{R} \cdot \dot{R}^{-1} \quad (28)$$

$R$  can be represented by Euler angle  $(\phi, \theta, \psi)$  [14], thus  $\hat{\omega}^s$  can be represented by Euler angle according to (27). Substituting (26) to (27), (28) and then to (23), we can obtain the relationship between  $\hat{U}$  and Newton Mechanics. Thus,  $\hat{U}$  related to the control system dynamics. We can take  $\hat{U}$  as the dynamic input of the system.

Model (23) is much more compact than with homogeneous coordination in [14]. It simplifies the control law design in configuration space. We will take the Variable Structure Control (VSC) as an example in the following.

## 4 VSC Controller on $DQ_u$

It involves two aspects to design a VSC controller. The first is to design an appropriate switching function  $S(x)$  for a desired sliding mode dynamics, and the second is to design a control input for the reaching mode such that a reaching condition is met [15,16]. In this section, we firstly design the switching function and the reaching law respectively, then synthesize them to obtain the new VSC controller.

### 4.1 Switching Function

The desired sliding mode dynamics is usually a fast and stable response. Considering the relationship between Lie-group and Lie-algebra, we propose a switching function in  $\hat{v}$  space as follows.

$$S = \xi^s + 2\hat{k} \cdot Ad_{\hat{q}}(\ln \hat{q}) \quad (29)$$

where  $\hat{k} = (k_{r1}, k_{r2}, k_{r3}) + \epsilon(k_{d1}, k_{d2}, k_{d3})$  is a dual vector with each entry greater than zero. The responsiveness of the sliding mode lies on the parameter  $\hat{k}$ . Here we will prove its stability.

*Proof.* Let  $S = 0$ , then we get the equivalent control

$$\xi^s = -2\hat{k} \cdot Ad_{\hat{q}}(\ln \hat{q}) \quad (30)$$

Thus,

$$\begin{aligned} \xi^s &= -\hat{k} \cdot (q + \frac{\epsilon}{2} p^s \circ q) \circ (\theta + \epsilon Ad_{q^*} p^s) \circ (q^* - \frac{\epsilon}{2} q^* \circ p^s) \\ &= -\hat{k} \cdot (Ad_q \theta + \epsilon(p^s + p^s \times Ad_q \theta)) \\ &= -K_r Ad_q \theta - K_d(p^s + p^s \times Ad_q \theta) \end{aligned}$$

where  $K_r = diag(k_{r1}, k_{r2}, k_{r3})$  and  $K_d = diag(k_{d1}, k_{d2}, k_{d3})$ .

Compare with (19), then

$$\begin{cases} \omega^s = -K_r Ad_q \theta \\ \dot{p}^s = -K_d p^s + (K_r - K_d)p^s \times Ad_q \theta \end{cases} \quad (31)$$

It is easy to validate  $\omega^s = Ad_q \dot{\theta}$ , thus we have

$$\begin{cases} \dot{\theta} = -K_r \theta \\ \dot{p}^s = -K_d p^s + (K_r - K_d)p^s \times Ad_q \theta \end{cases} \quad (32)$$

Considering the Lyapunov function candidate

$$V = \alpha \theta \theta^T + \beta p^s (p^s)^T \quad (33)$$

where  $\alpha, \beta > 0$ . Obviously,  $V \geq 0$ .

Differentiate (33), then we have

$$\begin{aligned}\dot{V} &= 2\alpha\theta\dot{\theta}^T + 2\beta p^s(\dot{p}^s)^T \\ &= -2\alpha K_r\theta\theta^T - 2\beta p^s(K_d p^s - (K_r - K_d)p^s \times Ad_q\theta)^T \\ &= -2\alpha K_r\theta\theta^T - 2\beta K_d p^s(p^s)^T \\ &\leq 0\end{aligned}$$

Thus, this switching function is stable.  $\square$

## 4.2 Reaching Law

As  $DQ_u$  is not commutative, the traditional reaching laws of VSC, especially the proofs of the stability, are not tenable in this case. In this study, we propose a new reaching law in Theorem II on  $DQ_u$ .

**Theorem 1.** Suppose sliding surface is  $S \in \hat{v}$ . Then  $\dot{S} = -\hat{m} \cdot S$ , where  $\hat{m}$  is a dual vector with each entry greater than zero, is a reaching law.

Theorem II will be proven with Lyapunov's direct method. But before constructing the Lyapunov function candidate, we give the following definition first.

**Definition 2 (Dual Conjugation).** Let  $\hat{q} = q_r + \epsilon q_d$  be dual quaternion, then the dual conjugation of  $\hat{q}$  is

$$\hat{q}^* = q_r - \epsilon q_d \quad (34)$$

The conjugation in (6) comes from the definition of quaternion conjugation; whereas in Definition 2, the conjugation is defined from the definition of dual number conjugation (3) by treating dual quaternion as a dual number with quaternions being real part and dual part. Dual conjugation owns a series properties. One of them is listed as Lemma II, which is easy to validate directly.

**Lemma 1.** Let  $\hat{q}_1 = q_{r1} + \epsilon q_{d1}$  and  $\hat{q}_2 = q_{r2} + \epsilon q_{d2}$  be dual quaternions, then

$$\hat{q}_1 \circ \hat{q}_2^* + \hat{q}_1^* \circ \hat{q}_2 = 2q_{r1} \circ q_{r2} \quad (35)$$

According to Lemma II, we obtain

**Corollary 1.** Let  $\hat{v} = v_r + \epsilon v_d$  be dual vector quaternion, then

$$\hat{v} \circ \hat{v}^* + \hat{v}^* \circ \hat{v} = -2v_r \cdot v_r \leq 0 \quad (36)$$

Now, let's prove Theorem II

*Proof.* Considering the following Lyapunov function candidate:

$$V = -S \circ S^* - S^* \circ S \quad (37)$$

According to Corollary II,  $V \geq 0$ .

Differentiating (37), we obtain

$$\dot{V} = -\dot{S} \circ S^* - S \circ \dot{S}^* - \dot{S}^* \circ S - S^* \circ \dot{S}$$

As  $\dot{S} = -\hat{m} \cdot S$ , we obtain  $\dot{S}^* = -\hat{m} \cdot S^*$ , then

$$\begin{aligned}\dot{V} &= \hat{m} \cdot S \circ S^* - S \circ (-\hat{m} \cdot S^*) - (-\hat{m} \cdot S^*) \circ S - S^* \circ (-\hat{m} \cdot S) \\ &= 2\hat{m} \cdot (S \circ S^* + S^* \circ S) \\ &= -2\hat{m} \cdot V \\ &\leq 0\end{aligned}$$

So,  $\dot{S} = -\hat{m}S$  is a stable reaching law on  $DQ_u$ .  $\square$

### 4.3 Synthesis Switching Function and Reaching Law

Considering the sliding surface  $S = \xi^s + 2\hat{k} \cdot Ad_{\hat{q}}(\ln \hat{q})$ , we obtain

$$\dot{S} = \dot{\xi}^s + 2\hat{k} \cdot (Ad_{\hat{q}}(\ln \hat{q}))'$$

Substituting  $\dot{S} = -\hat{m} \cdot S$ , we get

$$\hat{U} = -\hat{m} \cdot (\xi^s + 2\hat{k} \cdot Ad_{\hat{q}}(\ln \hat{q})) - 2\hat{k} \cdot (Ad_{\hat{q}}(\ln \hat{q}))' - \epsilon(\dot{p}^s \times \omega^s) \quad (38)$$

Equation (38) is the generalized VSC-like controller in configuration space based on unit dual quaternion Lie-group. Obviously, this new proposed controller preserves the natural coupling between the translation and rotation with compact form.

## 5 Application on Omnidirectional Mobile Vehicles

Omnidirectional mobile vehicle can move along any direction, irrespective of the orientation of the vehicle. It is an attractive option in dynamic environments for its superior maneuvering capability compared to the more common car-like vehicles. The RoboCup competition, where teams of fully autonomous robots engage in soccer matches, is a typical example of where omnidirectional mobile vehicles have been used in dynamic environments. NuBot team, founded by National University of Defense Technology in 2004, is one of the RoboCup Middle Size League teams [17]. Inspiring by [18], the dynamic model of NuBot robots is as (39).

$$\begin{bmatrix} (M+m)\ddot{x} \\ (M+m)\ddot{y} \\ (J+2mL)\ddot{\theta} \end{bmatrix} + 2\beta \begin{bmatrix} \dot{x} \\ \dot{y} \\ 2L^2\dot{\theta} \end{bmatrix} + C \begin{bmatrix} \dot{x}\dot{\theta} \\ \dot{y}\dot{\theta} \\ \dot{\theta} \end{bmatrix} = \alpha D \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} \quad (39)$$

where  $x$  and  $y$  are the moving distance in x-axes and y-axes respectively,  $\theta$  is the rotation angle in spatial-frame,

$$C = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} -\sin(\theta) & -\cos(\theta) & \sin(\theta) & \cos(\theta) \\ \cos(\theta) & -\sin(\theta) & -\cos(\theta) & \sin(\theta) \\ L & L & L & L \end{bmatrix}$$

and  $M = 21kg$ ,  $m = 0.7kg$ ,  $L = 0.205m$ ,  $J = 0.4413kgm^2$ ,  $\alpha = 7.009$ ,  $\beta = 1.8138$ ,  $-24V < U_i (i = 1, 2, 3, 4) < 24V$ .

In this section, we take model (39) as the example to validate our new control law (38). The configuration space of a planar robot like NuBot robot contains two translational and one rotational dimensions, which is homeomorphic to  $SE(2)$ , thus the control law (38) is specified into this space firstly.

The following relations exist in the configuration space of planar robot:

$$\begin{cases} \omega^s = (0, 0, \dot{\theta}) \\ \xi^s = \omega^s + \epsilon(p^s + p^s \times \omega^s) = (0, 0, \dot{\theta}) + \epsilon(\dot{x} + y\dot{\theta}, \dot{y} - x\dot{\theta}, 0) \end{cases} \quad (40)$$

Thus, we get

$$\ln \hat{q} = \frac{1}{2}(\theta + \epsilon Ad_{q^*} p^s) = \frac{1}{2}(0, 0, \theta) + \frac{\epsilon}{2}(x \cos \theta + y \sin \theta, y \cos \theta - x \sin \theta, 0) \quad (41)$$

Thus,

$$Ad_{\hat{q}}(\ln \hat{q}) = Ad_q \theta + \epsilon(p^s + p^s \times Ad_q \theta) = \frac{1}{2}(0, 0, \theta) + \epsilon(x + \theta y, y - \theta x, 0) \quad (42)$$

Thus,

$$(Ad_{\hat{q}}(\ln \hat{q}))' = \frac{1}{2}((0, 0, \dot{\theta}) + \epsilon(\dot{x} + y\dot{\theta} + \dot{y}\theta, \dot{y} - x\dot{\theta} - \dot{x}\theta, 0)) \quad (43)$$

According to (40),

$$\dot{p}^s \times \omega^s = (\dot{y}\dot{\theta}, -\dot{x}\dot{\theta}, 0) \quad (44)$$

and

$$\hat{U} = \dot{\omega}^s + \epsilon(\ddot{p}^s + p^s \times \dot{\omega}^s) = (0, 0, \ddot{\theta}) + \epsilon(\ddot{x} + y\ddot{\theta}, \ddot{y} - x\ddot{\theta}, 0) \quad (45)$$

substitute (42), (43), (44), (45) to (38), then

$$\begin{cases} \ddot{\theta} = -(m_1 + k_1)\dot{\theta} - m_1 k_1 \theta \\ \ddot{x} = -(m_2 + k_2)(\dot{x} + y\dot{\theta}) - m_2 k_2(x + \theta y) - k_2 \dot{y}\theta - \dot{y}\dot{\theta} - y\ddot{\theta} \\ \ddot{y} = -(m_3 + k_3)(\dot{y} - x\dot{\theta}) - m_3 k_3(y - \theta x) - k_3 \dot{x}\theta + \dot{x}\dot{\theta} + x\ddot{\theta} \end{cases} \quad (46)$$

The new proposed controller (46) can be applied directly without decoupling model (39). First, we introduce three temporary variables  $F_x$ ,  $F_y$  and  $T$ , and then divide model (39) into two parts as (47) and (48).

$$\begin{bmatrix} (M+m)\ddot{x} \\ (M+m)\ddot{y} \\ (J+2mL)\ddot{\theta} \end{bmatrix} + 2\beta \begin{bmatrix} \dot{x} \\ \dot{y} \\ 2L^2\dot{\theta} \end{bmatrix} + C \begin{bmatrix} \dot{x}\dot{\theta} \\ \dot{y}\dot{\theta} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \\ T \end{bmatrix} \quad (47)$$

$$\begin{bmatrix} F_x \\ F_y \\ T \end{bmatrix} = \alpha D \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} \quad (48)$$

Then, we substitute (46) to (47),

$$\begin{cases} T = -(J + 2mL)(k_1 + m_1) + 4\beta L^2 \dot{\theta} - (J + 2mL)m_1 k_1 \theta \\ F_x = (-M_a(m_2 + k_2) + 2\beta)\dot{x} - M_a m_2 k_2 x - (M_a - 1)\dot{y}\dot{\theta} \\ \quad - M_a k_2 y\dot{\theta} - M_a(m_2 + k_2)y\dot{\theta} - M_a m_2 k_2 y\theta - M_a y\ddot{\theta} \\ F_y = (-M_a(m_3 + k_3) - 2\beta)\dot{y} - M_a m_3 k_3 y + (M_a - 1)\dot{x}\dot{\theta} \\ \quad - M_a k_3 x\dot{\theta} - M_a(m_3 + k_3)x\dot{\theta} + M_a m_3 k_3 x\theta + M_a x\ddot{\theta} \end{cases} \quad (49)$$

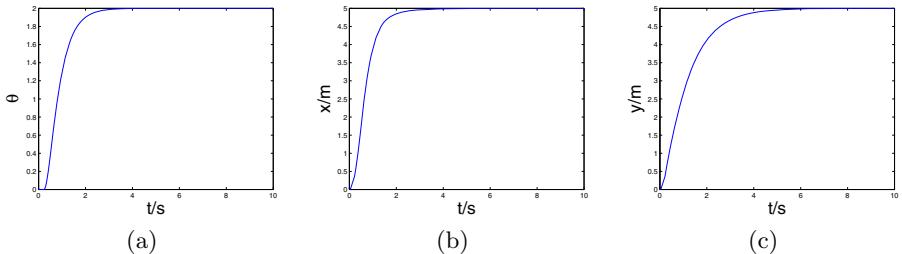
where  $M_a = M + m$ .

And then, we apply the pseu-inverse method to control (48) as follows.

$$\begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} = \frac{1}{\alpha} D^\dagger \begin{bmatrix} F_x \\ F_y \\ T \end{bmatrix} = \frac{1}{\alpha} \begin{bmatrix} -\sin \theta/2 & \cos \theta/2 & 1/4L \\ -\cos \theta/2 & -\sin \theta/2 & 1/4L \\ \sin \theta/2 & -\cos \theta/2 & 1/4L \\ \cos \theta/2 & \sin \theta/2 & 1/4L \end{bmatrix} \begin{bmatrix} F_x \\ F_y \\ T \end{bmatrix} \quad (50)$$

At last, combining (49) and (50), we obtain the full controller. In this study, we validate the new controller in simulink of Matlab. We assume that  $k_1 = 5, k_2 = 1, k_3 = 1; m_1 = 2, m_2 = 10, m_3 = 10$  and the original configuration is  $x_0 = 0, y_0 = 0, \theta_0 = 0$ , then simulate the control law with different target configurations. All the target configurations are reachable. Fig.1 shows the simulation results with a typical target configuration  $x_t = 5m, y_t = 5m, \theta_t = 2$ .

These simulations show that the controller (49) and (50) acting model (39) are stable. The near-optimal dynamic trajectory generation and control is proposed based on the Bang-bang control in [18]. But the third item of the left-side in (39) was omitted and the maximal common range of matrix D was used to substitute



**Fig. 1.** Simulation results in  $\theta$ , x-axes, y-axes with experiential parameters ( $k_1 = 5, k_2 = 1, k_3 = 1; m_1 = 2, m_2 = 10, m_3 = 10$ ). The original configuration is ( $\theta = 0, x = 0m, y = 0m$ ) and the target configuration is ( $\theta_t = 2, x_t = 5m, y_t = 5m$ ).

this matrix to decouple rotation and translation in model (39). So, the VSC controller based on dual quaternions proposed in this paper works well and can deal with the interconnection between rotation and translation.

## 6 Conclusion and Future Work

This paper proposed a new dynamic model with unit dual quaternion, which makes designing complex controller in configuration space much easier. Then a new VSC law on  $DQu$  is devised and validated in simulation with the NuBot omnidirectional robot. Compared with the conventional control law, this new control law preserves the natural coupling between translation and rotation.

This paper is an initial attempt to use the dual quaternion in designing complex dynamic control laws. There are several topics for future work. A compact dynamic model has been proposed, thus it is possible to apply other modern control laws, such as self-adaptive control law, fuzzy control law, especially the optimal control for better performance. In addition, tuning of parameters  $\hat{k}$  and  $\hat{m}$  is a hard work, thus it is a notable work to study the self-tuning problem of this controller. Genetic Algorithms(GAs) may be adoptable. Finally, unit dual quaternion owns a distinct advantage to deal with multi-robot control because of its compactness and computation efficiency, thus it is meaningful to represent the multi-robot motion with dual quaternion in  $SE(3) \times SE(3) \times \dots \times SE(3)$ .

## Acknowledgment

The authors would like to acknowledge the NuBot team members for their cooperation to establish and develop the RoboCup Middle Size League soccer robot team–NuBot.

## References

1. Pettersen, K.Y., Egeland, O.: Time-varying exponential stabilization of the position and attitude of an underactuated autonomous underwater vehicle. *IEEE Transactions on Automatic Control* 44(1), 112–115 (1999)
2. Servidua, P.A., Pena, R.S.S.: Thruster design for position/attitude control of spacecraft. *IEEE Transactions on Aerospace and Electronic Systems* 38(4), 1172–1180 (2002)
3. Featherstone, R.: Plücker basis vectors. In: Proceedings of the 2006 IEEE International Conference on Robotics & Automation, Orlando, Florida, USA, 1892–1897 (2006)
4. Bullo, F., Murray, R.: Proportional derivative (PD) control on the euclidean group. In: 1995 European Control Conference, Rome, Italy (1995)
5. Funda, J., Taylor, R.H., Paul, R.P.: On homogeneous transformations, quaternions, and computational efficiency. *IEEE Transactions on Robotics and Automation* 6(3), 382–388 (1990)

6. Aspragathos, N.A., Dimitros, J.K.: A comparative study of three methods for robot kinematics. *IEEE Transactions on Systems, Man and Cybernetics Part B: Cybernetics* 28(2), 135–145 (1998)
7. Daniilidis, K.: Hand-eye calibration using dual quaternions. *The International Journal of Robotics Research* 18(3), 286–298 (1999)
8. Goddard, J.S.: Pose and Motion Estimation from Vision Using Dual Quaternion-based Extended Kalman Filtering. Ph.D thesis, The University of Tennessee (1997)
9. Perez, A., McCarthy, J.: Dual quaternion synthesis of constrained robotic systems. *Journal of Mechanical Design* 126(3), 425–435 (2004)
10. Wu, Y., Hu, X., Hu, D., Lian, J.: Strapdown inertial navigation system algorithms based on dual quaternions. *IEEE Transactions on Aerospace and Electronic Systems* 41(1), 110–132 (2005)
11. Han, D., Wei, Q., Li, Z.: A dual-quaternion method for control of spatial rigid body. In: Proceedings of the 2008 IEEE International Conference on Networking, Sensing and Control, Sanya, China, pp. 1–6 (2008)
12. Park, F.: Distance metrics on the rigidbody motions with applications to mechanism design. *Transactions of the ASME* 117(3), 48–54 (1995)
13. Dooley, J., McCarthy, J.: Spatial rigid body dynamics using dual quaternion components. In: Proceedings of the 1991 International Conference on Robotics and Automation, Sacramento California, pp. 90–95 (1991)
14. Murray, R.M., Li, Z., Sastry, S.S.: *A Mathematical Introduction to Robotic Manipulation*. CRC Press, Boca Raton (1994)
15. Gao, W., Hung, J.C.: Variable structure control of nonlinear system: A new approach. *IEEE Transactions on Industrial Electronics* 40(1), 45–55 (1993)
16. Hung, J.Y., Gao, W., Hung, J.C.: Variable structure control:a survey. *IEEE Transactions on Industrial Electronics* 40(1), 2–22 (1993)
17. Zhang, H., Wang, X., Lu, H., Ji, X.: Nubot team description paper 2009. In: RoboCup 2009 Graz, Graz, Austria (2009)
18. Kalmár-Nagy, T., D'Andrea, R., Ganguly, P.: Near-optimal dynamic trajectory generation and control. *Robotics and Autonomous Systems* 46(1), 47–64 (2004)

# A 3-D Simulation of Unicycle Robot Based on Virtual Prototyping Technology

Xiaogang Ruan, Qiyuan Wang, and Naigong Yu

Beijing University of Technology,  
School of Electronic Information and Control Engineering, 100124 Beijing, China  
wangqiyuan@emails.bjut.edu.cn

**Abstract.** Co-simulation of ADAMS and MATLAB for designing and developing mechatronics and control system is presented in this paper. The co-simulation technology provides an avenue for realizing mechatronics system design loop and control system design synchronously without the need for building prototypes. The virtual prototype developed in ADAMS can be used to investigate dynamic behavior in a 3-dimensional environment. Though ADAMS has the capability of implementing a closed loop control of the virtual prototype, its capability is quite restricted. On the other hand, MATLAB is well known for designing control systems. Co-simulation allows us to get the benefit of both. The co-simulation platform was used to verify different features of a unicycle robot. Mechanical drawings of the robot are first created using CAD software, e.g., SOLIDWORKS and imported into the ADAMS. Dynamics of the robot in autonomous mode are simulated using the co-simulation platform in which the controller is designed and implemented using MATLAB. Simulation results can then be used to modify mechanical design and improve the control method. Although the virtual platform is tested for one specific system, it can be easily used for design of other mechatronics systems and control methods e.g., model free control, learning control and developmental robotics.

**Keywords:** Unicycle Robot; Virtual Prototyping Technology; Co-simulation; PID Controller.

## 1 Introduction

Simulation is the base of most design involving complex operations. Traditionally, simulations are computer based and require mathematical models to describe system dynamics and to find analytical solutions[1-2].

The use of virtual prototyping technology provides a different avenue for system design and simulation from traditional method. Automatic Dynamic Analysis of Mechanical Systems (ADAMS) started as a general purpose program that can analyze systems undergoing large non-linear displacements while under the effect of non-linear force and motion input. The methodology developed by Nicolae Orlandea was the basis of his PhD dissertation at the University of Michigan [3]. It can be categorized as a general purpose numeric code utilizing a non-minimal set of co-ordinates to

develop the equations of motion. It uses stiff integrators to solve these equations and sparse matrix algebra to solve the linear algebraic equations in its innermost computation loop[4]. Since its inception, significant development investments resulted in sophisticated virtual prototyping tools for a wide range of industrial applications[5-8].

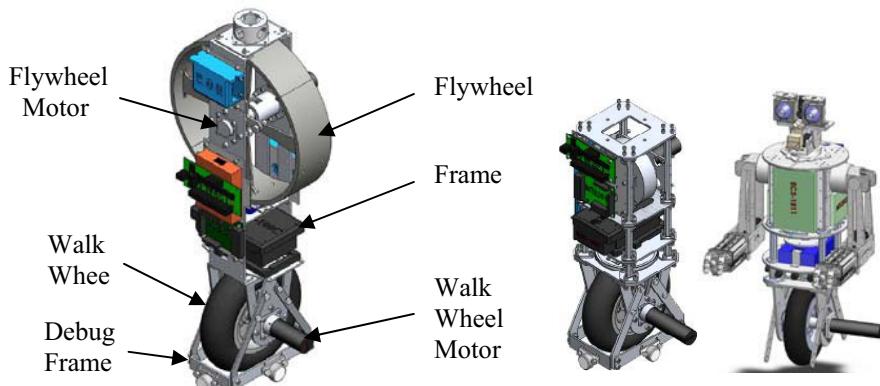
Unicycle riding is a kind of senior motor skills of human beings or other intelligent animals after a period of study. Construction of an autonomous unicycle riding robot system, imitation of human behavior of riding unicycles to achieve autonomous, dynamic, adaptive learning and even 3-D self-balance and motion is the research objective of unicycle robot. The unicycle robot system is an inherently unstable, coupled and highly nonlinear plant in 3-D. Although several unicycle robots have been invited[9-17], and these unicycle robots have their own details, most of them have similar body structure prototype and drive mechanism –the single walk wheel structure. And there is no research on the unicycle robot through ADAMS and MATLAB co-simulation method.

This paper presents the design, definition and simulation of unicycle robot in a virtual environment that integrates ADAMS with MATLAB. The actual system and its drawings are briefly described in section 2. Open loop and closed loop simulation of the virtual robot are introduced are described in section 3. Simulation results in Section 3 and 4 show that a good representation of the actual unicycle robot system and its control system is created in ADAMS and MATLAB co-simulation environment.

## 2 Unicycle Robot System

The mechanical and control structure of unicycle robot is designed as shown in Figure 1. The overall structure has three modules: the roll balance module on the top to tune the roll DOF and the pitch balance module with the walk wheel to adjust the pitch DOF, with robot frame to fix all modules. An adjustable and removable bracket with 4 castors is designed for debugging. The flywheel is fixed vertical through the higher axle of the frame. Its rotation provides the reaction force for the adjustment of roll DOF. It is spun by the flywheel DC motor in the back of the flywheel shaft. The walk wheel DC motor attached in the back of the walk wheel shaft tunes the walk wheel for the adjustment of pitch DOF. All motors are attached to the shaft by precision planetary gearheads.

The electrical system for the robot includes a. sensory system, b. central control system, c. actuator and power system. Sensory system measures pitch tilt, roll tilt, tilt velocity, position and other variables of interest for autonomous operation of unicycle robot, including: a. gyroscope, b. inclinometer, c. infrared sensor, d. sonar, e. encoder; Central control system calculates the feedback signal through signal processing algorithms according to sensory system output and figures out the control signal of motors through the balance control algorithm and then sends such command to the actuators, including: a. DSP as motion controller, b. MCU as assistant controller; Actuators take the movement by driving motors according to signals received from central control system, including all motor drivers. Power system supplies all the required energy for electrical modules, including power supply board and rechargeable Lithium batteries module.



**Fig. 1.** Schematic of Unicycle Robot (Type 1/ Type 2/ Type 3)

### 3 Dynamic Simulation in ADAMS

After created using SOLIDWORKS according to the shape and dimensions of the actual parts used in unicycle robot, the 3-D model can be imported into the ADAMS environment for settings of mechanical parameters.

#### 3.1 Unicycle Robot in ADAMS

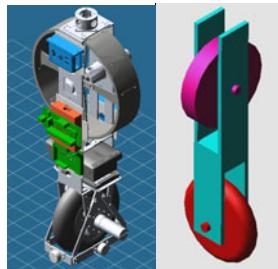
Three dimensional drawings of different parts of the unicycle robot are imported into ADAMS as parasolid file. Then material, mass, density and friction are defined. Mass and inertial matrices are automatically calculated. Main parameters of different parts used to define the model of the unicycle robot are listed in Table 1. Unicycle robot model in ADAMS takes several aspects into account, such as gravity, contact constraints, friction, inertial properties and reference markers. All these definitions and settings must be defined properly for a good approximation of real robot behavior in the virtual environment. Table 2 shows some of the main joints and motions that have been defined for unicycle robot model. Each joint is defined to have a particular motion with specific purpose. For example, the driving flywheel joint is specified as revolute joint and so the definition of the rotation along with it is the flywheel spinning.

**Table 1.** Main Inertial Data for Unicycle Robot Model

Parts	Mass (kg)	$I_{xx}$ (kg m <sup>2</sup> )	$I_{yy}$ (kg m <sup>2</sup> )	$I_{zz}$ (kg m <sup>2</sup> )
Walk Wheel	1.96	$7.78 \times 10^{-2}$	$4.04 \times 10^{-2}$	$4.06 \times 10^{-2}$
Flywheel	4.90	$4.37 \times 10^{-2}$	$4.37 \times 10^{-2}$	$7.49 \times 10^{-3}$
Frame	3.37	$1.86 \times 10^{-3}$	$1.86 \times 10^{-3}$	$3.98 \times 10^{-4}$
Total	10.24	NA	NA	NA

**Table 2.** Joints and Motions Defined for Unicycle Robot Model

Joint	Motion	Action
Frame_WalkWheel_joint	SForce(Revolute joint)	Drive the walk wheel
Frame_Flywheel_joint	SForce(Revolute joint)	Drive the flywheel
Ground_WalkWheel_contact	Frictional force	Contact

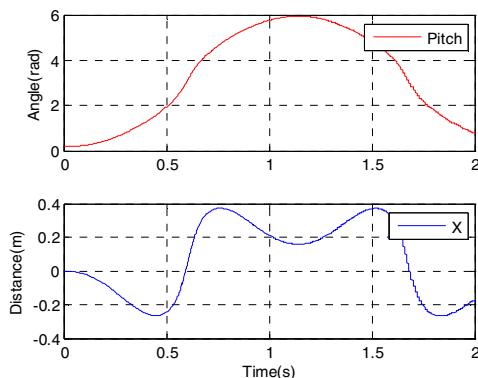
**Fig. 2.** Visual prototype in ADAMS (right: simplified version)

### 3.2 Open Loop Simulation

The unicycle robot is inherently unstable, coupled and highly nonlinear plant in 3-D. That is the robot initially standing upright will fall down if there is some disturbance. If there is a pitch or roll non-zero tilt, the robot will fall down too with some displacement. And because of the coupling effect, both some pitch and roll angle will engender some yaw tilt.

#### Simulation Test 1

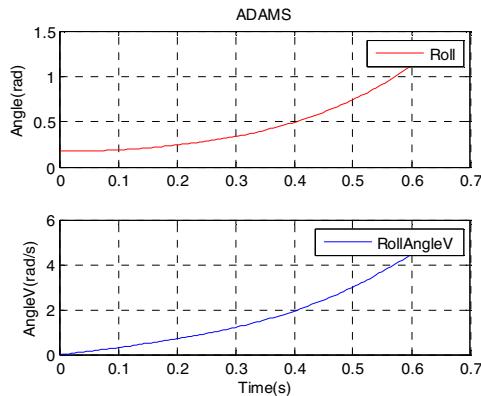
Set the initial point of simulation is: the roll tilt: 0rad; the pitch tilt: 0.1rad; the yaw tilt: 0rad, the initial position at (0 m, 0 m). The simulation results are shown in Figure 3.

**Fig. 3.** Simulation result

It is observed that in the uncontrolled case, when the robot deviates from the upright state, it will continue to fall along this direction until contact with the ground, the same with the actual situation. The ground condition of simulation is not set here. Therefore, the robot will continue to swing downward, like a pendulum.

### Simulation Test 2

Set the initial point of simulation is: the roll tilt: 0.175rad; the pitch tilt: 0rad; the yaw tilt: 0rad, the initial position at (0 m, 0 m). The simulation results are shown in Figure 4.



**Fig. 4.** Simulation result

It is observed that in the uncontrolled case, when the robot deviates from the upright state, it will continue to fall along this direction.

### Simulation Test 3

Robot initial pitch angle: 0.1rad; initial roll angle: 0.1rad; initial yaw angle: 0rad; Initial position: 0. The results shown in Figure 5.

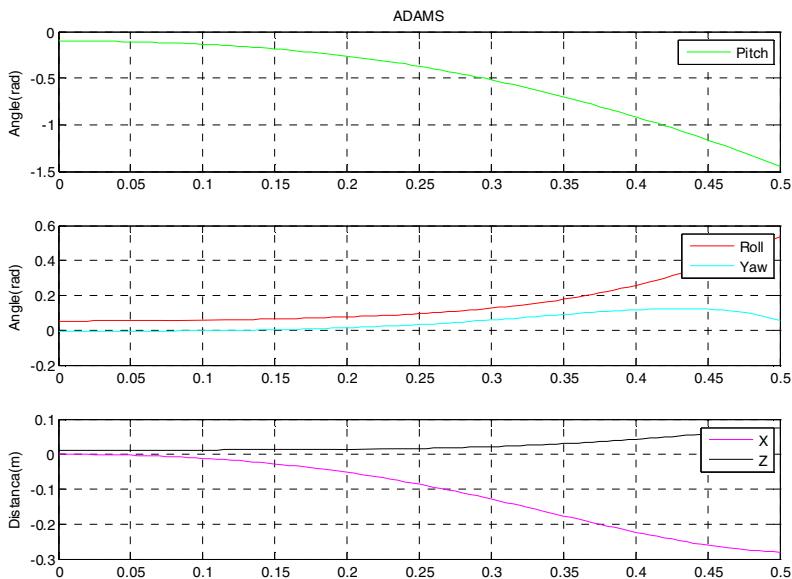
Robot falls along both directions of pitch and roll with some longitudinal displacement, lateral displacement and yaw angle.

The above results of three simulations coincide with the actual experience. Therefore, the ADAMS model is verified to represent the real robot.

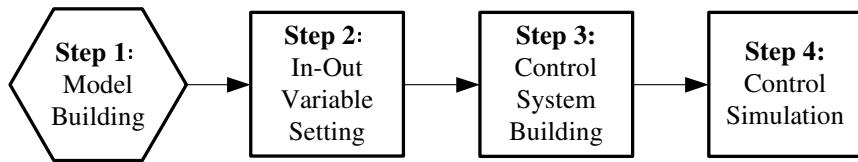
## 4 Closed Loop Simulation

### 4.1 Import the Unicycle Robot Model

Multibody system need to be established in the ADAMS for ADAMS and MATLAB co-simulation. Then the export model from the ADAMS should be imported into MATLAB environment. And MATLAB reads the output information of the plant and calculates control output while the ADAMS solver solves system equations for the plant dynamic. Such progress has 4 main steps.

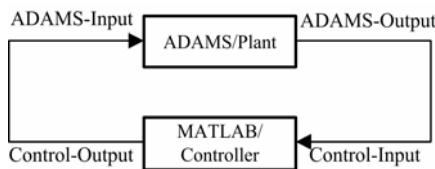


**Fig. 5.** Simulation result



**Fig. 6.** Flowchart of ADAMS/MATLAB co-simulation

After defining the input and output variables of ADAMS, it can form a closed information loop between MATLAB and ADAMS, as shown in Figure 7.



**Fig. 7.** Information flowchart of ADAMS/MATLAB co-simulation

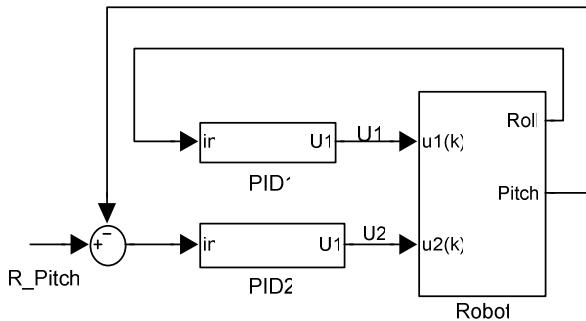
The input variables for the plant are walk wheel motor torque and flywheel motor torque. Output variables for the robot are roll, pitch, yaw angle, location and the

derivatives of them. After exporting control object to the MATLAB environment, exchange parameters need to be set between MATLAB and ADAMS.

## 4.2 Controller for the Unicycle Robot

Such factors must be taken into account: the actual self-balancing unicycle robot control system is discrete, memory and computing speed of its computing unit is limited and realizability of actual controller. Finally, the structure of dual-loop PID decoupling control is applied. In the control structure, pitch and roll angles are tuned respectively by a PID controller. The control diagram is shown in figure 8. Gains of the controllers are calculated respectively as:

$$u = k_P (\theta_{ref} - \theta) + k_I \int (\theta_{ref} - \theta) + k_D \dot{\theta} \quad (1)$$



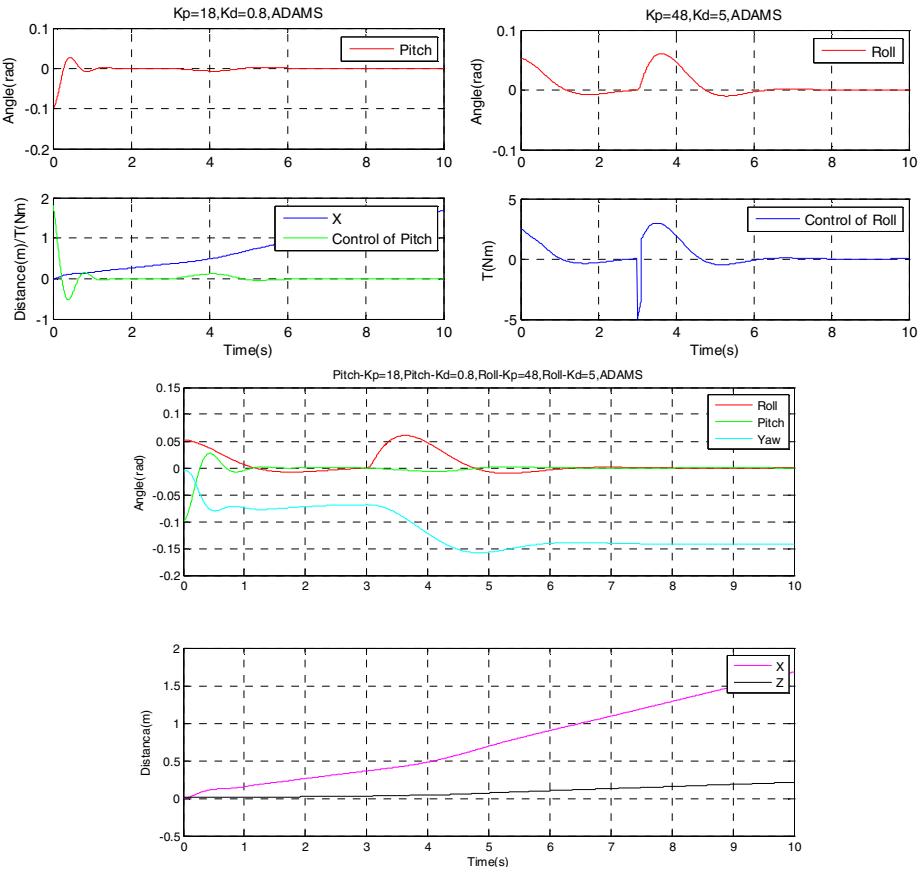
**Fig. 8.** Structure of controller

## 4.3 Control Experiment

Parameters of controller are obtained according to the critical oscillation method as Pitch:  $K_p=18$ ,  $K_d=9$ ; Roll:  $K_p=28$ ,  $K_d=1.6$ . And initial state of the robot is pitch angle: 0.1rad, roll angle: 0.0525rad, position 0m. At 3s a 1Nm torque for 0.1s is applied to the unicycle robot. Curves are as follows:

It can be seen that the final pitch and roll angle reach equilibrium. The unicycle robot remains stable after the impulse at 3s. Some yaw tilt emerges. That is a very important revelation to control the direction of the unicycle robot. The results demonstrate that the decoupling controller applied to the unicycle robot is feasible.

Simulation results shown here show reasonable response for controller realized in MATLAB. Capability of ADAMS in control design is quite restricted as it is meant for 3D simulation of mechanical systems but not for design and analysis of controller. Integrating ADAMS with MATLAB let us exploit the better aspects of both.



**Fig. 9.** Control result of co-simulation

## 6 Conclusion

Dynamic simulation of self-balancing unicycle riding robot using ADAMS has been presented here. Several tests have been carried out to verify principle of operation of the unicycle robot. Test results show that the virtual robot's behavior is in accordance with the basic principles and, therefore, the virtual unicycle robot created in ADAMS represents the actual mechanical system. Using the internal control template provided by ADAMS to design and implement closed loop operation limits the flexibility of this virtual system as an effective tool. So we integrated ADAMS with MATLAB /SIMULINK to exploit the strength of MATLAB in designing controller. The virtual prototype and co-simulation technology can be used for various investigations which are otherwise time-consuming or costly.

Future work available from the following four areas for further study:

1. Manufacture the designed unicycle robot, and then do the same experiment on it to compare test results.
2. Apply self-learning control of balance in co-simulation environment. Make the robot stable from unstable after learning in the simulation environment. Apply the same controller to the actual robot, compare the results.

**Acknowledgments.** This work is supported by 863 program of China (2007AA04Z226), NSFC (60774077), BMEC (KZ200810005002) and BMSF.

## References

1. Zhu, Z., Naing, M.P., Al-Mamun, A.: A 3-D Simulator using ADAMS for Design of an Autonomous Gyroscopically Stabilized Single Wheel Robot. In: Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, USA (2009)
2. Zhu, Z., Naing, M.P., Al-Mamun, A.: Integrated ADAMS+MATLAB environment for design of an autonomous single wheel robot. In: 35th Annual Conference of the IEEE Industrial Electronics Society, Porto, Portugal, pp. 2253–2258 (2009)
3. Orlandea, N.: Development and application of node-analogous sparsity-oriented methods for simulation of mechanical dynamic systems. PhD thesis, University of Michigan, Ann Arbor MI (1973)
4. Rampalli, R.: ADAMS – a Sparse Matrix Approach to Solving Multi-body Dynamics Problems. In: NASA Workshop on Multi-Body Simulation, California Institute of Technology (1987)
5. MathWorks Inc. Lockheed Martin Space Systems Uses SimMechanics with a Real-Time Simulator to Automate Mars Reconnaissance Orbiter Development (2010-04-2), <http://www.mathworks.com>
6. Kesner, S.B.: Mobility Feasibility Study of Fuel Cell Powered Hopping Robots for Space Exploration. Massachusetts Institute of Technology, Massachusetts (2007)
7. Lindemann, R.: Dynamic Testing and Simulation of the MARS Exploration Rover. In: Proceedings of IDETC/CIE 2005 2005 ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Long Beach, California, USA (2005)
8. Xiaohua, Z.: System Modelling and Simulation. Tsinghua University Press, Beijing (2006)
9. Tsai, C.-C., Chan, C.-K., Shih, S.-C., Lin, S.-C.: Adaptive Nonlinear Control Using RBFNN for an Electric Unicycle. In: 2008 IEEE International Conference on Systems, Man and Cybernetics, SMC 2008 (2008)
10. University of California San Diego, <http://sicl.ucsd.edu/jaschavp/Blog/Blog.html>
11. Murata Manufacturing Co., Ltd., [http://www.murata.co.jp/corporate/boy\\_girl/girl/index.html](http://www.murata.co.jp/corporate/boy_girl/girl/index.html)
12. Nakajima, R., Tsubouchi, T., Yuta, S., Koyanagi, E.: A Development of a New Mechanism of an Autonomous Unicycle. In: IEEE Proc. IROS 1997 (1997)
13. Sheng, Z., Yamafuji, K.: Postural stability of a human riding a unicycle and its emulation by a Robot. IEEE Transactions Robotics and Automation 13(5), 709–720 (1997)

14. Xu, Y., Au, K.W.: Stabilization and path following of a single wheel robot. *IEEE/ASME Transactions on Mechatronics* 9(2), 407–419 (2004)
15. Schoonwinkel, A.: Design and test of a computer stabilized unicycle. Ph.D. dissertation, Stanford Univ., CA (1987)
16. Majima, S., Kasai, T.: A controller for changing the yaw direction of an underactuated unicycle robot. In: International Conference on Technology and Automation, vol. 5, pp. 73–88 (2005)
17. Sheng, Z., Yamafuji, K.: Postural Stability of a Human Riding a Unicycle and Its Emulation by a Robot. *IEEE Transaction on Robotics and Automation* 13(5) (1997)

# A Study of the Electric Wheelchair Hands-Free Safety Control System Using the Surface-Electromyogram of Facial Muscles

Hiroki Tamura, Takao Manabe, Takafumi Goto, Yuki Yamashita, and Koichi Tanno

Faculty of Engineering, University of Miyazaki  
Miyazaki, 889-2192 Japan  
htamura@cc.miyazaki-u.ac.jp

**Abstract.** The goal of Human-Computer Interface (or called Human-Robot interface) research is to provide humans with a new communication channel that allows translating people's intention states via a computer into performing specific actions. This paper presents a novel hands-free control system for controlling the electric wheelchair, which is based on Bio-signals as surface electromyogram signals. The Bioelectric signals are picked up from facial muscles then the Bio-signals are passed through an amplifier and a high pass filter. Motion control commands (Forward, Left, Right, Forward to the Right, Forward to the left and Stop) are classified by simple rule. These commands are used for controlling the electric wheelchair.

**Keywords:** Surface Electromyogram Signal, Electric Wheelchair, Motion Control Command Rule, The Obstacle Evading Function.

## 1 Introduction

Surface electromyogram signals (abbr. s-EMG) are detected over the skin surface and are generated by the electrical activity of the muscle fibers during contraction [1]. Moved muscle can be presumed by analyzing s-EMG. S-EMG is used to control artificial leg etc. s-EMG recognition of using the conventional neural network is a method which learns the relation between s-EMG patterns and is reproduced using a neural network. Our previous works [4] [5] [6] showed that technique of integrated EMG is much easier and superior than the fast fourie transform (FFT) analysis. Integrated EMG is a running average of the rectification s-EMG.

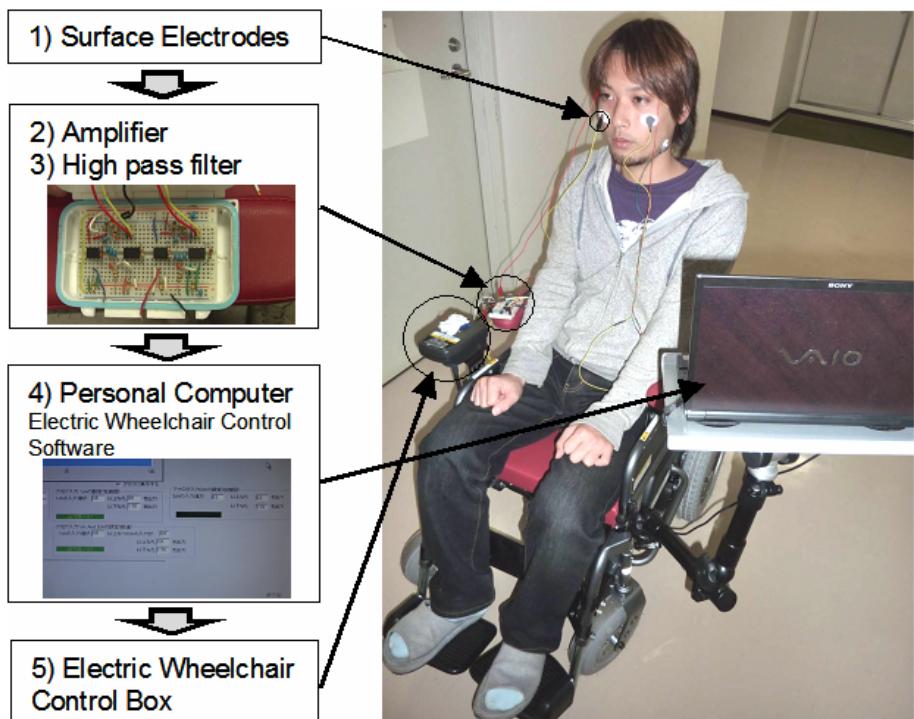
One of the major challenges for prosthesis or Human-Robot Device development is to produce devices to perfectly mimic their natural counterparts. A very popular approach for prosthesis or Robot control is based on the use of Bio-signals (s-EMG, electroencephalogram (abbr. EEG), electrooculogram (EOG) ,etc ). One of them is s-EMG collected from remnant or normal muscles and use them as control inputs for the artificial limb or Robot Device. As these devices, known as s-EMG-based Hands-Free Wheelchair [2] use a biological signal to control their movements, it is expected that they should be much easier to control [3]. Other techniques have the method of

using gesture [8]. However, the system of gesture is not used by the severely handicapped human.

In this paper, we applied to the electric wheelchair hands-free safety control system for the severely handicapped human using facial muscles s-EMG pattern recognition system. The six motion control commands are classified by our proposal simple rule. We tried the experiment of driving in the course by using our proposal system. The experimental subjects were three healthy men in twenties whose physique looks like. One person was an expert of the system, and others were the inexperienced person. From their experiments, we showed that the control of our proposal system is easy. In addition, the obstacle evading function was added to our system for the safety control.

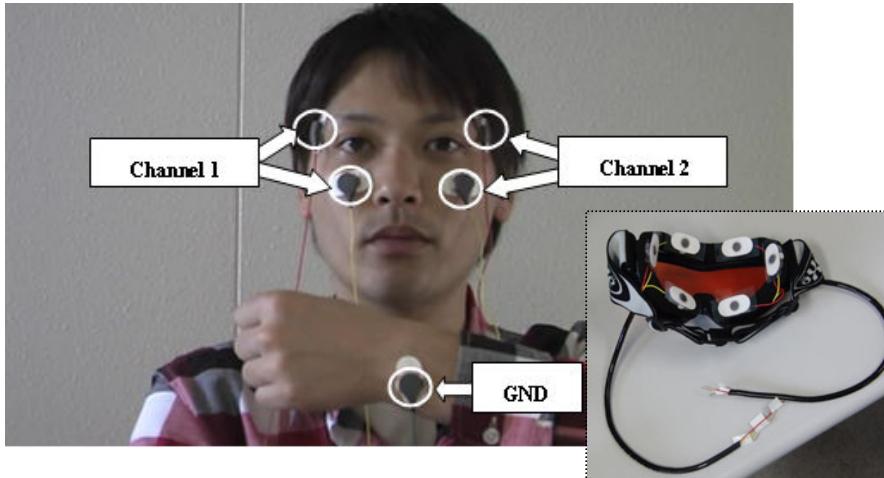
## 2 Electric Wheelchair Hands-Free Control System Using S-EMG Signals

Head movement is a natural form of gesture and can be used to indicate a certain direction [2]. Serious disabled people can not move neck and head, but can get face figures. s-EMG is a way of studying facial muscles activities by recording action potentials from contracting fibres. S-EMG can be detected with surface electrodes. Surface electrodes are easy to apply. This is a non-invasive way to record s-EMG while posing no health and safety risk to the users.



**Fig. 1.** The Flow of Electric Wheelchair Control System using surface s-EMG

Figure 1 shows the formal scheme for the acquisition and analysis of the s-EMG signal for the control organization and flow of information through the system. Our proposal system is based on the five features, 1) surface electrodes, 2) Amplifier, 3) High pass filter, 4) Personal Computer for s-EMG signal classification and motion control command output, and 5) Electric Wheelchair Control.



**Fig. 2.** The Position of Surface Electrodes and The Goggle for Electrodes Fixation

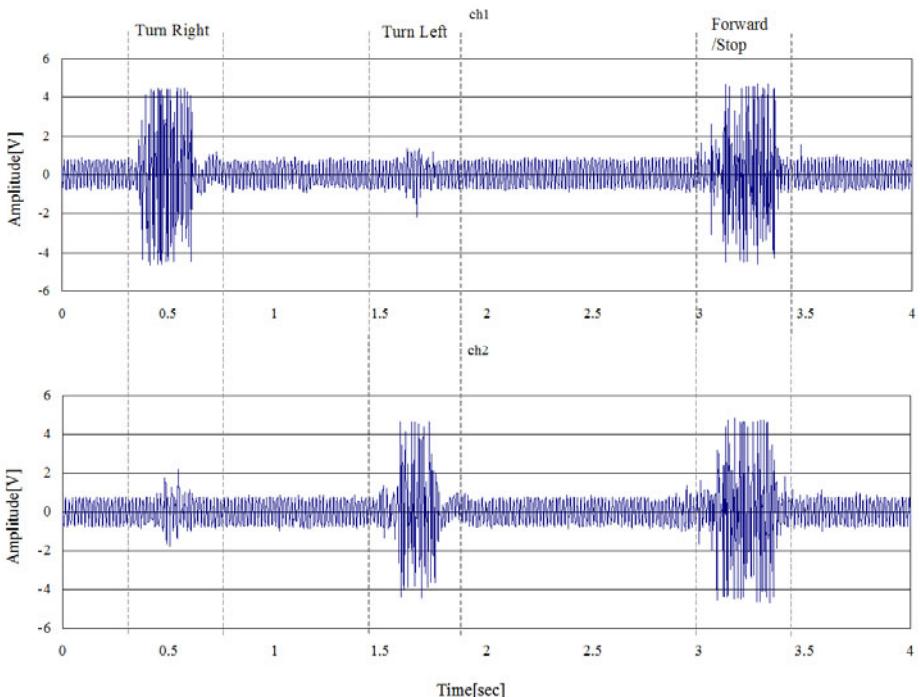
**Table 1.** The Input Pattern and Classes of Our Rules

Command	Motion	Active Channel
Turn Right	Wink with right eye	1
Turn Left	Wink with left eye	2
Forward	Bite	1 and 2
Stop	Bite in the state of the Forward, Turn Right and Turn Left Commands	1 and 2 * *These thresholds are smaller than Forward command.
Forward to the Right	Wink with right eye in the state of the Forward Command	1
Forward to the Left	Wink with left eye in the state of the Forward Command	2

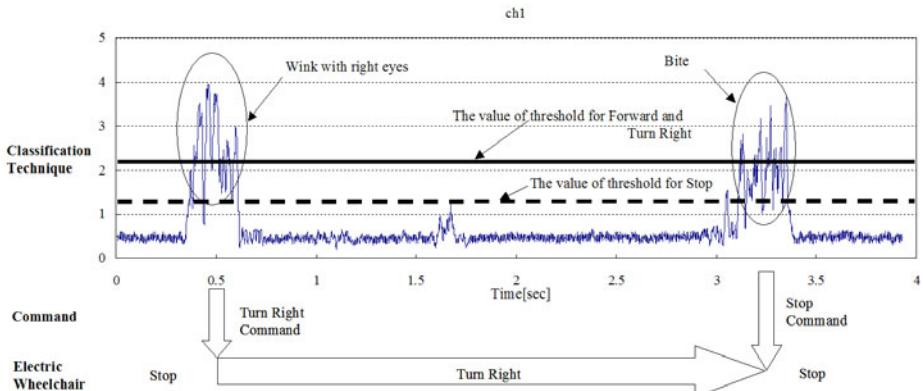
The s-EMG signal detected by surface electrodes is amplified and filtered prior to data acquisition, in order to reduce noise artefacts and enhance spectral components that contain information for data analysis. Two channels of s-EMG signals can be used to recognize face figure movement (Figure 2). 5 Ag/AgCl electrodes are used (two for each channel and one is for ground). In order to remove the DC level and 60 Hz power line noise, the high pass filter is used. The cutoff frequency of high pass

filter is 66.7Hz. After filtering and the amplification (about 470 times) stages, the s-EMG signals are digitized (16 bit) and then transferred to the personal computer. The sampling frequency of the measurement data is 1 KHz. And the band is from 0 KHz to 500 KHz. The s-EMG signals are then processed by a classification technique which is based on the threshold algorithm and the simple moving average method. This proposal method is necessary to set the value of threshold of each user. At the finally, the electric wheelchair is controlled by the output command from personal compute using the D/A converter.

The utilized facial movements, motion control command classes and effective channels of bio-signals are defined in Table 1. The electric wheelchair is controlled according to the rule of Table 1. From Table 1, it is easy to select Stop Command in this rule. Therefore, our system is safety, because Stop Command is selected in the false identification case. Figure 3 shows activation levels after filtering and the amplification stages of two channels s-EMG in commands (Right, Left and Forward/Stop). Figure 3 showed that three commands are different in activation levels of two channels s-EMG. Figure 4 showed that the image figure of motion command processing and classification technique. All parameters (the thresholds of Figure 4) of classification of each user are set first. Moreover, this system doesn't react to the usual blink.



**Fig. 3.** Activation Levels of Two Channels s-EMG



**Fig. 4.** The Motion Command and Classification Technique

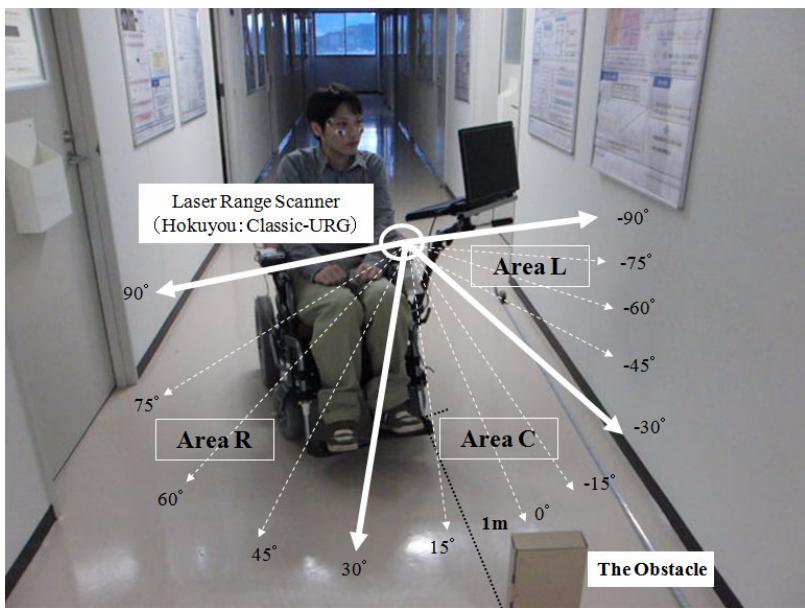
In the previous our work using  $k$ -Nearest Neighbor method, the system's correct classification rate was between 93% (5 class, the spoken motion could be distinguished) and 98% (3 class, same as Table 1) [7]. In this paper's technique, our system's correct classification rate is nearly the same as that in [7]. However, users have the ability to do good wink and bite motion. Moreover, users while operating this system are not possible to laugh and speak. The electric wheelchair stops when user speaks or laughs.

### 3 Functional Safety

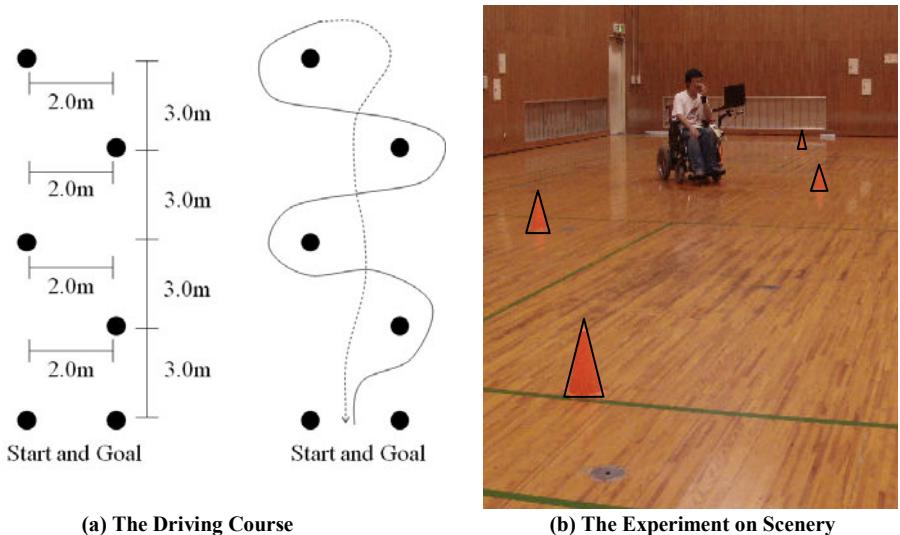
The obstacle evading function is important for the safety control of the electric wheelchair. In this paper, we develop the obstacle evasion function using a laser range scanner (HOKUYOU Automatic Co.,Ltd : Classic-URG) for the electric wheelchair safety control. The URG laser range scanner is the most compact obstacle detection sensor with distance measurement principle using laser. Scanning angle of our system is  $180^\circ$  (max  $240^\circ$ ). Our system measures the distance by  $15^\circ$  intervals. When the obstacle evading function detects the obstacle, our system sounds the alarm and speed control (Table 2). Our system is shown in Figure 5. Our system could detect about 1

**Table 2.** The Speed Control Rules

Area R / C / L (0 : no obstacle, 1: obstacle)	Power of Turn Right / Forward / Turn Left (100%:max speed , 0%:stop)
0 / 0 / 0	100% / 100% / 100%
0 / 0 / 1	100% / 100% / 75%
0 / 1 / 0	100% / 75% / 100%
0 / 1 / 1	100% / 75% / 75%
1 / 0 / 0	75% / 100% / 100%
1 / 0 / 1	75% / 100% / 75%
1 / 1 / 0	75% / 75% / 100%
1 / 1 / 1	75% / 75% / 75%



**Fig. 5.** Functional Safety



**Fig. 6.** The Experiment of Driving

meter before obstacle. Our system slows down the speed with the obstacle, and wards off a collision. It became easy to evade the obstacle with this function without stop. However, our system still has two or more problems in the detection of the wall etc. These are our future works.

## 4 Experimental Results

We tried the experiment of driving in the course by using our proposal system. The experimental subjects were three healthy men in twenties whose physique looks similar. One subject was an expert of the system, and others were the inexperienced subject. The driving course is shown in Figure 6 (a)(b). This driving course is about 32 seconds by walking. The task is to drive the electric wheelchair from position Start to Goal by contraction of facial muscles. The speed of electric wheelchair used middle and high speed for experienced subject. Inexperienced subject used middle speed only. We performed 5 trials. We compare our proposal system by average time [sec] and average miss count. The miss count is the number of error detection of classification in one trial.

The simulation results are shown in Table 3. The result of inexperienced subject 1 was almost same as the result of experienced subject in middle speed. From their experiments, we think that the control of our proposal system is easy. From these experiment results, the electric wheelchair control system using facial muscles s-EMG is a good system for the severely handicapped human. In addition, we think that the merit of our proposal system does not need a lot of training for user. Moreover, when the obstacle (the road cones) became near, our proposal system was able to be evaded smoothly without Stop command because the speed slowed by the obstacle evasion function.

**Table 3.** The experiment results

Subject	Average time [sec]	Average miss count
Experienced subject	97	4.0
Inexperienced subject 1	111	3.4
Inexperienced subject 2	135	3.4

## 5 Conclusions

In this paper, we applied to the electric wheelchair hands-free safety control system for the severely handicapped human using facial muscles s-EMG pattern recognition system. We tried the experiment of driving in the course by using our proposed system. The experimental subjects were three healthy men in twenties whose physique looks like. One person was an expert of the system, and others were the inexperienced person. From their experiments, we showed that the control of our proposed system is easy and our proposed system does not need a lot of training for user. Moreover, we think that our proposed system is useful than the system using the EEG signals.

In the future works, we plane to develop of algorithm that can be turned more smoothly, improve the obstacle evading function and test the severely handicapped human.

## Acknowledgement

This research has been supported by TATEISHI Science and Technology Foundation (No.1081019). The authors would like to thank Mr. Dohi for his constructive comments.

## References

1. Tomohiro, K., Tadashi, M., Tohru, K., Tsugutake, S.: Practical Usage of Surface Electromyogram, Biomechanism Library (2006)
2. Assareh, A., Konjkav, S., Fallah, A., Firoozabadi, S.M.P.: A New Approach for Navigating Automatic Wheelchairs using EMG signals Feature Extraction and Classification with an Adaptive Controller. In: Proc. The 12th International Conference on Biomedical Engineering, Singapore (December 2005)
3. Firoozabadi, S.M.P., Asghari Oskoei, M., Hu, H.: A Human-Computer Interface based on Forehead Multi-channel Bio-signals to control a virtual wheelchair. In: Proceedings of the 14th Iranian Conference on Biomedical Engineering (ICBME), pp. 272–277. Shahed University, Iran (February 2008)
4. Tamura, H., Okumura, D., Tanno, K.: A Study of Motion Recognition without FFT from Surface-EMG. The Journal of IEICE D J90-D(9), 2652–2655 (2007)
5. Okumura, D., Tamura, H., Tanno, K.: Proposal of the Motion Recognition system corresponding Sarface-EMG changes. In: Proceedings of Forum on Information Technology, FIT 2007, pp. G-022 (2007) (Japanese)
6. Tamure, H., Gotoh, T., Okumura, D., Tanaka, H., Tanno, K.: A Study on the s-EMG Pattern Recognition using Neural Network. International Journal of Innovative Computing, Information and Control 5(12), 4877–4884 (2009)
7. Manabe, T., Tamura, H., Tanno, K.: The Control Experiments of the Electric Wheelchair using s-EMG of Facial Muscles. In: Proceedings of Forum on Information Technology (FIT 2009), (CD-ROM), pp. K-008 (2009) (Japanese)
8. Jia, P., Hu, H.H., Lu, T., Yuan, K.: Head gesture recognition for hands-free control of an intelligent wheelchair. Industrial Robot: An International Journal 34(1), 60–68 (2007)

# A New Triple-Based Multi-robot System Architecture and Application in Soccer Robots

Xiangke Wang, Hui Zhang, Huimin Lu, and Zhiqiang Zheng

College of Mechatronics and Automation, National University of Defense Technology,  
Changsha 410073, China  
[xkwang@nudt.edu.cn](mailto:xkwang@nudt.edu.cn)

**Abstract.** Inspired by the organizational structures of some biological communities and social activities, a new multi-robot system architecture is proposed after the structure of multi-robot system being modeled as a triple  $\tau = (I, G, S)$ , which consists of the Integrated characters  $I$ , the Group characters  $G$  and the Single charters  $S$ . The proposed system architecture is divided into Integrated Intelligent Layer  $IIL$ , Group Intelligent Layer  $GIL$  and Single Intelligent Layer  $SIL$  sequentially. And then the information flow, the control flow and the response to dynamic environment are all analyzed. At last we report detailedly the application of proposed system architecture in NuBot team, a RoboCup Middle-Size League soccer robots team. In the view of triple-based system architecture, we designed different layers specially by imitating human soccer team naturally. This application, which has been run successfully for three years and is still developing, shows that the proposed triple-based system architecture is adaptable with distinct information/control flow, perfect compatibility and excellent flexibility in dynamic environment.

## 1 Introduction

Multi-robot systems have been attracted considerable attention in recent years, which naturally extends the research on single-robot systems. Because of intrinsic distribution and parallelism related to sensors and actuators, multi-robot systems have more advantages than single-robot systems, such as they can perform tasks more efficiently, then can accomplish the tasks that single one is disable to execute. Moreover, multiple-robot systems also increase the fault tolerance, provide more flexibility and so on [1–4].

System architecture is an importance issue in multi-robot field, which provides the infrastructure upon which collective behaviors are implemented and determines the capabilities and limitations of the system [5]. Existing system architectures generally fall into two categories: individual architecture and group architecture. The individual architecture mainly focuses on the decisions and the behaviors of the single robot. Many such architectures have been studied such as multi-layer hierarchical structure, behavior-based structure and hybrid structures [6–9]. The group architecture mainly pays attention to the logical

and physical control or information relationships between robots to lead cooperation, including centralized architecture, layered architecture and distributed architecture [1, 5].

Multi-robot system, as an artificial system, essentially imitates the group activities in nature or in society; and moreover the system architecture decides the operational mechanisms and cooperative efficiency of multi-robot system as the action of social institution to human society [1, 10]. Researches on system architecture can parallel to the related results in social-science, life-science and cognitive-science to some extent. In this study, parallel the organizational structure of group activities, we first propose a triple-based system architecture for multi-robot system. And then we report the application of the proposed system architecture in our RoboCup Middle-Size League soccer robots team – NuBot, which has run successfully for nearly three years in many competitions. The application shows that the proposed triple-based system architecture is adaptable with distinct control/information flow, perfect compatibility and excellent flexibility in dynamic environment.

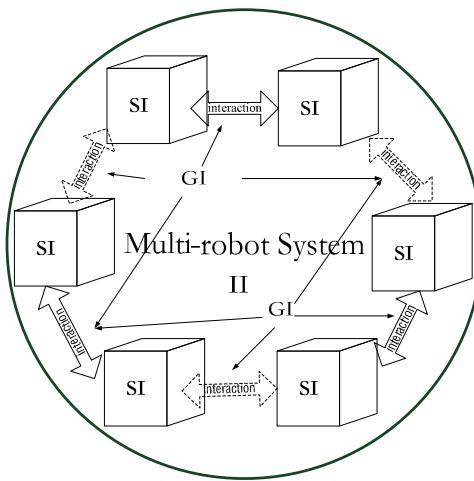
## 2 System Architecture of Multi-robot Based on Triple

Many individuals in natural or social group-activities can be organized and coordinated well to accomplish the tasks in the dynamic, competitive or cooperative environment, whose organizational structures, especially their perfect adaptability to dynamic environment, are available for multi-robot system architecture.

Let us investigate some lower biological communities with obvious social characters first. Here we take ant or bee colony as examples. Ants or bees build up their unique ways of life in thousands of years, which distinguish them from other insects. These ways of life are their integrated characters as species. Clear assignments and co-working also exist in ant or bee colony. By implicit or explicit communications, ants or bees cooperate to nest, prey and so on after different tasks are assigned to. These assignments and cooperations reflect their group characters. To accomplish assigned task, each ant or bee performs its respective function, which embodies the single character. Thus the organizational structure of some typical biological communities includes three aspects, viz. integrated character, group character and single character.

In the following we consider some soundly organized social activities, such as football game, basketball game, . . . . In general, each notable team is branded with special, vivid properties. As soon as these properties are mentioned, the related team will be associated with naturally. These properties are the integrated characters of the teams. When playing, teamwork is necessary for players to defense, score, . . . . Thus, group character is embodied in the teamwork. Each player shows its skill and specialty in competition, which reflects the single character. Similarly, the organizational structure of some soundly organized social activities also includes integrated character, group character and single character.

From above analysis, we know integrated character, group character and single character are all included in some lower biological communities and soundly organized social activities. This result could be borrowed to propose a new system architecture for multi-robot system. Abstractly multi-robot system could be modeled as a triple  $\tau = (I, G, S)$ . As an autonomous system, multi-robot system should own its intrinsic property, i.e. integrated character  $I$ , which distinguishes multi-robot system with other systems. Capacity of multi-robot system should be larger than single-robot, which means teamwork or group character exists in multi-robot system.  $G$  describes this property to embody teamwork into multi-robot system and make multiple robots harmonious. Each robot in multi-robot system is relatively independent and owns itself capacity, which is described by  $S$ . Sequentially these three characters foster three kinds of intelligences, namely Integrated Intelligence ( $II$ ), Group Intelligence ( $GI$ ) and Single Intelligence ( $SI$ ). Based on this triple model, multi-robot system architecture can be divided into three layers, viz. Integrated Intelligent Layer ( $IIL$ ), Group Intelligent Layer ( $GIL$ ) and Single Intelligent Layer ( $SIL$ ), as demonstrated in Fig. 1. Each intelligent layer relates to one character in the triple  $\tau$ . In the following, we will describe the details of three layers.



**Fig. 1.** Multi-robot system architecture based on triple.  $II$ : Integrated Intelligence;  $GI$ : Group Intelligence;  $SI$ : Single Intelligence.

*Integrated Intelligent Layer (IIL):* This layer addresses designer's expectations to multi-robot system and decides which 'intelligent level' is owned by system. It includes the objectives, performances and preferences of system. Operational mechanisms and basic styles of multi-robot system are also designed in this layer.

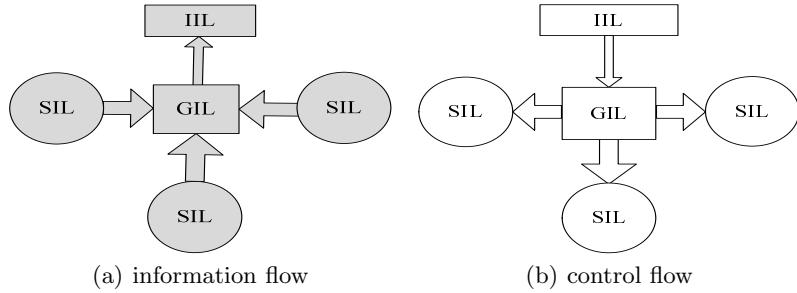
*Group Intelligent Layer (GIL):* This layer ensures the capacity of multi-robot system is larger than that of single-robot system. The environment should be evaluated in real time and different tasks should be decomposed and allocated to each robot after conflict resolutions in this layer. Communications, including implicit and explicit communications, often exist in multi-robot system, which are useful to improve the group intelligence and make the system harmonious. Thus designing proper mechanism for communications is also an important issue in this layer.

*Single Intelligent Layer (SIL):* Intelligence of single robot embodies in this layer. How to accomplish efficiently the allocated task for single robot is the main task of this layer. It may be with homogeneous or nonhomogeneous robots and can adopt the studies on individual architecture.

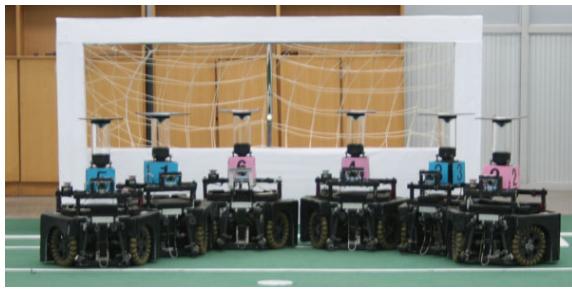
This triple-based multi-robot system architecture reflects the organizational structures of some biological communities and social activities to some extent; in the following, we will further design the information flow, the control flow and the response to dynamic environment. Just as that in biological communities or social activities, information flow in triple-based system architecture is bottom-up, i.e. from *SIL* to *IIL* (as Fig. 2(a)) and control flow is from *IIL* to *SIL* (as Fig. 2(b)). For information flow, single robot senses its environment in *SIL*, and then abstracts and sends the sensed and itself states to *GIL*; information from different robots is fused in *GIL*, and then the valuable part is sent to *IIL* with the group states. Whereas control flow directs reversely, which is up-down. Basic response to environment is generated in *IIL* first, and then transferred to *GIL*; based on the received basic response and the teamwork requirements, tasks of single robot are allocated and transmitted to *SIL* in *GIL*; accordingly, single robot plans and accomplishes its received task in *SIL*. Both the information flow and the control flow are more abstract whereas transferring or controlling frequency is lower in higher layer. The response to dynamic environment of proposed system architecture is as follows. The basic response is decided in *IIL*, and then in the view of receiving information, adjustments are performed in *GIL* until varies are fitted. If the vary is out of the range that *GIL* can deal with, then the new basic response will be regenerated in *IIL* and new the adjustments are run in *GIL*, ..., until varies are all fitted well.

### 3 Application in NuBot Team

In this section, we report the application of triple-based multi-robot system architecture in NuBot team. NuBot team is the RoboCup Middle Size League robot soccer team of National University of Defense Technology and mainly focuses on multi-robot cooperation, robust robot vision, motion control etc [11]. There are eight homogeneous robots in NuBot team, each of which is with omni-directional vision system and omni-directional motion system. Every robot is full autonomous and equipped with notebook, DSP TMS320F2812 control board, encoders and so on. Some NuBot robots are shown in Fig. 3. For convenience, we also develop an *ODE*-based simulation platform.



**Fig. 2.** Information flow and control flow of triple-based system architecture. *IIL*: Integrated Intelligent Layer; *GIL*: Group Intelligent Layer; *SIL*: Single Intelligent Layer.



**Fig. 3.** Robots of NuBot team

### 3.1 System Architecture of NuBot

We take a specified social actives, human being soccer team, as the natural model of NuBot team. An obvious fact is that human being soccer team is much more intelligent than robot soccer team currently, thus the inspiration is rational. Inspired by the organizational architecture of human being soccer team, we materialize the triple-based system architecture in NuBot. In the following, we elaborate the design of Integrated Intelligent Layer, Group Intelligent Layer and Single Intelligent Layer one by one.

**Integrated Intelligent Layer.** Integrated Intelligent Layer is also called Coach Layer straightforward in this study, because it mainly imitates the coach's intelligence. In human being soccer team, the coach responds to decide the basic strategy aiming at different opponents before the game begins, and adjusts the statics real time in the game. So, Coach Layer of NuBot includes two aspects: Out-Field Coach and In-Field Coach. Out-Field Coach responds the basic strategy imitating the coach's arrangements before game. Till now, three strategies are designed in NuBot, which are Attack First, Defence First and Balance between Attack and Defence. And eight roles are contributed, which are Goalkeeper, Attacker,

Passive, Mid-Field, Acid-Passive, Assistant, Blocker and Gazer. The number of the roles (equals 8) is always more than the number of active robots (no more than 6), so different role sets can be assigned to realize different strategies by Out-Field Coach. In-Field Coach evaluates the states between opponents and teammates and then adjusts the statics in the game. A simple adjustment in NuBot is: if the score keeps ahead and the remaining time is not too long, the ‘Defence First’ strategy will be chosen by priority; accordingly if the score keeps behind and the remaining time is not too long, the ‘Attack First’ strategy will be chosen.

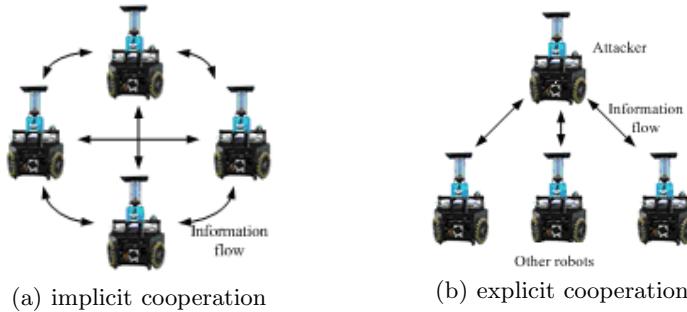
**Group Intelligent Layer.** Group Intelligent Layer mainly imitates the captain or the decision-maker in the field. It allocates the roles of the robots and initiates the cooperation.

A hybrid distributed role allocation method is developed for NuBot, which includes role performance evaluation, assignment based on rule and dynamic application. Every robot firstly evaluates the costs of being treated as Attacker and Passive for itself and all the other active robots, and selects the most proper robots as Attacker and Passive [12]. Then, the remaining roles are assigned based on rules [13]. At last, the dynamic application is used to avoid the repeated role assignments for the inconsistent information between different robots.

Cooperation includes implicit cooperation and explicit cooperation in NuBot. All robots maintain a common world model. And in implicit cooperation, robots realize cooperations by sharing the information in the world model. Taking ‘defence action’ as an example. Each active robot decides its defensive action based on its self-location, the distance to ball and the teammates’ actions in the sharing world model. Every robot is equal in implicit cooperation and the information flows between robots are shown as Fig. 4(a). In explicit cooperation, robots complete cooperations by direct communication. The cooperation is initiated by one of the robots, called initiator. The initiator selects and informs one or more teammates to complete the tactics. When the tactics is finished, the cooperative relationship relieves immediately. For example, the tactics of free kick in NuBot team is an explicit cooperation. Robots are not equal in explicit cooperation and initiator dominates the cooperation. The information flows of explicit cooperation are shown as Fig. 4(b).

**Single Intelligent Layer.** Single Intelligent Layer imitates the single player, and is driven by role. Single Intelligent Layer executes different action according to the role allocated by Group Intelligent Layer or the action assigned in explicit cooperation. The structure of Single Intelligent Layer is shown as Fig. 5, which includes Sensor Hierarchy and Action Hierarchy [14]. Sensor Hierarchy processes the data from different sensors, constructs and maintains the World Model. Action Hierarchy controls the robot to complete the tasks.

Action Hierarchy is divided into four different parts, and each part includes different intelligence and complexity behavior sets. The four parts are Action



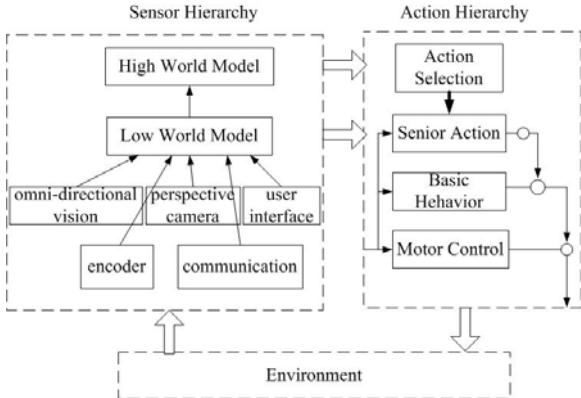
**Fig. 4.** Information flow in NuBot Team

Selection, Senior Action, Basic Behavior and Motor Control. Every part gets its required information from World Model respectively and is just influenced by the lower part.

**Information Flow and Control Flow.** Information flow and control flow are both important issues for system architecture. The information flow is from bottom (*SIL*) to up (*IIL*), just as the analysis in above section. Especially, information in *SIL* flows more and more abstractly from bottom to up in Sensor Hierarchy too. The lowest includes the original information obtained from different sensors. Low World Model includes the basic knowledge abstracted from original information, such as the robot self-location, the teammates' locations, the ball's position and the obstacles' positions. High World Model abstracts further the information from Low World Model, which includes the ball trace estimation, the ball speed estimation and so on. *GIL* includes information from different robots and then fuses to obtain the teammates' and enemies' locations, the exact ball's location, which robot is controlling or will control ball and what are the current roles of teammates, ..., . *IIL* records and deals with the highest-level information such as the team's strategy, states or scores. Control commands direct from *IIL* to *SIL*. Which strategy (Attack First, Defence First, and Balance between Attack and Defence) should be performed is selected in *IIL*, that is the basic response for NuBot in competitions. Based on the selected strategy, roles are allocated to every robot in *GIL*. And at last every robot selects its behavior and runs action according to its role. Obvious, distinct information flow and control flow exist in this system architecture.

### 3.2 Application in Competition

Triple-based system architecture was developed in NuBot Team from 2007 and now is still developing. Since 2007, NuBot Team has participated in RoboCup 2008 Suzhou and RoboCup 2009 Graz with entering the second round each time. It also took part in RoboCup China Open 2007 - 2009 and wined the first-place prizes in 2007, 2008, third-place in 2009. These achievements are certainly

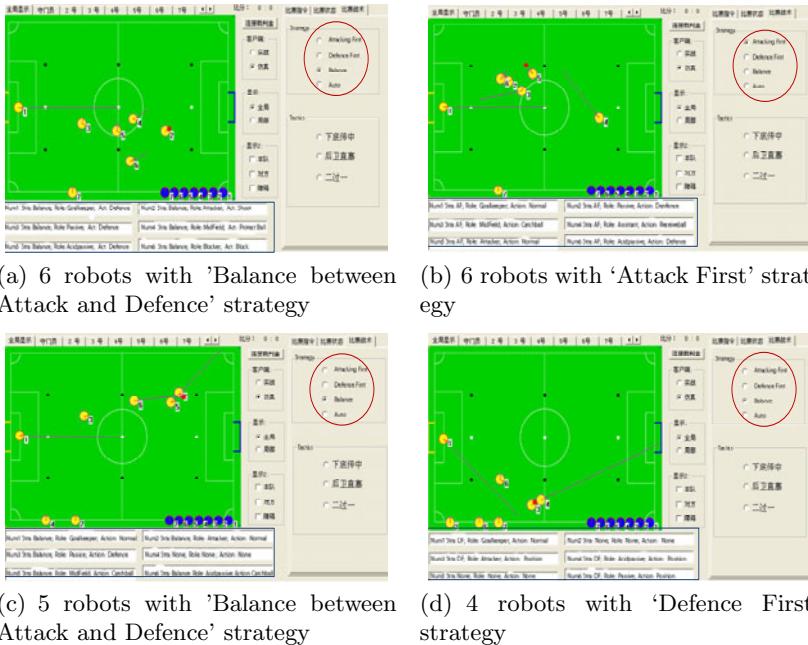


**Fig. 5.** The structure of Single Intelligent Layer

dependent on the whole of multi-robot system, including mechanism, sensing, control and so on; whereas there is no denying that this triple-based system architecture also plays an importance role. This architecture can adopted in such a complex and dynamic completion, which proves its excellent flexibility in dynamic environment. In the three-year developing, different ideals and plans are embodied in this frame with perfect compatibility. To report clearly the effect of proposed system architecture, we catch and explain some typical scenarios from our *ODE*-based simulation first.

In Fig. 6, items marked by red-oval are the strategies appointed in Coach Layer; and items marked by blue-box are the roles allocated in *GIL* and the senior behaviors selected in *SIL*. There are different appointed strategies or different robot numbers in these four scenarios. Fig. 6(a) shows a typical scenario that 6 active robots with the ‘Balance between Attack and Defence’ strategy appointed by Coach Layer. Sequentially, six roles are allocated for these robots automatically in *GIL*, which are Goalkeeper, Attacker, Passive, Mid-Field, Acid-Passive and Blocker. Fig. 6(b) also shows 6 active robots scenario, whereas the appointed strategy is ‘Attack First’. The role Blocker is removed and role Assistant is added by *GIL* accordingly. The scenario in Fig. 6(c) is appointed the same strategy as that in Fig. 6(a), but just with 5 active robots, thus the role Blocker disappears. Fig. 6(d) shows 4 active robots with ‘Defence First’ strategy scenario, in which the allocated roles are just Goalkeeper, Attacker, Passive and Acid-Passive. These scenarios show different roles are allocated with different strategies and robot numbers to adapt different competitions. So system architecture in NuBot reflect the intelligences of Coach Layer and *GIL* in some sense.

In Fig. 6, yellow circles in green field represent robots. The centers of circles mean robots’ locations and the red radiuses directed from centers mean robots’ attitudes. Teamwork between multi-robots are reflected from the distribution of robots’ locations and attitudes. For example, scenarios in Fig. 6(a) and 6(c)



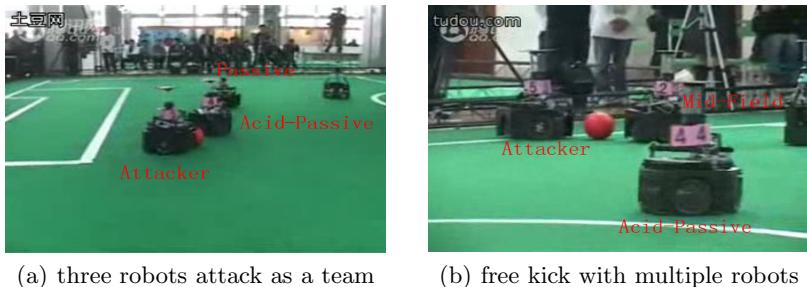
**Fig. 6.** Typical scenarios of applying triple-based system architecture in NuBot

show multiple robots attack as a team and scenario in Fig. 6(b) shows multiple robots defense as a team. These cooperations are all implicit. Playing free kick with multiple robots is shown in Fig. 6(d), in which robot 3 shoots after the ball is push gently to by robot 4. Obviously, the cooperation in this scenario is explicit.

The blue-box marked items also explain the senior behaviors selected in *SIL* in Fig. 6. For example, when multiple robots attack as a team in Fig. 6(a), the selected action of Attacker is ‘shoot’, i.e. dribbling and preparing to shoot; the selected action of Mid-Field is ‘protect ball’, i.e. protecting the ball not to catch by enemy; . . . Every robot will select different senior action in different situation automatically.

System architecture of NuBot also works well in competitions. Here we give two scenarios caught from the video of the final in RoboCup China Open 2008. (see <http://www.tudou.com/programs/view/8XCpS44yX18/>).

Fig. 7(a) shows three robots initiate their attack, where Attacker dribbles the ball, Acid-Passive protects the ball and Passive positions rightly to block the enemy. Fig. 7(b) shows the free kick tactics, where Mid-Field pushes the ball gently to Attacker, and then Attacker shoots. Acid-Passive is in right position to protect the ball. These two scenarios exhibits partly the proposed system architecture work well in practise and adapt in dynamic environment.



**Fig. 7.** Typical scenarios of NuBot Team in Competition

## 4 Conclusion

Paralleled to the organizational architecture of some biological communities and social activities, we proposed a new triple-based multi-robot system architecture and reported its application in NuBot Team. The triple  $\tau = (I, G, S)$  consists of the Integrated characters  $I$ , the Group characters  $G$  and the Single charters  $S$ ; system architecture is divided into Integrated Intelligent Layer, Group Intelligent Layer and Single Intelligent Layer sequentially. An implementation of the proposed system architecture, which has been run successfully for three years and is still developing, is elaborated. We design the *IIL*, *GIL* and *SIL* in NuBot team by imitating human being soccer team naturally. The application in simulations and competitions shows that proposed triple-based system architecture is adoptable with distinct control/information flow, perfect compatibility and excellent flexibility in dynamic environment.

## References

1. Tan, M., Wang, S., Cao, Z.Q.: Multi-robot System. Tsinghua University Press, Beijing
2. Zhang, F., Lin, L.M.: Architecture and related problems concerning cooperative mobile robot system. Robot 23(6), 554–558 (2001)
3. Arkin, R.C., Balch, T.: Cooperative multiagent robotic systems. In: Artificial Intelligence and Mobile Robots. MIT/AAAI Press, Cambridge (1998)
4. Yuan, K., Li, Y., Fang, L.X.: Multiple mobile robot systems: A survey of recent work. Acta Automatica Sinica 33(8), 785–794 (2007)
5. Cao, Y.U., Fukunaga, A.S., Kahng, A.B.: Cooperative mobile robotics: Antecedents and directions. Autonomous Robotics 4(1), 7–27 (1997)
6. Saridis, G.: Toward the realization of intelligent controls. Proceedings of the IEEE 67(8), 1113–1115 (1979)
7. Brooks, R.A.: A robust layered control system for a mobile robot. IEEE Journal of Robotics and Automation 2(1), 14–23 (1986)
8. Arkin, R.C.: Motor schema-based mobile robot navigation. The International Journal of Robotics Research 8(4), 92–112 (1998)

9. Connell, J.H.: Sss: A hybrid architecture applied to robot navigation. In: Proceedings of the 1992 IEEE Conference on Robotics and Automation, Los Alamitos, pp. 2719–2724 (May 1992)
10. Gao, Z.J., Yan, G.Z., Ding, G.Q., Yan, D.T., Chen, Z.Z.: Current status and development of multirobot coordinative and cooperative systems. Optics and Precision Engineering 9(2), 99–103 (2007)
11. Zhang, H., Wang, X., Lu, H., Ji, X.: Nubot team description paper 2009. In: RoboCup 2009 Graz, Graz, Austria (2009)
12. Liu, L., Ji, X.C., Zheng, Z.Q.: Multi-robot task allocation based on market and capability classification. Robot 28(3), 338–343 (2006)
13. Ji, X.C., Cui, L.H., Zheng, Z.Q.: Application of a hybrid distributed multi-agent task allocation mechanism to soccer robot system. Computer Application 28(3), 706–709 (2008)
14. Ji, X.C., Liu, L., Zheng, Z.Q.: A modular hierarchical architecture for autonomous robots based on task-driven behaviors. In: Proceedings of the International Conference on Sensing, Computing and Automation, Chongqing, China, pp. 631–636 (May 2006)

# Pursuit Evasion in Dynamic Environments with Visibility Constraints

Ibrahim Al-Bluwi<sup>1</sup> and Ashraf Elnagar<sup>2</sup>

<sup>1</sup> LAAS-CNRS, Universite de Toulouse, F-31077, Toulouse, France  
[a1.bluwi@laas.fr](mailto:a1.bluwi@laas.fr)

<sup>2</sup> Dept. of Computer science, Univ. of Sharjah, Sharjah, UAE  
[ashraf@sharjah.ac.ae](mailto:ashraf@sharjah.ac.ae)

**Abstract.** This paper presents a novel approach for the problem of tracking a moving target in a dynamic environment. The robot has to move such that it keeps the target visible for the longest time possible, and at the same time, avoid colliding with any of the moving obstacles. This paper presents a solution that is based on the idea of three interacting components which perform: tracking, collision avoidance and motion selection. The proposed solution is validated using a comprehensive set of simulations, which show that transition from tracking in static environments to tracking in dynamic environments can be done without much loss in robot safety or tracking ability.

## 1 Introduction

A recent study estimated the number of robots world-wide that were in operation in 2007 to be close to 6.5 million [4]. Today, the use of robots is not restricted to factories and highly structured environments any more. Robots today perform daily-life tasks like, vacuum cleaning and grass mowing. They also guide people in museums and hospitals assist the elderly and blind and work as security watchmen. In fact, applications of robots are not restricted to the aforementioned, and are continuously increasing in number and type. This expansion puts today's robots against two main challenges. Namely, the ability to perform more complex tasks and to work in unpredictable and unknown environments. This paper considers the task of target following, which is a natural task performed by humans. It also assumes that moving objects exist around the robot performing the tracking task, which reflects the nature of our daily-life environments. Having robots with such a capability opens the door for many applications in several domains such as medicine, security, and home entertainment.

The importance of this study is of two fold. The first is that it addresses a lively problem that has applications in various domains and the second is that it contributes to filling the gap in current research on the problem of autonomous tracking. Current research assumes that tracking occurs in a static environment, where the target is the only moving object in the environment. This assumption allows concentrating on finding a solution for the tracking problem, however, it is very restricting as most application domains are in dynamic environments.

The proposed solution in this paper is based on an architecture that is made of three distinct but interacting components. These components are called the occlusion advisor, collision advisor and decision maker. The occlusion advisor is dedicated for tracking, the collision advisor for collision avoidance and the decision maker for motion selection. The robot moves in the environments trying to minimize two risks: the risk of loosing the target and the risk of colliding with a moving object.

The paper is organized as follows: Section 2 surveys previous work on tracking in dynamic environments. The problem is then further described and formalized in Section 3. Occlusion advisor, collision advisor, and decision maker components are explained in Sections 4 and 5, respectively. Simulations and results are discussed in Section 6 and the paper then ends with a section for conclusions and possible future work.

## 2 Previous Work

Research related to the problem of tracking in dynamic environments can be categorized into two groups: work done on tracking in static environments and work done on motion-planning in dynamic environments. The following is a survey of work in these groups.

The problem of tracking in a static environment is formally defined and analyzed using Game Theory, [16]. In other works [19], the problem is modeled as a motion planning problem of a rod, of variable length.

It was shown that the problem of tracking a target around one corner is completely decidable, [2]. On the other hand, the decidability of the general problem of tracking in a cluttered environment was found to be at least NP complete, [20]. In [3], the environment is divided into decidable and undecidable regions and heuristics were used to approximate the bounds of the decidable regions.

All of the above mentioned results are of theoretical significance for the problem of autonomous tracking. However, practical solutions are also reported. In [10] a tracking algorithm is introduced that maximizes the target's shortest distance to escape. Similarly, [4] uses a greedy local strategy that is based on a different risk function, which tries to achieve balance between not loosing the target while keeping the target visible for the longest time.

A plethora of algorithms have been introduced to address the problem in dynamic environments. For obstacles with a predictable motion track, several extensions of available planning algorithms in static environments are proposed. A joint time-configuration space is built which captures the motions of the moving obstacles and planning is done offline in the joint space (e.g., see [11]). Another approach is to use the velocity decomposition technique, [14].

Online algorithms also exist which assume that the motions of the obstacles are not known in advance. Some adopt a plan-then-fix strategy, like the D\* [15] and the DRRT [9] which repeatedly observes the environment and fixes an initial RRT plan by adding and removing nodes. A similar approach is used with the

PRM [12]. Other algorithms adopt a plan-then-improve strategy. A rough plan is initially created and then is improved over time to adopt for the changes in the environment, [21]. Such planners are called anytime planners since they are ready to give a plan any time. On the other hand, algorithms also exist for predicting the motions of the obstacles [7]. These algorithms can be used to enhance the performance of the above mentioned algorithms.

Previous work has addressed the two distinct problems: tracking in static environments and motion planning in dynamic environments separately. This paper studies the effect of addressing both problems together and how the performance of the robot is affected while tracking in a global dynamic environment amidst randomly-moving obstacles.

### 3 Problem Description

We use the terms pursuer and evader to refer to the robot and the target respectively. Both, the pursuer and the evader are assumed to be rigid objects that are not restricted by any constraints other than that they have the same velocity bound. The evader is assumed to move unpredictably in the environment and the pursuer has the ability to detect the evader and the dynamic obstacles and determine their relative locations. Next, we formally state our problem similar to [8].

Let  $p$  and  $e$  denote the pursuer and evader respectively. Without loss of generality, we assume  $p$  and  $e$  to be points moving in the configuration space. We refer to a time step in the tracking problem with  $t_k$ , where  $0 \leq k \leq T$ . The current state of the pursuer at time  $t_k$  is denoted by  $x_k^p$ . It belongs to the state space of the pursuer  $\chi^p$ . Similarly, the current state of the evader is denoted by  $x_k^e$  and it belongs to  $\chi^e$ ; the evader's state space.

At each time step, the pursuer and evader choose an action  $a_k^p$  and  $a_k^e$  from their respective action spaces  $\mathcal{A}^p$  and  $\mathcal{A}^e$ . Tracking takes place in the Euclidean workspace  $\mathcal{W} \subset \mathbb{R}^2$ . This workspace is cluttered with a set of static obstacles which occupy the space  $\mathcal{W}_{obs}$ , and the free part of  $\mathcal{W}$  is referred to as  $\mathcal{W}_{free}$ . Thus,  $\mathcal{W} = \mathcal{W}_{obs} \cup \mathcal{W}_{free}$ . The set of dynamic obstacles in  $\mathcal{W}$  is referred to as  $\mathcal{O}$ . At each time step, we are interested only in dynamic obstacles that are currently visible to the pursuer. We denote this subset of  $\mathcal{O}$  with  $\mathcal{O}_k$ .

We say a point  $q$  is visible to  $p$  if and only if it can be connected to it with a straight line  $\ell$ . Two constraints govern  $\ell$ . First, it should lie completely inside  $\mathcal{W}_{free}$ . Second, the length of  $\ell$  should be less than  $\rho$ , where  $\rho$  is the maximum reachable distance by the pursuer's sensors. We are now able to define  $\mathcal{V} \subseteq \mathcal{W}_{free}$  as the *visibility region* of  $p$ . Any point  $q \in \mathcal{V}$  can be connected to  $p$  with  $\ell$ . Note that our definition of  $\mathcal{V}$  prohibits the intersection of  $\ell$  with obstacles in  $\mathcal{W}_{obs}$  and not in  $\mathcal{O}$ . This means that we assume in our problem that dynamic obstacles do not cause occlusion. Another important note is that when tracking starts, we assume that the evader is initially detected, i.e.  $e \in \mathcal{V}$ .

It is now possible to define our tracking problem as a decision problem over the triplet  $\{\Omega, \mathcal{D}, \mathcal{C}\}$ , where:

- $\Omega$  is the *problem state space*. It represents all possible combinations of evader and pursuer states with possible states of dynamic obstacles. At time  $t_k$  the problem state is denoted as  $\omega_k$  and it captures  $x_k^p, x_k^e$  and the states of each obstacle in  $\mathcal{O}$  at that time instance.
- $\mathcal{D}$  is the set of possible *decisions*. A decision  $d_k \in \mathcal{D}$  made at time  $t_k$  is a choice of action  $a^p$  that the pursuer makes after observing  $\Omega$ .
- $\mathcal{C}$  is the set of possible *consequences*. A consequence  $c(d_k) \in \mathcal{C}$  is a result or an outcome that occurs when the pursuer makes the decision  $d_k$ .

Depending on the goal which the pursuer would like to achieve, it will have a preference for some consequences over others. Based on this preference, the pursuer should choose its decisions from  $\mathcal{D}$  at each time step. More specifically, each consequence  $c$  has a utility value that quantifies to what extent the pursuer likes the consequence. This utility is defined as the function  $\mathcal{U}(d_k)$  over the decision that caused this consequence. We use the notation  $\succeq$  to denote preference between decisions and pronounce  $d_k \succeq d'_k$  as  $d_k$  is preferred to  $d'_k$ . We say  $d_k \succeq d'_k$  only if  $\mathcal{U}(d_k) \geq \mathcal{U}(d'_k)$ .

The goal of the pursuer in our tracking problem now becomes to choose  $d_k$  at each time step such that:  $d_k = \arg \max_{d \in \mathcal{D}} \{\mathcal{U}(d)\}$ .

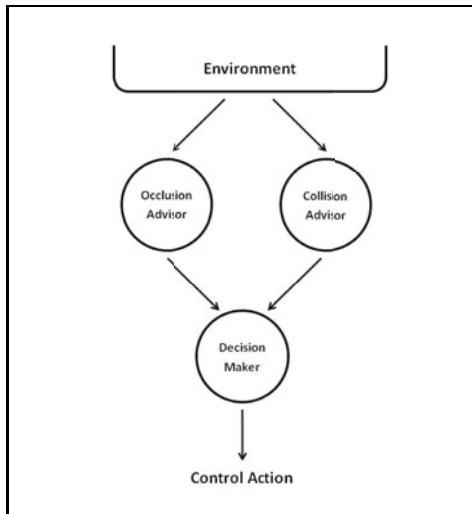
Our tracking problem is now clearly dependent on how we define the utility function  $\mathcal{U}$ . Roughly speaking,  $\mathcal{U}$  should award more decisions that make the pursuer less susceptible to loosing the evader and less susceptible to colliding with a dynamic obstacle at the same time. In the following, we give two general definitions for  $\mathcal{U}$  depending on the tracking goal which the pursuer tries to achieve.

## 4 Solution Overview

Our solution is intuitively divided into three separate components. The first component helps the pursuer evaluate the tracking status. The second helps it evaluate the safety status. The third component helps it decide where to move, based on information acquired from the first and second components. We refer to these components as the "occlusion advisor", the "collision advisor" and the "decision maker" respectively. The relationship between these components is depicted in Figure II.

At each time step, the robot will observe the current state of the environment and two risk values are produced. The first is by the occlusion advisor and it represents the likelihood of loosing the evader and the second is by the collision advisor and it represents the likelihood of colliding with a moving obstacle. The decision maker then uses these two risk values to choose a control action that is best in terms of tracking effectiveness and robot safety. Such a solution has several advantages:

- Analysis and testing of each of the three components can be done separately without affecting the others. Future modifications to the algorithms can also be done separately.
- Since the solution is reactive (i.e. repeatedly observes the environment and then reacts), it is possible to apply it in new and unknown environments.
- No need to worry about localization issues as the robot does not use any map and relies only on observed information.



**Fig. 1.** Solution Architecture

#### 4.1 The Occlusion Advisor

Given the current states of both the pursuer and the evader, the responsibility of the occlusion advisor is to compute a risk factor,  $\mathcal{R}^{occ}$ , which describes the likelihood of loosing sight of the evader.

Occlusion occurs when the evader moves outside the visibility range of the pursuer, or when  $e \notin \mathcal{V}$ . The goal of the pursuer in the tracking problem is to prevent occlusion from happening. Therefore, the first step to compute the occlusion risk is to identify obstacle vertices that can cause occlusion. Each occlusion vertex will contribute a risk value. The sum of such risks will make the overall risk value. The choice of risk functions is determined by the tracking objective. We consider two risk functions, one for maximizing escape time and another for maximizing exposure time.

Generally speaking, occluding vertices are defined by the shape of the obstacle and the position of the pursuer facing the obstacle at that time. Formally, given a pursuer  $p$  and an obstacle  $o$ , a vertex  $v \in o$  is said to be an occluding vertex if and only if  $v$  is visible to  $p$ , and the left and right neighboring vertices of  $v$  both strictly lie on the same side of the infinite line  $\ell$  passing through  $p$  and  $v$ .

In order to compute the occluding risk, we need to locate occluding vertices, codetermine the risk factors, and compute the risk by aggregating all risk functions (details are in [5]). After finding all occluding vertices and computing the risk for each of these vertices separately, the total occlusion risk,  $\mathcal{R}^{occ}$ , can be computed which should be more than or equal to the maximum of these risk values. This can be done using the following formula:

$$\mathcal{R}^{occ} = maxR + (1 - maxR) \left( \frac{1}{n-1} \sum_{\substack{i=1, \rho_i \neq maxR}}^n \rho_i \right) \quad (1)$$

where  $maxR$  is the maximum risk value among the computed risks and  $n$  is the number of risk values. What this function does is that it adds a factor to the maximum risk value. This factor depends on the other computed risk values; the higher these risk values are, the higher the factor is.

We call this method the *pessimistic risk computation* and we use it in our simulations for computing the value of  $\mathcal{R}^{occ}$ . It is important, however, to note that applying this method may cause jerks in the motion of the robot as shown in [17]. As the robot moves, it encounters new occluding vertices. These new vertices may have radically higher risk values, which will cause the robot to adjust its motion abruptly towards minimizing the new risk value. Using an average will lessen the sharpness of these jerks but will not eliminate it.

## 4.2 Collision Advisor

When dealing with unknown dynamic obstacles, it is very useful to be able to predict the future movements of these obstacles. This allows planners to create plans that are safer and less likely to change. In fact, motion prediction is a natural task that we perform in our daily life as we move around and avoid colliding with other moving entities like people and cars. In the collision risk function to be introduced in the next section, we assume that a motion prediction algorithm exists and is used by the pursuer.

Many algorithms exist in the literature for motion prediction. For example, Kalman Filters [13], Brownian models [18] and Auto-regressive models ([6] and [7]) have been proposed to produce predictions based on an observed short history of the obstacle's movements. Other approaches also follow a learn-then-predict approach [22]. In such approaches, the environment is first observed and motion patterns are recorded offline. A model is then created based on the observed data and is used to produce predictions on-line.

Assume that a robot is placed at some room and due to some unknown error it starts moving around itself randomly. To be safe from being hit by this robot, the intuitive behavior is to keep away from it. The farther, the better, and the direction does not matter since it is not known where the robot will move next. Consider on the other hand the case of crossing a road. Cars move in one direction and to be safe from being hit by a moving car, one has to avoid being in front of it. Therefore, it is usually safe to get close to the car from its sides

and its rear. These two scenarios depict the difference between the case of predictable and unpredictable obstacle motions. In the first case, it is not possible to predict where the malfunctioning robot is going to move next, so it is required to maintain a clearance distance from the robot in all directions. However, it is possible in the second case to predict where the car will move next, so a clearance distance can be maintained based only on the direction of the predicted motion. Hence we distinguish between two risk types: random risk and prediction risk.

Random risk is independent of the motion direction of the obstacle. It represents the likelihood of colliding with an obstacle regardless of where this obstacle is moving. Such a risk can be represented as follows:

$$\text{RandomRisk} = 1 - \frac{d^m}{d_{max}^m} \quad (2)$$

where  $d$  is the distance to the obstacle,  $d_{max}$  is the maximum distance possible to the obstacle and  $m$  is a scaling factor. This equation creates a normalized risk value, where a closer distance to the obstacle yields a higher risk value.

Prediction risk, on the other hand, is dependent on where the observer anticipates the obstacle will move next. If the observer knows for sure the very location of the next move, then the risk will be 1 at that location and 0 everywhere else. However, it is usually not possible to predict the exact location of the next move. There is always a degree of uncertainty based on the amount of information available and the prediction method used.

We model prediction risk as a circle around the prediction point, where the radius of this circle depends on the confidence in this prediction. If the prediction is guaranteed to be correct, then the radius is 0, otherwise, the radius of the circle will grow proportional to the lack of confidence. Therefore, Prediction risk can be modeled using the following equation:

$$\text{PredictionRisk} = (1 - \text{confidence}) * (1 - \frac{dP^m}{dP_{max}^m}) \quad (3)$$

where  $dP^m$  is the distance to the prediction point,  $dP_{max}^m$  is the maximum distance possible to the prediction point, and  $\text{confidence}$  is a value between 0 and 1 that represents our belief in the validity of the prediction.

We represent our collision risk function for one dynamic obstacle as a weighted sum of these two risks, where increasing the weight of one risk causes a decrease in the weight of the other.

$$\varphi = (1 - \lambda)(\text{RandomRisk}) + (\lambda)(\text{PredictionRisk}) \quad (4)$$

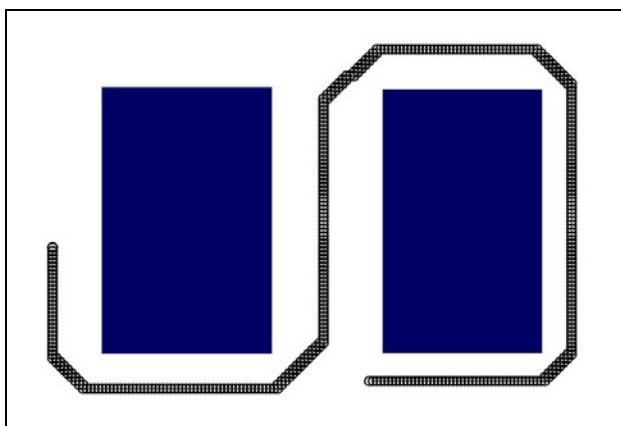
The weight factor  $\lambda$  may be interpreted as the prediction confidence.

### 4.3 Decision Maker

The overall risk is computed by taking a weighted average of the two major risks of occlusion and collision. Namely,  $\text{overallRisk}$  is:

$$\text{overallRisk}(\omega_k) = (\lambda)(\mathcal{R}^{occ}) + (1 - \lambda)(\mathcal{R}^{coll}) \quad (5)$$

where  $\lambda$  is between 0 and 1. Such a weighted average allows a lot of flexibility when using the algorithm in different domains. Depending on the application, the pursuer can be tuned with a  $\lambda$  that suits it. For example, if tracking is more important than the safety of the pursuer, then  $\lambda$  can be given a value greater than 0.5. If the opposite is true, then  $\lambda$  can be given a value that is less than 0.5. Moreover, dynamic obstacles in some environments can be soft in the sense that they do not cause a lot of harm if they collide with the pursuer, whereas dynamic obstacles in other environments can be very dangerous. For example, an environment cluttered with humans is not equal in its danger to an environment that is cluttered with speedy cars. A pursuer tracking in the human-cluttered environment can use a high value for  $\lambda$  whereas a pursuer tracking in a car-cluttered environment needs to use a low value for  $\lambda$ .



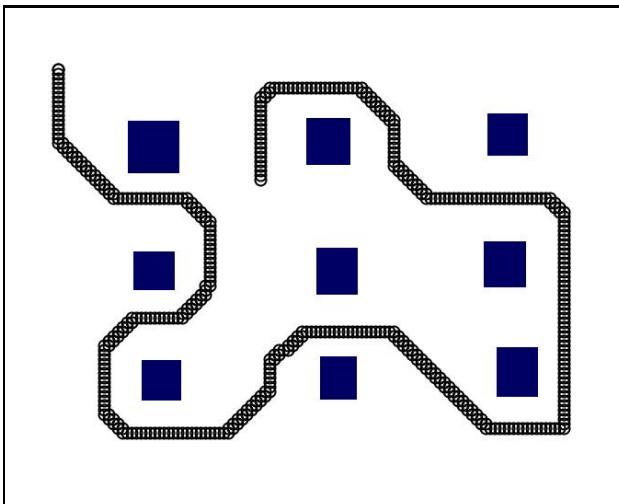
**Fig. 2.** The first tracking scenario

## 5 Simulations and Results

It was noticed that changing the evader's path in the same environment changes the results of the simulation. In addition, changing the environment's shape affects the results of the simulation. Many such parameters affect the performance of the pursuer. The following are the parameters to be considered in the simulations of this work.

- *Number of corners in the environment:* The higher the number of corners is, the harder tracking becomes. This is because the evader has more escape choices and the pursuer has to take them all in mind.
- *Environment Density:* It is the percentage of the environment that is occupied by obstacles. The higher the density is, the harder tracking is. This is because the pursuer has less motion choices in a dense environment.
- *Number of Dynamic Obstacles:* The higher the number is, the harder tracking becomes. This is because the pursuer has to perform more collision avoidance motions, which may negatively affect the tracking process.

The goal of this work is to propose a method for tracking in dynamic environments, therefore, the simulations of this chapter will study the effect of the presence of dynamic obstacles in different environments on the tracking process. We would like to study the difference between the pursuer's performance in static environments and its performance in dynamic environments. To answer



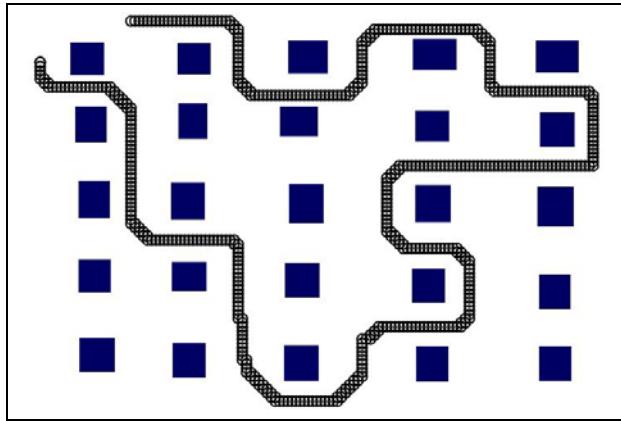
**Fig. 3.** The second tracking scenario

these questions, we have used the scenario builder to create three environments, each of which is 800x600 pixels in size and with a pursuer that has a visibility range of 150 pixels. The following is a description of these environments.

1. *High density, few corners:* This environment has two large static obstacles (with 8 corners) covering 36% of the workspace. The evader moves between these two obstacles 404 steps. Figure 2 shows the environment and the evader's path.
2. *Low density, moderate corners:* This environment is made up of a 3x3 matrix of static obstacles covering around 4% of the workspace. This environment has 36 corners and the evader moves between them a total of 440 steps. Figure 3 shows the environment and the evader's path.
3. *High density several corners:* This environment is made up of a 5x5 matrix of static obstacles covering around 10% of the workspace. This environment has 100 corners and the evader moves between them a total of 494 steps. Figure 4 shows the environment and the evader's path.

For each of these environments the simulation is first run without any dynamic obstacles. After recording the results, three sets of simulations are performed for each of these environments, such that the first set has 10 dynamic obstacles, the second has 20 and the third has 30 dynamic obstacles.

Dynamic obstacles are placed randomly in the environment. Each dynamic obstacle moves in a straight line in an arbitrary direction. The dynamic obstacle moves until it collides with a static obstacle or it reaches the sides of the environment. Thence, the dynamic obstacle changes its direction to a new arbitrary one provided that the new direction does not make it head out side the environment or into a static obstacle.



**Fig. 4.** The third tracking scenario

This random behavior of the dynamic obstacles makes their motions unpredictable, which makes also the results of different runs for the same simulation unpredictably different. For this reason, we have repeated simulations that have dynamic obstacles 30 times each. For example, the 3x3 scenario is first run without dynamic obstacles once, then it is run 30 times with 10 dynamic obstacles, 30 times with 20 dynamic obstacles and 30 times again with 30 dynamic obstacles. The same is also done for the other two environments, which makes the total number of simulation sets nine, plus three simulations in static environments.

**Table 1.** Summary of simulation results

Averages	Scenario	No Dyn. Obs.	10 Dyn. Obs.	20 Dyn. Obs.	30 Dyn. Obs.
% of Time Visible	First	97.8%	89.5%	89.5%	86.6%
Number of Collisions		0	0.6	2.6	5.1
% of Time Visible	Second	100%	96.5%	92.9%	90.6%
Number of Collisions		0	1	1.6	2.9
% of Time Visible	Third	99.6%	88.2%	74.3%	54.4%
Number of Collisions		0	0.86	1.2	3.8

Table 1 summarizes the results of the 12 simulation sets. The table compares between the pursuer's performance in the three scenarios and shows how the

**Table 2.** Visibility Results

Result	Scenario	10 dyn. obstacles	20 dyn. obstacles	30 dyn. obstacles
A	First	21	21	20
B		4	4	7
C		5	5	3
A	Second	29	21	20
B		1	8	7
C		0	1	3
A	Third	20	11	1
B		4	5	7
C		6	14	22

pursuer's performance changes when dynamic obstacles are added to the static scenarios. As mentioned before, the simulation was run once for the static scenarios and 30 times for each of the dynamic scenarios. For each run, the table shows the number of collisions between the pursuer and the dynamic obstacles on average in the 30 runs. The table also shows the percent of time, on average, for which the evader remained visible to the pursuer in the 30 runs.

Table 2 gives another view of the results by emphasizing on the tracking ability of the pursuer. The table shows three categorizations: A (evader is visible  $\geq 90\%$  of the time), B (evader is visible  $\geq 80\%$  and  $< 90\%$  of the time) and C (evader is visible  $< 80\%$ ). The table shows the number of simulation runs that fall within each of these categories. Table 3 also shows a similar categorization for the results by considering the pursuer's collision avoidance ability. In this table, *A* stands for less than or equal to 3 collisions, *B* for collisions between 3 and 6 and *C* for more than 6 collisions.

But why can't the pursuer avoid all collisions despite of having an adaptive collision avoidance advisor? In many simulation runs, the pursuer succeeds to complete the tracking task without any collision. However, in other runs, it may collide with some dynamic obstacles. This is attributed to either of the following reasons. The first one is the fact that any dynamic obstacle moves in a straight line until it collide with a static obstacle or environment boundary, then it will bounce in a new random direction. This abrupt motion may lead to a collision if the pursuer is close by. Second, the pursuer may be trapped among several dynamic obstacles to the extent that collision becomes inevitable even if motion prediction is correct. This is because dynamic obstacles are assumed to be non-cooperative and, therefore, such obstacles may head to the pursuer. The third reason is due to the degree of trade-off between safety and tracking. This is inherent in the definition of the overall risk function which assigns weights for the collision and occlusion risks. In reality, sometimes it is acceptable to allow collisions during tracking. This depends on the type of pursuer used (fragile, flexible, solid, etc.), type of dynamic obstacles present (soft, hard, etc.), and type of application in hand. It is fair to compare our robot pursuer with a human pursuer. A human pursuer tracking in a cluttered environment like a show room

**Table 3.** Collision Results

Result	Scenario	10 dyn. obstacles	20 dyn. obstacles	30 dyn. obstacles
A	First	29	22	12
B		1	6	9
C		0	2	9
A	Second	27	26	19
B		3	6	9
C		0	1	6
A	Third	29	27	19
B		0	2	5
C		1	1	6

is also susceptible to collisions with objects/humans in the environment. The criticality of these collisions relies on the situation and the type of people/objects with which collisions occur. In fact, it is almost impossible to completely avoid collisions and, therefore, the goal should always be to minimize these collisions as much as possible.

## 6 Conclusions

In this paper, we have shown how occlusion risk can be computed. Occlusion risk represents the ability of the evader to escape at a certain instance of time. To compute occlusion risk, we need to know which vertices can cause occlusion. Next, we compute the occlusion risk for each of these vertices individually and then aggregate all of these risks into one final risk value. Finding occlusion risk at some time instance allows the decision maker component in our solution architecture to evaluate the criticality of the tracking situation.

In this paper, a novel approach for the problem of tracking a moving target in a dynamic environment with visibility-range constraints is presented. The solution has three interacting components: tracking, collision avoidance and motion selection. The occlusion advisor which computes the occlusion risk that represents the ability of the evader to escape at a certain instance of time. To compute occlusion risk, we first determine vertices that cause occlusion before computing its occlusion risk and then aggregate all of these risks into one final risk value. The collision advisor produces an adaptive collision risk value which guides the pursuer to avoid colliding with the moving obstacles present in the environment. To compute this risk factor,  $\mathcal{R}^{coll}$ , visible dynamic obstacles, to the pursuer, along with a short history of their random motion track are made available. This risk represents the likelihood of the pursuer colliding with any of the moving obstacles. Based on this risk and the occlusion risk, the decision maker component evaluates the criticality of the tracking situation. The proposed solution is validated using a comprehensive set of simulations, which show that transition from tracking in static environments to tracking in dynamic environments can be done without much loss in robot safety or tracking ability.

## References

1. Bandyopadhyay, T., Li, Y., Ang Jr., M., Hsu, D.: A Greedy Strategy for Tracking a Locally Predictable Target among Obstacles. In: Proceedings of the IEEE International Conference on Robotics and Automation, pp. 2342–2347 (2006)
2. Bhattacharya, S., Candido, S., Hutchinson, S.: Motion Strategies for Surveillance. In: Proceedings of Robotics: Science and Systems III, Atlanta, GA, USA, pp. 249–256. MIT Press, Cambridge (June 2007)
3. Bhattacharya, S., Hutchinson, S.: Approximation Schemes for Two-Player Pursuit Evasion Games with Visibility Constraints. In: Proceedings of Robotics: Science and Systems IV, Zurich, Switzerland. MIT Press, Cambridge (June 2008)
4. The IFR Statistical Department. 2007: 6,5 million robots in operation world-wide (October 2008). <http://www.worldrobotics.org> (Retrieved February 2009)
5. Elnagar, A., AlBluwi, I.: Maintaining Visibility of a Moving Target: The Case of An Adaptive Collision Risk Function. To appear in the Proceedings of the International Conference on Soft Computing, Intelligent System and Information gTechnology, ICSIIT (2010)
6. Elnagar, A., Gupta, K.: Motion prediction of moving objects based on autoregressive model. *IEEE Transactions on Systems, Man and Cybernetics, Part A* 28(6), 803–810 (1998)
7. Elnagar, A., Hussein, A.: An adaptive motion prediction model for trajectory planner systems. In: Proceedings of the IEEE International Conference on Robotics and Automation, vol. 2, pp. 2442–2447 (2003)
8. Fabiani, P., Gonzalez-Banos, H., Latombe, J., Lin, D.: Tracking a partially predictable target with uncertainties and visibility constraints. *Journal of Robotics and Autonomous Systems* 38(1), 31–48 (2002)
9. Ferguson, D., Kalra, N., Stentz, A.: Replanning with RRTs. In: Proceedings of the IEEE International Conference on Robotics and Automation, pp. 1243–1248 (2006)
10. Gonzalez-Banos, H., Lee, C., Latombe, J.: Real-time combinatorial tracking of a target moving unpredictably among obstacles. In: Proceedings of the IEEE International Conference on Robotics and Automation, vol. 2, pp. 1683–1690 (2002)
11. Hsu, D., Kindel, R., Latombe, J., Rock, S.: Randomized Kinodynamic Motion Planning with Moving Obstacles. *The International Journal of Robotics Research* 21(3), 233–255 (2002)
12. Jaillet, L., Simeon, T.: A prm-based motion planner for dynamically changing environments. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 2, pp. 1606–1611 (2004)
13. Kalman, R.: A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82(1), 35–45 (1960)
14. Kant, K., Zucker, S.: Toward Efficient Trajectory Planning: The Path-Velocity Decomposition. *The International Journal of Robotics Research* 5(3), 72–89 (1986)
15. Koenig, S., Likhachev, M.: D\* lite. In: Proceedings of the Eighteenth national conference on Artificial intelligence, pp. 476–483. American Association for Artificial Intelligence (2002)
16. LaValle, S., Gonzalez-Banos, H., Becker, C., Latombe, J.: Motion strategies for maintaining visibility of a moving target. In: Proceedings of IEEE International Conference on Robotics and Automation, vol. 1, pp. 731–736 (1997)
17. Lee, C.: Real-time Target Tracking in an Indoor Environment. PhD thesis, Stanford, CA, USA, Advisor-J.C. Latombe (2002)

18. Montemerlo, M., Thrun, S., Whittaker, W.: Conditional particle filters for simultaneous mobile robot localization and people-tracking. In: Proceedings of the IEEE International Conference on Robotics and Automation, vol. 1, pp. 695–701 (2002)
19. Murrieta, R., Sarmiento, A., Bhattacharya, S., Hutchinson, S.: Maintaining visibility of a moving target at a fixed distance: the case of observer bounded speed. In: Proceedings of the IEEE International Conference on Robotics and Automation, vol. 1, pp. 479–484 (2004)
20. Murrieta-Cid, R., Monroy, R., Hutchinson, S., Laumond, J.: A Complexity result for the pursuit-evasion game of maintaining visibility of a moving evader. In: Proceedings of IEEE International Conference on Robotics and Automation, pp. 2657–2664 (2008)
21. van den Berg, J., Ferguson, D., Kuffner, J.: Anytime path planning and replanning in dynamic environments. In: Proceedings of the IEEE international Conference on Robotics and Automation, pp. 2366–2371 (2006)
22. Vasquez, D., Fraichard, T.: Motion prediction for moving objects: a statistical approach. In: Proceedings of the IEEE international Conference on Robotics and Automation, vol. 4, pp. 3931–3936 (2004)

# Wall Following with a Single Ultrasonic Sensor

Shérine M. Antoun and Phillip J. McKerrow

**Abstract.** We seek to get better insights into how blind people navigate with the K-Sonar mobility aid and to translate this insight into an autonomous navigation strategy useful to mobile robotics. Current mobile robots that use ultrasonic sensing to follow walls use multiple sensors and use only range to the nearest reflecting point making assumptions about the reflected echo. In this paper, we describe an advanced wall following algorithm, where the robot follows a wall using a single directed Continuous Transmission Frequency Modulated ‘CTFM’ Ultrasonic Sensor. The sensor is mechanically panned to track the wall, avoid collisions, and check for navigable space. We present results from our mobile robot wall following experiments. These experiments allowed us to compare our robot sensing and navigation system to the approach we have observed blind people use.

**Index Terms:** CTFM Directed sensing, Collision avoidance, Discontinuous sensing, Navigation by blind people, Ultrasonic Sensing, Mimicking, Wall following.

## 1 Single Sensor Navigation

When a blind person travels along a corridor, with a single CTFM ultrasonic sensor (K-Sonar) for guidance, he uses the sensor data for multiple purposes: wall following, hole avoidance, collision avoidance, motion tracking, people following and object recognition. He obtains the information for each purpose by pointing the sensor in different directions at different times with a panning motion synchronised with his gait. As a result of the panning the data required for wall following (for example) is updated every couple of seconds. One of the goals of the research reported in this paper is to gain insight into this method of navigation by developing a mobile robot system modelled on it. This insight is helpful both in design of mobile robots and in teaching humans to navigate with the K-Sonar [1].

Wall, corridor, and path following are essential behaviours for robots in many workspaces, requiring viable obstacle avoidance techniques. In many robots ultrasonic sensors are used for wall and corridor following. On these robots, the range from robot to wall is commonly measured by one or more ultrasonic sensor(s) and used for steering control. We note however, that for vehicle steering, the bearing to the wall is often more useful than distance to the wall when it comes to wall following [2]. Wall following robots that use a real-time obstacle avoidance technique rely on multiple sensors and advanced algorithms such as the Virtual Force Field (VFF) method [3]. Other robots use techniques that require reducing sensor data to point features so that corrections can be calculated with Kalman filtering. Robots that

perceive the environment from multiple sensors such as in a sonar ring [4] and those that rely on complex genetic algorithms have to contend with the need to fuse the different inputs of range and directionality from the various sensors into a set of usable data [4]. Others construct features called Regions of Constant Depth (RCD) semicircles whose radii are the ranges to the nearest perpendicular surface and use them to track a wall [5].

We are able to obtain high quality range data by controlling the direction/orientation of a single ultrasonic sensor and consequently controlling the region of ensonification where the sound wave is reflected from. As a result we avoid the need to use exact geometric maps of the environment, and the problems of data fusion from multiple sonar sensors [4]. At the same time we mimic a human navigation strategy that relies mostly on approximation and time to collide estimates. Our strategy differs from force-based obstacle-avoidance methods [6] that do not allow the robot to pass through narrow passages. Force based obstacle avoiders can suffer from instability of motion in confined spaces [6].

For a robot to successfully navigate with only one sensor a key question is: Is it possible to move safely in one direction as it pans the sensor to a different one? We rephrase this question with respect to the task requirements of a wall following robot in mind, namely that: it has to perceive sufficient information about its environment to determine where there is a safe path to travel, and a continuous wall to follow. This is not unlike the strategy employed by blind people; they scan (long cane or ultrasonic mobility aid) a short distance ahead to verify clear safe footing, and regularly pan to achieve a single contact to the wall being followed to ensure tracking.

We aim to demonstrate that for both tasks a robot does not need to contiguously perceive the wall its tracking and the floor its travelling on, rather it needs to confirm the presence of both at regular intervals. We presented examples of collision avoidance with directed sensing in a previous papers [7][1], and we have derived a navigation model of a blind person travelling using a K-Sonar mobility aid and developed a theoretical set of algorithms for safe translation, path planning and landmark navigation using ultrasonic sensing [7][1].

The significance of this approach is that it is based on observations that both sighted and vision impaired people navigate by reducing the sensed information to the minimum required for the task. This enables them to safely navigate based on a constant update of environment data for the region immediately ahead of them [8][9].

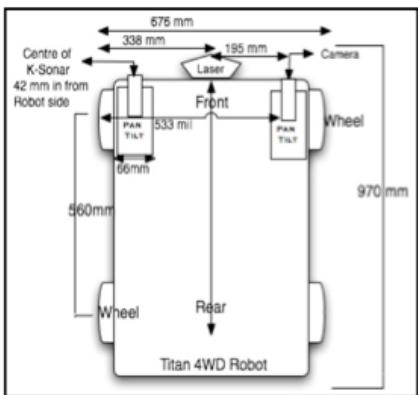
Individuals do not walk looking down at their feet; rather they scan ahead looking over a 120° field of view. Blind people do it with a long cane or with a K-Sonar mobility aid to scan the volume of space in front of them and the ground over a ± 30° field. Humans often use an oval panning pattern between 4 waypoints of front, left, floor, right and repeat [7]. In this work, we simulate the strategies that blind people rely on, and demonstrate that the robot indeed can travel safely using a single directed Ultrasonic Sensor. We deliberately choose to use a single ultrasonic sensor because we observe that blind people do not use laser or vision systems, but rather they rely on touch and/or auditory perception. In our research we used lasers and vision to understand the sources of the echoes perceived by the ultrasonic sensor as it panned to the four waypoints during its wall following mission.

A wall-following robot mission is a tasking of a mobile robot to start a journey along side a wall, to navigate following the contours of that wall at some distance without straying too close, or too far from the wall, and to stop at the end of the wall. For this task our robot had no need to map its environment, rather it relies on surface perception of its workspace combined with a sensing model.

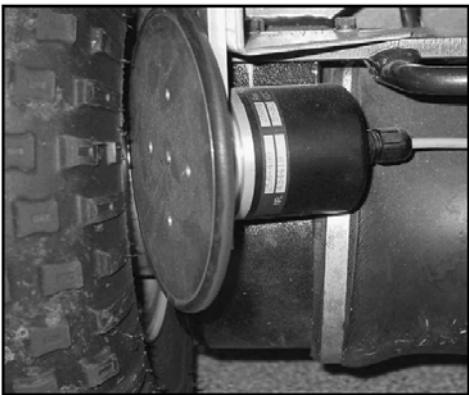
## 2 Titan 4WD Mobile Robot

Titan is a reengineered 4-wheel drive wheelchair. We replaced the chair with a superstructure to hold electronic racks and sensors (Fig. 1). The wheeled base comprises the drive system and batteries. The drive system consists of four motorised wheels for traction and steering. All wheels are fitted with low-pressure pneumatic tyres that act as suspension, compensating for irregularities in the terrain, thus ensuring optimal traction. Two encoders mounted on the rear wheels, measure robot position with a resolution of 0.49 mm (Fig. 2). A third encoder measures the steering angle of the left front wheel. Steering is achieved with differential velocity between the two sides, without skid steering [10]. The front wheels are free to turn but constrained to achieve Ackerman steering by a 4 bar linkage. As a result, differential velocity indirectly causes the front wheels to steer. This design makes the steering susceptible to being back-driven by bumps. We reduce this problem with a steering control loop using feedback from the steering encoder. The sensor suite also includes a K-Sonar' CTFM ultrasonic sensor [11]. The focus of our work is: wall following with a single ultrasonic sensor. To that end, data from the remainder of the sensor suite is not gathered or utilised by our software.

Titan's sensors and actuators are connected to a Mac Mini by: USB interface devices (Fig. 3). The software is written as a set of applications in LabVIEW. The



**Fig. 1.** Titan 4WD robot geometry, sensor mount point can be swapped to suit different tasks



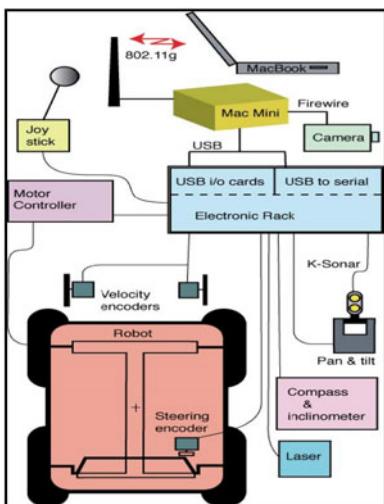
**Fig. 2.** Motion Encoder mounted on rear wheel driven by a friction coupling

applications use functions from four libraries: input/output library, control library, ultrasonic sensing library, and state library.

### 3 K-Sonar

In this research, we are using the K-sonar CTFM sensor developed by BAT [11] as a mobility aid for blind people (Fig. 4). The K-Sonar combines two transducers to form a transmitter and receiver. At 75 kHz a single 19 mm diameter transducer has a theoretical beam angle of  $\beta = 19.32^\circ$  from axis to first minima (Fig. 5), the vertical diameter is 47 mm and the theoretical horizontal beam angle is  $7.6^\circ$ . The CTFM system (Fig. 6) is set to transmit a downward swept sine wave (fsweep is 100 kHz to 50 kHz) every 100 MSc (sweep period  $t_s$ ) (Fig. 6). The ultrasound energy reflects from objects and returns to the receiver as an echo. The echo is a delayed and filtered version of the transmitted signal. A demodulation sweep, derived from the transmitted sweep, is multiplied with the received echo in the time domain. The outputs of this multiplication are sum and difference frequencies (Fig. 6)[12]. The distance of flight information is contained in the difference frequencies ( $f_a$  is 0 to 5 kHz), where frequency is proportional to range (Fig. 6, 7) and amplitude is proportional to orthogonal surface area. This time domain signal is converted to a power spectrum with a Fast Fourier Transform (FFT) to give a range-energy echo (Fig. 7).

The amplitude in frequency-bin  $i$  is the energy reflected from surfaces in a spherical annulus at range  $r_i$  (Fig. 5). To characterise the sensor we measured the echo from a large flat surface, at 1 metre range, as we panned the sensor by  $\pm 40^\circ$ .

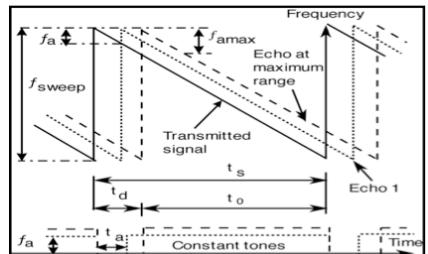
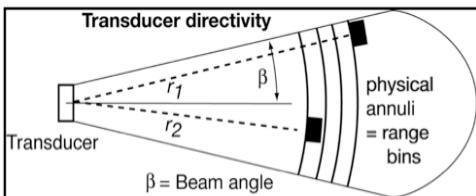


**Fig. 3.** Block diagram of Titan's sensors, actuators, and electronics



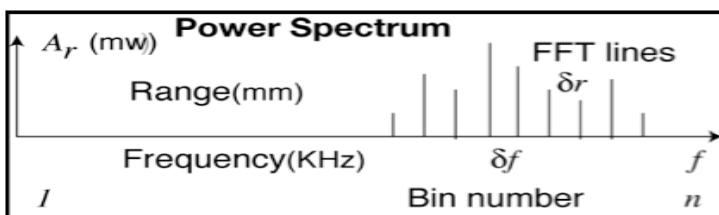
**Fig. 4.** K-Sonar ultrasonic sensor is designed with a mount point to fit on a blind person's long cane

The measurements for a pan of the surface (Fig. 8 at end of paper) show that the horizontal beam angle of the main lobe is  $+18^\circ$  to  $-20^\circ$ , which is close to the theoretical  $19.32^\circ$ . Similarly the measurement for the vertical beam angle is  $\pm 15^\circ$ . The measurements for a pan of the surface (Fig. 8 at end of paper) show that the horizontal beam angle of the main lobe is  $+18^\circ$  to  $-20^\circ$ , which is close to the theoretical  $19.32^\circ$ . Similarly the measurement for the vertical beam angle is  $\pm 15^\circ$ .



**Fig. 5.** An ultrasonic transducer emits a beam of energy.  $r$  = range to orthogonal surface One transducer is used for transmission and one for reception.

**Fig. 6.** CTFM demodulation – multiplying the echo by the transmitted signal produces a set of different tones where frequency is proportional to range to object



**Fig. 7.** Power spectrum of echo – frequency (bin number) is proportional to range and amplitude to echo energy at that range

## 4 Directed Sensing

The model of directed sensing we use is based on observations of the way blind humans pan the K-Sonar to achieve directed sensing. The workspace is sensed discontinuously at regular intervals so the sensing pattern is as follows.

1. Front, looking for obstacles, overhangs, and safe traversable space.
2. Left, to detect the wall being tracked. The expected range to the wall at the target tracking distance and sensor orientation is calculated. Tolerances based on beam angle, are added to this range to form a search window (Fig. 9). Then we look in this window for a wall echo. If an echo is recognised then the distance to the wall is calculated using the geometry in Fig. 9.
3. Floor, sense bumps, holes, obstructions, and oncoming traffic.

4. Right, sense intruding obstacles and their intrusion distance (does corridor become too narrow).

The model of directed sensing we are using consists of a pan cycle of four Pan-Tilt (PT) movements to waypoints executed as the robot moves. The cycle is repeated at a rate of one pan cycle every 2.4 seconds in a counter clock wise direction. From an initial position of ahead (PT 0°;0°) the unit moves to left (PT-40°;0°), floor (PT 0°;40°, right (PT 40°;0°), back to ahead (PT 0°;0°). During this pan cycle, sonar echo data are captured every 100 msec. So for example, during the continuous motion between waypoints ahead and left echoes are captured at 3°, 10°, 20°, 30°, 37°, and 40°.

We observed that blind people scan only in the direction in which they are travelling relying on overtaking people to avoid colliding with them. Noting the utility of this approach we adopted similar techniques. Titan scans the space ahead in the direction of travel along the wall it is tracking. In Australia we drive on the left, and we are encouraged to keep left when walking or climbing stairs, so we investigate left wall tracking. Our work can readily be applied to right wall tracking with a simple mount point shift for the CTFM sonar and minor software changes to the state machine controller(Fig. 10).

In this paper we consider only echoes reflected off a flat wall to the left of Titan they are used to control the wall following mission. The four waypoints denote one each of four main states in a state machine software architecture [7] used to control Titan's navigation (Fig. 10). In each state, 6 echoes are captured, one every 100 msec, then a command is sent to move the sonar to next waypoint and the state machine to the next state (Fig. 10).

The look ahead state considers the space ahead of the robot in the corridor. It adjusts the linear velocity based on the clear distance ahead, decelerating if an object is in the path of the robot at range 1.8 m or less and stopping at range  $\leq 300$  mm [7].

The look left state estimates the distance to the wall and together with the distance estimate from the previous cycle calculates an orientation angle to the wall. Then it issues a steering angle command to adjust Titan's trajectory. It is calculated to steer the robot to reduce the error in its orientation to, and range from, the wall.

In addition when an echo between the front, left and floor waypoints indicates that the robot is closer to the wall than 80% of the target tracking distance from the wall, a left touch command is triggered to modify the steering reference by up to 1° to the right. This allows the robot to detect and start correcting for variations in the wall before the robot reaches them.

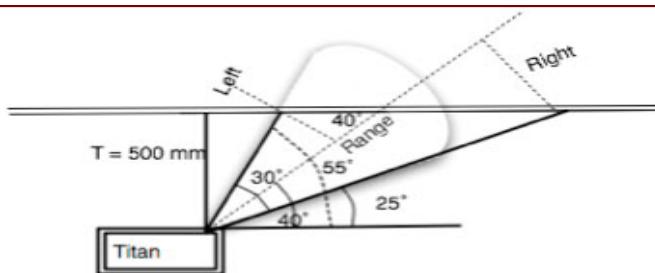
The look floor state considers the surface on which the robot is travelling to check that the floor is safe to traverse i.e. it looks for steps up, down, and or other low lying objects and it behaves in a similar manner to the look ahead state.

The look right state checks for intruding objects and adjusts the tracking target to ensure that there is sufficient clearance on the right of the robot for the robot to pass the object. Often objects to the right are moving in the opposite direction to the robot, so to detect them well ahead echoes between floor, right, and front waypoints and check each for clearance too. If too close a right touch is activated and functions similarly to left touch.

Every scan cycle, we collect a data vector of the 80 features that the controller's logic used to issue the steering commands for the mission. With this data we can replay the operation of the robot with a simulator and analyse our navigation experiment. Typically, we consider the geometry of the environment, which state the controller is in, the ultrasonic sonar data and the consequent decisions by the controller. The distance to the wall is calculated from the sensor geometry and the echo. Examples of experiment reconstruction for the experiments in Fig. 12 are shown in Fig. 11, 13 (end of paper) to illustrate the results of navigation experiments.

## 5 Wall Tracking with Directed Sensing

The challenge with discontinuous sensing is to achieve stable control. As the time between sensing gets longer this challenge becomes more difficult. The wall-tracking controller calculates a new steering angle reference each time period. Two parallel controllers are used (Equation 1). One controls the angle to the wall with a simple proportional gain; the other controls the range to the wall with a proportional/Integral controller. It required many runs of the robot to achieve the tuning for the results for the 7 metre runs shown in Fig. 11 and 13. Our initial settings of the control loop gains ( $g_r$ ,  $g_i$ , and  $g_a$ ) were too high causing the robot to over-steer to the point where it drove the robot into the wall.



The conical shape of sonar signal with the beam angle of  $\pm 30^\circ$

Our target range to the wall is  $T = 500$  mm

The range to the wall at  $55^\circ$  = Left

The range to the wall at  $25^\circ$  = Right

Therefore we window the echo at ranges Left to Right

-----  
|-----|-----|  
Left      Right

We use threshold to reduce the window

-----|---|---|=|---|-----  
Left      ^      Right  
Signal

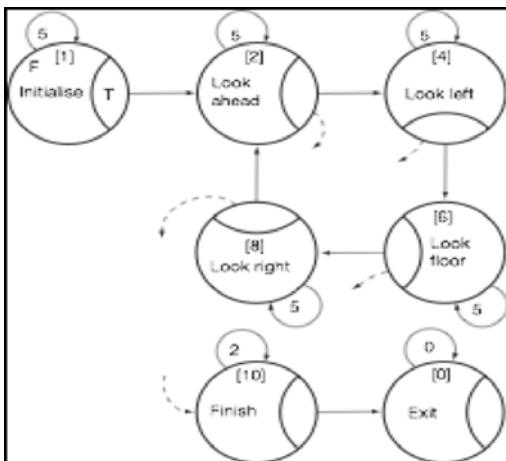
-find peak = range to wall

Also we can add a tolerance to beam angle to cater for the robot being not parallel to the wall and not at Target from the wall (Pythagorus Theorem gives us Target range error)

**Fig. 9.** Geometry for calculating the window in which to scan for wall echo

Conversely, when the gains are too low the robot doesn't correct quickly enough and runs into the wall. Once we achieved stable steering control we ran a suite of wall following experiments to test for the effects of various initial start positions and wall geometries (Figs. 11, 12, 13), in all cases successfully following the wall for the length of the mission. The data plotted on Fig 11 and 13 are: range to left wall (brown vertical bars), steering angle reference (red solid line), steering angle (blue dotted line) linear velocity (green dashed line), magnetic bearing (black dots dashes line), the states plotted with dots are: found left wall (purple at  $y = 11$ ), left touch (red at  $y = 8$ ) right touch (red at  $y = -6$ ).

The wall-tracking controller calculates a new steering angle reference each time period. Two parallel controllers are used (Equation 1). One controls the angle to the wall with a simple proportional gain; the other controls the range to the wall with a proportional/Integral controller..



$$S_t = g_r e_t + g_i \int e_t - g_a (r_t - r_{t-1})$$

where  $S_t$  = steering angle command  
 $e_t = ref_t - r_t$  = range error  
 $g_r$  = range proportional gain  
 $g_i$  = range integral gain  
 $g_a$  = angle proportional gain  
 $I_t = g_i \int e_t - g_a e_t + I_{t-1}$  = integral gain

**Equation 1.** Steering control

**Fig. 10.** Time based state machine for path scanning. [Value] is the number of 100 msec time steps in state.

## 6 Experiment Results

In the experiment in Fig 11, to follow a straight wall (Fig 12 A), the robot starts with a large initial error. The vertical bars (brown) show the range to the wall in (mm/100) calculated from the sonar echo. The robot starts from a stationary position, and is at 750 mm from the wall ( $x = 11$ ) with a target tracking distance of 500 mm. Thus in its initial position it has a tracking error of -250 mm, so it is commanded to steer (solid line) towards the wall with an angle of -8.5° ( $x = 35$ ). Note it also has an initial steering angle of -1.70° and is oriented 8.9° away from the wall (measured manually). As a result of the initial errors and its initial motion, at the next measurement its distance to the wall ( $x = 59$ ) is 420 mm, so it corrects with a command of +9° to turn away from the wall.

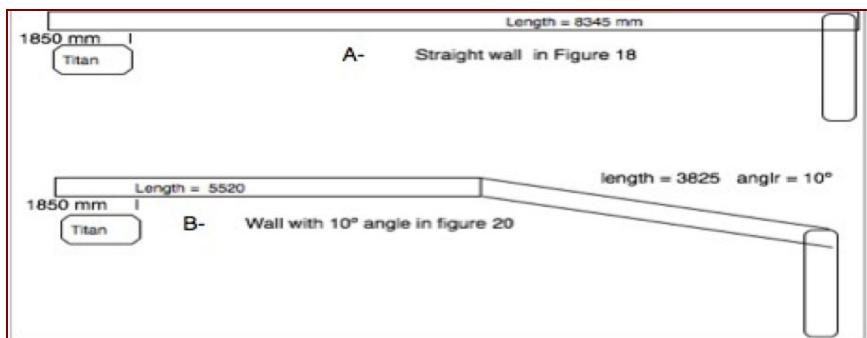
As Titan is a relatively large robot, it requires time to respond to commands. The dashed line (green at  $y = 4$ ) shows that the robot takes 1.5 seconds to accelerate to 400 mm/sec. The dotted blue line shows that it takes half a second to change steering angle when moving at 400 mm/sec. The plot of magnetic bearing (black dash dot starting at  $y = 9$ ) shows that changes in orientation of the robot occur quickly. Notice, everything slows down when the robot slows down in response to detecting an object ahead ( $x = 178$ ).

The initial errors combined with the effects of accelerating to target velocity resulted in both orientation and range errors. Once the robot reaches target velocity, it can then start to follow the wall. From  $x = 83$  to  $x = 178$  the commanded steering angle converges, as the robot tracks the wall. At  $x = 180$ , the robot is oriented slightly towards the wall, detects the wall ahead, and slows down to 250 mm/sec and continues to track the wall until it stops at a barrier placed at the end of the wall for this purpose (at  $x = 260$ ).

Along with controlling the steering, the robot reports a number of states that can be used in higher-level navigation decisions, and for evaluating performance. Three of them are shown as dots on the graph. The circular dots at  $y = 11$  (purple) indicate that the sensor has found the wall each time (every 2.4 sec) it looked for it, so the range readings (brown vertical bars) are considered valid. The square (red) dots at  $y = 8$  indicate that on those scans the robot found that it is closer than 80% of the tracking distance from the left wall. As these scans have a lower incident angle to the wall they in fact provide a look ahead for changes in the wall geometry, and in response change the steering angle by up to 1°.

The square (red) dots at  $y = -6$  and  $x = 240$  indicate detection of a right surface ahead and change the steering reference. In this run, the robot is detecting the barrier placed at the end of the wall to block its path to stop it. Hence, the detection of a right surface is an erroneous state. Also the robots ability to steer declines at low velocities. These two problems are counteracted by only allowing steering command and control when the robot velocity is greater than 50 mm/sec.

A second experiment involved starting the robot approximately parallel to the wall and then when it should be running at speed, introducing a bend in the wall. The results in Fig. 13 (at end of paper) are for a 10° bend (Fig 12 B). By  $x = 83$  the robot is tracking the wall.



**Fig. 12.** Walls Titan followed in the experiment in Fig 13 & 15

At  $x = 131$ , it detects the wall at the other side of the bend. So from  $x = 131$  to  $x = 230$  it is commanded to turn away from the wall by both left wall steering reference calculation and left wall look ahead. At  $x = 227$  it has passed the bend but is too far from the wall so it is commanded to turn back towards the wall, before being commanded to stop at  $x = 248$ .

## 7 Discussion of Results

The two experiments demonstrate that the robot is able to track a wall, when travelling at 400 mm/sec, and sensing the wall every 2.4 sec (960 mm). The control oscillates due to the time delay between sensing, but it converges showing that it is stable. Hence it is evidence that regular intermittent sensing can be used to navigate along a wall.

The reliability of this control is improved by looking ahead on non-waypoint scans for variations in the left wall and intruding objects on the right. This models the ability of the blind human to perceive such objects between panning waypoints.

The robot stopped as commanded at the end of each run when it detected an intersecting wall (object in the way). Also the collision avoidance software slowed the robot by detecting obstacles on other scans during the pan cycle.

In the first experiment the robot was positioned with large initial errors in both range and orientation. Once up to speed it corrected for those errors in 6 sensing cycles and stopped at the obstacle. By contrast, in the second experiment, the initial errors were small, and the robot corrected for those errors in 2 sensing cycles. Its navigation of a bend with a 10° change in orientation shows that the system is capable of following walls that are not straight, and handling step changes in wall orientation.

## 8 Conclusions

Part of our investigation is to determine the safety of our directed sensing for navigation strategy. As stated at the outset we have observed vision impaired people using this strategy safely. Here we have presented initial experimental results in navigation with directed sensing that demonstrates the validity of our contention: namely that it is possible for a robot to move safely in one direction as it scans in a different one. The key is to break the processes into tasks and to scan for the information needed for each task at the point in time that it is required.

The requirements of the sensing system for a wall following robot are to perceive sufficient information about the environment to determine where the wall is, and the robot's orientation to the wall. At each sensor scan during the pan cycle, the human or the robot has to execute functions pertinent to the direction the sensor is pointing such as: collision avoidance if pointing ahead, and steering if pointing left. When our robot is moving at speeds below 50mm/sec steering commands are ineffective so they are not issued.

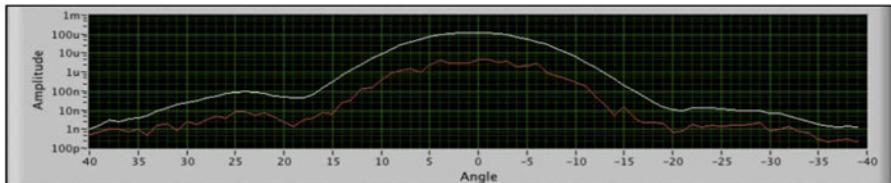
In this research we have demonstrated that the strategies employed by humans (both sighted and blind) in look/scan ahead, and collision avoidance can be used successfully by a robot to follow a wall with no need for precise prior or geometric

mapping of the workspace. Also, we achieved this with a single directed sensor. Our appreciation of the human's ability to both perceive the wall and to navigate along it with limited sensing has increased significantly as a result of our experience in trying to program a mobile robot to achieve the same task.

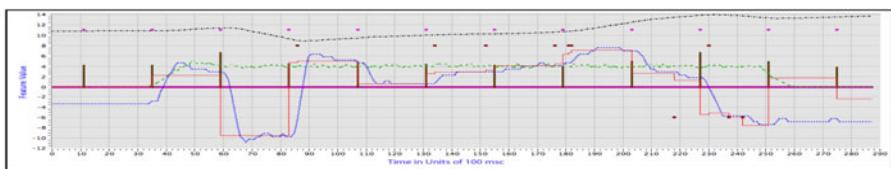
## References

- [1] Vita, F., McKerrow, P.: DVD Use of echo location visual aid devices educational video shown at a Sight Village Exhibition in Birmingham, U.K. (July 15-17, 2008)
- [2] Yata, T., Kleeman, L., Yuta, S.: Wall Following Using Angle Information Measured by a Single Ultrasonic Transducer. In: Proceedings of the 1998 IEEE International Conference on Robotics & Automation Leuven, Belgium (May 1998)
- [3] Borenstein, J., Koren, Y.: Real-time Obstacle Avoidance for Fast Mobile Robots. *IEEE Transactions on Systems, Man, and Cybernetics* 19(5), 1179–1187 (1989)
- [4] Braunstingl, R., Mujika, J., Uribe, J.P.: A Wall Following Robot With A Fuzzy Logic Controller Optimized By A Genetic Algorithm. In: Proceedings of 1995 International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium (1995)
- [5] Maksarov, D., Durrant-Whyte, H.: Mobile vehicle navigation in unknown environments: a multiple hypothesis approach. In: IEE Proc.-Control Theory Appl., July 1995, vol. 142(4) (1995)
- [6] Arkin, R.C.: Motor Schema-Based Mobile Robot Navigation. *The International Journal of Robotics Research*, 92–112 (August 1989)
- [7] McKerrow, P.J., Antoun, S.M., Worth, P.: A software architecture for mobile robot navigation. In: TAROS 2008, Edinburgh, September 2008, pp. 2–7 (2008)
- [8] Berthoz, A.: *The Brain's Sense of Movement*, pp. 81–83. Harvard Univ. Press, Cambridge (2000)
- [9] Berthoz, A.: Simplifying principles of perception, action, locomotion and navigation. In: Plenary Talk, Digest IEEE ICRA 2007, Rome, pp. XVIII - XIX (April 2007)
- [10] McKerrow, P.J., Ratner, D.: Calibrating mobile robot odometry with ultrasonic sensors, Bay Advanced Technology, <http://www.batforblind.com> (Last visited December 2008)
- [11] Bay Advanced Technology 'BAT', <http://www.batforblind.com> (Last visited December 2008)
- [12] Ratner, D., McKerrow, P.J.: Navigating an outdoor robot along continuous landmarks with ultrasonic sensing. In: *Robotics and Autonomous Systems*, vol. 45, pp. 73–82 (2003)
- [13] Antoun, S.M., McKerrow, P.J.: Perceiving A Corridor With CTFM Ultrasonic Sensing. In: Proceedings of Australasian Conference on Robotics and Automation, Brisbane (2007); Proceedings of IROS 2002, Lausanne, pp. 859–864 (October 2002)

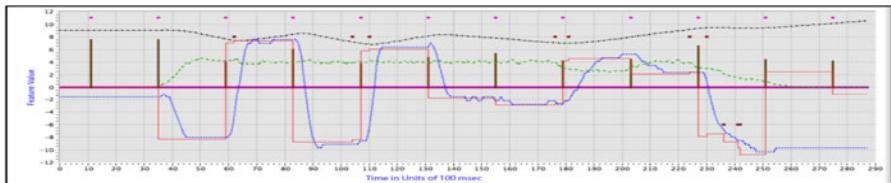
## Large Figures



**Fig. 8.** Measurement of the horizontal beam angle of the K-Sonar using reflection from a flat surface at 1 m range



**Fig. 11.** Data recorded during a 7 m run along a straight wall, range target 500 mm. Key - Continuous values: Steering angle reference - red solid line (scale = 1), Steering angle - blue dotted line (scale = 1), Magnetic bearing - black dots and dashes (scale = 0.13), Linear Velocity - green dashed line (scale = 10), Range to wall - vertical brown bars (scale = 0.01). States: Found wall - purple dots at  $y = 11$ , left touch - red dots at  $y = 8$ , and right touch - red dots at  $y = -6$ .



**Fig. 13.** Data recorded during a 7 m run along a wall with a 10° bend near midway point ( $x = 120$ ), range target 500 mm. Key - Continuous values: Steering angle reference - red solid line (scale = 1), Steering angle - blue dotted line (scale = 1), Magnetic bearing - black dots and dashes (scale = 0.13), Linear Velocity - green dashed line (scale = 10), Range to wall - vertical brown bars (scale = 0.01). States: Found wall - purple dots at  $y = 11$ , left touch - red dots at  $y = 8$ , and right touch - red dots at  $y = -6$ .

# Kinematics Modeling and Analysis of a Five-Wheeled Rover with Caster and Camber

He Xu\* and Zhenyu Zhang

College of Mechanic and Electric Engineering, Harbin Engineering University,  
150001, Harbin, P.R. China

[railway\\_dragon@163.com](mailto:railway_dragon@163.com), [zhang.phoenix@gmail.com](mailto:zhang.phoenix@gmail.com)

**Abstract.** Because of articulated all terrain rovers (ATR) have adaptivity of rough terrain, they are playing increasingly important role in many areas. Based on a castered-and-cambered 5-wheel-steering and 5-wheel-driving (5WS-5WD) ATR prototype, its kinematics model is derived for full 6 Degree of Freedom (DOF) motion, in this process, the influence of wheel-terrain contact including contact angle and slippage are taken into account. The kinematical impact of caster and camber is elaborated.

**Keywords:** Mobile robot, rover kinematics, caster and camber, fifth wheel.

## 1 Introduction

Articulated all terrain rovers (ATR) are playing increasingly important role in many areas, such as planet exploration and detecting<sup>[1]</sup>, search and rescue, agriculture<sup>[2]</sup>, mining industries<sup>[3]</sup>, and military mission like defense and hazardous material handling. NASA has been developing articulated all terrain rovers for planet exploration, such as, Rocky series, Sojourner, Spirit, Phoenix and newly developed Mars Science Laboratory<sup>[4]</sup>.

There is a tremendous development in mobile robot, the navigation and motion control of the rovers are based on kinematics modeling and analysis. Research of kinematics modeling have been fenced in modeling on flat surface. In recent years, all terrain rovers kinematics in the environment of uneven terrain is surging. In general, there are two methods for kinematics modeling, geometric and transformation. The geometric approach is restrictive if used on its own. Transformation is commonly adopted by researchers. For instance, Muir and Newman accomplished some fundamental work [5], Rajagopalan [6] utilizes a transformation approach to develop the kinematic model of ordinary mobile

---

\* This work was supported by the National Science Foundation of China under Grant 60775060, the Foundation Research Fund of Harbin Engineering University under Grant HEUFT07027, Natural Science Foundation of the Heilongjiang Province of China under Grant F200801, and Specialized Research Fund for the Doctoral Program of Higher Education under Grant 200802171053.

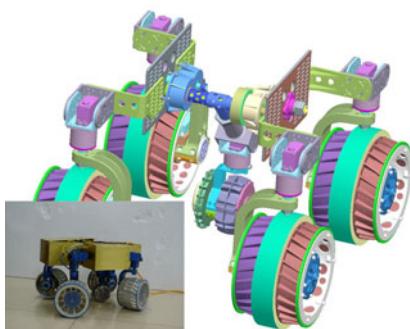
robots (OMRs) with an inclined steering column and different combinations of driving and steering wheels, the forward kinematics of a specific OMR with omnidirectional wheels is derived in B.J.Yi's article [7]. A simple kinematics model and a state observer is adopted for estimate position, orientation, velocities, and contact angles of a rover. Mahmoud Tarokh describes a general approach to kinematics modeling and analysis of articulated rovers traversing uneven terrain [8], moreover, he take wheel-terrain contact into account [9].

However, all the forgoing achievement have not take caster and camber into consideration. Because of caster and camber can enhance the stability and kinematics performance, so, the study of caster and camber may benefit the development of new ATRs. He Xu did some work on kinematics modeling of rover with caster and camber, but wheel-terrain contact is neglected [10].

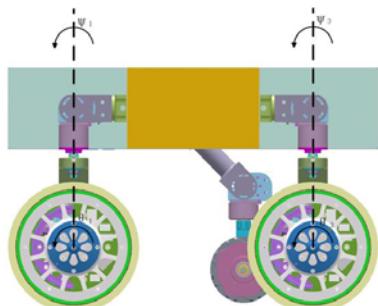
Based on the achievements illustrated previously, the study of castered-and-cambered rover kinematics which considers wheel-terrain contact and the impact of caster and camber to the kinematics is reasonable.

## 2 Structure of a Five Wheel Rover

As shown in Fig. 1, the ATR this paper based on adopts a 5-wheel-driving, 5-wheel-steering mode. The main body connected to wheels via a set of linkages and joints, and it is suspended on a separate differential mechanism. Chassis system is composed of two rockers on left and right side of the rover, left side rocker is shown in Fig. 2. The fifth wheel is articulated with the gear axis of the separate differential mechanism. By rolling of the wheels and rotating its joints, the ATR can traverse uneven terrain. The conjunction between the rocker and wheel-arm is reconfigurable, by adjusting it, caster and camber angle could be set.



**Fig. 1.** A 5WD and 5WS rover



**Fig. 2.** Left side rocker

### 3 Kinematics Model Development

#### 3.1 Assumptions

In the development of kinematics model, we adopt three assumptions: (1)there is no deformation of wheels as well as terrain surface;(2)the wheels of the rover are simplified to wafers;(3)wheels and surface keep point contact.

#### 3.2 Definition of Coordinate Frame and Variables

Fig.3 shows the coordinate frames of left side of the rover and the left rocker. The number of wheels are as: wheel-1 is left front wheel, wheel-3 is the left rear wheel. Wheel-2 and wheel-4 are right front wheel and right rear wheel respectively. The meanings of coordinate frames are:  $L$  denotes terrain frame,  $R$  denotes reference frame of the rover body,  $D$  denotes the middle point frame of differential mechanism,  $T$  denotes the middle point frame of the axis of differential mechanism,  $H_i(i = 1, 2, 3, 4)$  denotes the frame fixed on the ends of left and right rockers according to the wheel number,  $S_i(i = 1, 2, 3, 4, 5)$  denotes steering frame of each wheel,  $A_i(i = 1, 2, 3, 4, 5)$  denotes wheel frame of wheel 1,2,3,4 and 5 respectively;  $R, D, T, S_i$  and  $A_i$  are the frames at time  $t$ .  $\bar{R}, \bar{D}, \bar{T}, \bar{S}_i$  and  $\bar{A}_i$  are the frames at time  $t - \Delta t$ , besides, they are instantaneously coincidence with  $R, D, T, S_i$  and  $A_i$  respectively. The frame parameters are listed below:

$k_1$  - vertical offset between rover reference frame  $R$  to differential frame  $D$ ;

$k_2$  - forward offset between rover reference frame  $R$  to differential frame  $D$ ;

$k_3 = 165\text{mm}$  - Distance between  $D$  and rocker;

$k_4 = 161.5\text{mm}$  - Distance between the middle point of rocker and frame  $H$ ;

$k_5 = 178.1\text{mm}$  - Distance between  $H$  and the center of the wheel;

$k_6$  - forward offset between rover reference frame  $R$  to fifth wheel arm frame  $T$ ;

$r = 88.5\text{mm}$  - Wheel radius. Fig.4 illustrates the camber and caster transform

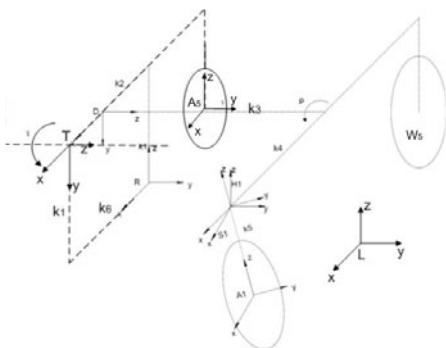


Fig. 3. Left side and 5<sup>th</sup> wheel frame

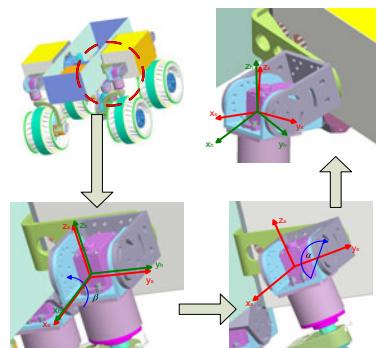


Fig. 4. Transformation from H to S

from the origin pose, the transformation matrix is:

$$T_{H_i, S_i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c\beta_i & -s\beta_i & 0 \\ 0 & s\beta_i & c\beta_i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c\alpha_i & 0 & -s\alpha_i & 0 \\ 0 & 1 & 0 & 0 \\ s\alpha_i & 0 & c\alpha_i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

In which,  $s$  and  $c$  stand for sin and cosine respectively ( $i = 1, 2, 3$  and  $4$ ).

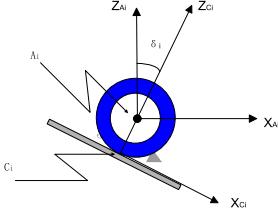


Fig. 5. Contact frame

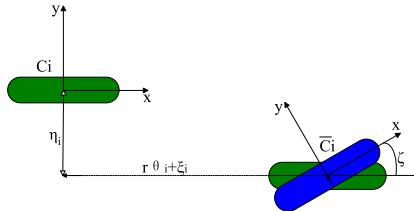


Fig. 6. Incremental motion by rolling and slip

Fig 5 illustrates the coordinate frame  $C_i$  for terrain contact point of wheel- $i$ , where its  $x_{ci}$  axis is tangent to the terrain at the point of contact, and its  $z_{ci}$  axis is normal to the terrain. From axle-coordinate frame  $A_i$ , the contact coordinate frame  $C_i$  can be obtained: rotating  $\delta_i$  about the axle, then translating by the wheel radius  $r$  in the negative  $z$  direction. The transformation matrix from  $A_i$  to  $C_i$  denoted by  $T_{A_i, C_i}$  is:

$$T_{A_i, C_i} = \begin{bmatrix} c\delta_i & 0 & s\delta_i & -rs\delta_i \\ 0 & 1 & 0 & 0 \\ -s\delta_i & 0 & c\delta_i & -rc\delta_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Fig 6 illustrates the slipping model of wheel- $i$ .  $C_i$  ( $i = 1, 2, 3, 4, 5$ ) is the contact frame of each wheel, and  $\bar{C}_i$  ( $i = 1, 2, 3, 4, 5$ ) denotes  $C_i(t - \Delta t)$  ( $i = 1, 2, 3, 4, 5$ ). Coordinate transformation from  $\bar{C}_i$  to  $C_i$  is based on a lateral slip  $\eta_i$  along the  $y$  axis, a turn-slip rotation  $\zeta_i$  about the  $z$  axis and a wheel-rolling translation ( $r\theta_i + \xi_i$ ) along the  $x$  axis,  $\theta_i$  and  $\xi_i$  denote angular rotation and rolling slip respectively. The transformation matrix from  $\bar{C}_i$  to  $C_i$  is [9]:

$$T_{\bar{C}_i, C_i} = \begin{bmatrix} c\zeta_i & -s\zeta_i & 0 & r\theta_i + \xi_i \\ s\zeta_i & c\zeta_i & 0 & \eta_i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

The pose of reference point in terrain frame  $L$  could be illustrated as a vector  $U = [X \ Y \ Z \ \Phi_x \ \Phi_y \ \Phi_z]^T$ , where,  $[X \ Y \ Z]^T$  is the position vector, and

$[\Phi_x \Phi_y \Phi_z]^T$  is the orientation vector. The lower-case quantities are with respect to the instantaneous rover frame  $R$ .

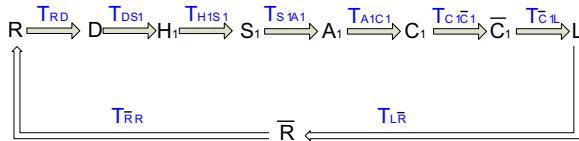
Two rockers are connected to the body via a differential mechanism, The left and right side rocker angles denoted by  $\rho_1$  and  $\rho_2$  respectively, which could be represented by a single variable  $\rho$  and  $\rho = \rho_1 = -\rho_2$ , when the terrain is complanate,  $\rho_1 = \rho_2 = 0$ ,  $\rho$  is measured by an encoder. The fifth wheel rocker angle is  $\tau$ , also measured by an encoder,  $\tau = 0$  if the ground is flat. Because servo motors are used for steering of four wheels, the steer angles  $\psi_i$  are controlled by the width of the width of pulse. The joint angles could be illustrated by a vector  $q = [\rho \tau \psi_1 \psi_2 \psi_3 \psi_4 \psi_5]^T$ . The range of  $\rho$  and  $\psi_i$  is:

$$-90^\circ \leq \rho \leq 90^\circ; -45^\circ \leq \tau \leq 45^\circ; -90^\circ \leq \psi_i \leq 90^\circ.$$

Every wheel is driven independently, and its angular rotation  $\theta_i$  is measured by an encoder.

### 3.3 Coordinate Transformation

Prescribe  $\dot{X}$  is the differential of matrix  $X$ .  $T_{M,N}$  stands for the transforming matrix from coordinate M to N. From the properties of transforming matrix, we can obtain:  $T_{M,N} = (T_{N,M})^{-1}$ ;  $T_{M,M}$  is identity;  $T_{M,\bar{N}} = T_{\bar{M},N} = T_{\bar{M},\bar{N}} = T_{M,N}$ ;  $T_{M,N} = T_{M,X}T_{X,N}$ ;  $\dot{T}_{M,N} = \dot{T}_{M,X}T_{X,N} + T_{M,X}\dot{T}_{X,N}$ .



**Fig. 7.** Sketch map of coordinate transformation for wheel-1

Fig. 7 is the sketch map of coordinate transformation of wheel-1, according to Fig.7, transformation matrix from  $\bar{R}$  to R is:

$$T_{\bar{R},R} = T_{\bar{R},L}T_{L,C1}T_{C1,C1}T_{C1,A1}T_{A1,S1}T_{S1,H1}T_{H1,D}T_{D,R} \quad (4)$$

also

$$T_{L,C1} = T_{L,\bar{R}}T_{\bar{R},R}T_{R,D}T_{D,H1}T_{H1,S1}T_{S1,A1}T_{A1,C1}T_{C1,C1} \quad (5)$$

From (4), derivative of  $T_{\bar{R},R}$  could be computed:

$$\begin{aligned} \dot{T}_{\bar{R},R} &= T_{\bar{R},L}T_{L,C1}\dot{T}_{C1,C1}T_{C1,A1}T_{A1,S1}T_{S1,H1}T_{H1,D}T_{D,R} \\ &\quad + T_{\bar{R},L}T_{L,C1}\dot{T}_{C1,C1}[T_{C1,A1}T_{A1,S1}T_{S1,H1}T_{H1,D}] \\ &\quad + T_{C1,A1}(T_{A1,S1}T_{S1,H1}T_{H1,D} + T_{A1,S1}T_{S1,H1}\dot{T}_{H1,D})]T_{D,R} \end{aligned} \quad (6)$$

Substituting (5) into (6)

$$\begin{aligned}\dot{T}_{\bar{R}R} = & T_{R,D}T_{D,H1}T_{H1,S1}T_{S1,A1}T_{A1,C1}\dot{T}_{C1,\bar{C}1}T_{C1,A1}T_{A1,S1}T_{S1,H1}T_{H1,D}T_{D,R} \\ & + T_{R,D}T_{D,H1}T_{H1,S1}T_{S1,A1}T_{A1,C1}[T_{C1,A1}T_{A1,S1}T_{S1,H1}T_{H1,D} + T_{C1,A1} \\ & (\dot{T}_{A1,S1}T_{S1,H1}T_{H1,D} + T_{A1,S1}T_{S1,H1}\dot{T}_{H1,D})]T_{D,R}\end{aligned}\quad (7)$$

$T_{R,D}$ ,  $T_{D,S}$ ,  $T_{S,A}$ ,  $T_{A,C}$ ,  $T_{A,S}$ ,  $T_{S,D}$ ,  $T_{D,R}$ ,  $\dot{T}_{C,\bar{C}}$ ,  $\dot{T}_{C,A}$ ,  $\dot{T}_{A,S}$ , could be calculated. Substituting these matrix into (7), evaluating the matrix product, then,  $\dot{T}_{\bar{R}R}$  is calculated. Comparing with the variables as  $\rho$ ,  $\theta_i$  and  $\delta_i$  already exist while the rover is traversing uneven terrain, the slippage variables  $\eta_i$ ,  $\xi_i$  and  $\varsigma_i$  are high-order quantities, So the derivative of  $T_{\bar{C}i,Ci}$

$$\dot{T}_{\bar{C}i,Ci} = \begin{bmatrix} -s\varsigma_i\dot{\varsigma}_i & -c\varsigma_i\dot{\varsigma}_i & 0 & -r\dot{\theta}_i + \dot{\varsigma}_i \\ c\varsigma_i\dot{\varsigma}_i & -s\varsigma_i\dot{\varsigma}_i & 0 & \dot{\eta}_i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

can derive to:

$$\dot{T}_{\bar{C}i,Ci} = \begin{bmatrix} 0 & -\dot{\varsigma}_i & 0 & -r\dot{\theta}_i + \dot{\varsigma}_i \\ \dot{\varsigma}_i & 0 & 0 & \dot{\eta}_i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\quad (8)$$

$\dot{T}_{\bar{R},R}$  can also be found for a general body in motion, using the position and orientation rates as

$$\dot{T}_{\bar{R},R} = \begin{bmatrix} 0 & -\dot{\phi}_z & \dot{\phi}_y & \dot{x} \\ \dot{\phi}_z & 0 & -\dot{\phi}_x & \dot{y} \\ -\dot{\phi}_y & \dot{\phi}_x & 0 & \dot{z} \\ 0 & 0 & 0 & 0 \end{bmatrix}\quad (9)$$

Note that the rotation part of  $\dot{T}_{\bar{R},R}$  is a skew-symmetric matrix, and the transformation-product matrix on the right-hand side of (7) also has the structure of (9). Then substituting (9) into (7), equating the matrix elements on both side of the equation, we can determine the rover configuration-rate vector [9]  $\dot{u}$  in terms of the joint angular-rates vector  $\dot{q}$ , contact-angle rate  $\dot{\delta}_i$ , wheel-rolling rate  $\dot{\theta}_i$  and the wheel slip rate vector  $\dot{\varepsilon}_i$ . Thus the Jacobian matrix and motion equation could be computed:

$$[\dot{x} \dot{y} \dot{z} \dot{\phi}_x \dot{\phi}_y \dot{\phi}_z]^T = J_i [\dot{q} \dot{\theta}_i \dot{\varepsilon}_i \dot{\delta}_i]^T$$

For short:

$$\dot{u} = J_i \dot{p}\quad (10)$$

The resulting equation of wheel-1 is as:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{\phi}_x \\ \dot{\phi}_y \\ \dot{\phi}_z \end{bmatrix} = \begin{bmatrix} J_{1,1} & J_{1,2} & J_{1,3} & J_{1,4} & J_{1,5} & J_{1,6} & J_{1,7} \\ 0 & J_{2,2} & J_{2,3} & J_{2,4} & J_{2,5} & J_{2,6} & J_{2,7} \\ J_{3,1} & J_{3,2} & J_{3,3} & J_{3,4} & J_{3,5} & J_{3,6} & J_{3,7} \\ 0 & J_{4,2} & 0 & 0 & J_{4,5} & 0 & J_{4,7} \\ J_{5,1} & 0 & 0 & 0 & J_{5,5} & 0 & J_{5,7} \\ 0 & J_{6,2} & 0 & 0 & J_{6,5} & 0 & J_{6,7} \end{bmatrix} \times \begin{bmatrix} \rho_1 \\ \dot{\psi}_1 \\ \dot{\theta}_1 \\ \dot{\xi}_1 \\ \dot{\varsigma}_1 \\ \dot{\eta}_1 \\ \dot{\delta}_1 \end{bmatrix} \quad (11)$$

in which,

$$\begin{aligned} J_{1,1} &= k_1; \\ J_{1,2} &= -c\rho c\alpha_1 c\beta_1 k_3 - s\rho k_3 s\alpha_1 - c\alpha_1 k_1 s\beta_1 + s\rho c\alpha_1 k_4 s\beta_1; \\ J_{1,3} &= c\rho r c\alpha_1 c\delta_1 c\psi_1 + r s\rho c\beta_1 c\delta_1 c\psi_1 s\alpha_1 - r s\rho c\alpha_1 c\beta_1 s\delta_1 + c\rho r s\alpha_1 s\delta_1 + r s\rho c\delta_1 s\beta_1 s\psi_1; \\ J_{1,4} &= c\rho c\alpha_1 c\delta_1 c\psi_1 + s\rho c\beta_1 c\delta_1 c\psi_1 s\alpha_1 - s\rho c\alpha_1 c\beta_1 s\delta_1 + c\rho s\alpha_1 s\delta_1 + s\rho c\delta_1 s\beta_1 s\psi_1; \\ J_{1,5} &= c\rho c\alpha_1 c\beta_1 c\delta_1 k_3 + s\rho c\delta_1 k_3 s\alpha_1 + c\alpha_1 c\delta_1 k_1 s\beta_1 - s\rho c\alpha_1 c\delta_1 k_4 s\beta_1 \\ &\quad - s\rho c\alpha_1 c\psi_1 k_3 s\delta_1 + c\rho c\beta_1 c\psi_1 k_3 s\alpha_1 s\delta_1 - s\rho c\psi_1 k_5 s\beta_1 s\delta_1 + c\psi_1 k_1 s\alpha_1 s\beta_1 s\delta_1 \\ &\quad - s\rho c\psi_1 k_4 s\alpha_1 s\beta_1 s\delta_1 - c\beta_1 k_1 s\delta_1 s\psi_1 + s\rho c\beta_1 k_4 s\delta_1 s\psi_1 + c\rho c\alpha_1 k_5 s\delta_1 s\psi_1 \\ &\quad + s\rho c\beta_1 k_5 s\alpha_1 s\delta_1 s\psi_1 + c\rho k_3 s\beta_1 s\delta_1 s\psi_1; \\ J_{1,6} &= s\rho c\psi_1 s\beta_1 - c\rho c\alpha_1 s\psi_1 - s\rho c\beta_1 s\alpha_1 s\psi_1; \\ J_{1,7} &= c\beta_1 c\psi_1 k_1 - s\rho c\beta_1 c\psi_1 k_4 - c\rho c\alpha_1 c\psi_1 k_5 - s\rho c\beta_1 c\psi_1 k_5 s\alpha_1 \\ &\quad - c\rho c\psi_1 k_3 s\beta_1 - s\rho c\alpha_1 k_3 s\psi_1 + c\rho c\beta_1 k_3 s\alpha_1 s\psi_1 - s\rho k_5 s\beta_1 s\psi_1 \\ &\quad + k_1 s\alpha_1 s\beta_1 s\psi_1 - s\rho k_4 s\alpha_1 s\beta_1 s\psi_1. \end{aligned}$$

Similarly, equations like (II) could be computed for wheel-2,3,4 and 5. Equation (II) describes the contribution of an individual wheel motion and the connecting joints rotation to the rover body motion. The net body motion is the composite effect of all wheels, and could be obtained by combining each equation for each wheel into a single matrix equation as:

$$\begin{bmatrix} I_6 \\ I_6 \\ I_6 \\ I_6 \\ I_6 \\ I_6 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{\phi}_x \\ \dot{\phi}_y \\ \dot{\phi}_z \end{bmatrix} = J \begin{bmatrix} \dot{q} \\ \dot{\theta} \\ \dot{\epsilon} \\ \dot{\delta} \end{bmatrix} \quad (12)$$

where  $J$  is formed of the each individual wheel Jacobian  $J_i$ . Besides,  $J$  is a sparse matrix.

## 4 Forward Kinematics

The purpose of forward kinematics is to solve the movement of reference point of rover body based on given or measured parameters, so that to obtain the pose and position of rover body in inertial coordinate. The pose and position usually used in navigation and tracking control.

As shown in equation (10), in which,  $\dot{u} = [\dot{x} \quad \dot{y} \quad \dot{z} \quad \dot{\phi}_x \quad \dot{\phi}_y \quad \dot{\phi}_z]^T$ , and  $\dot{p}_i = [\dot{\rho}_i \quad \dot{\psi}_i \quad \dot{\theta}_i \quad \dot{\xi}_i \quad \dot{\zeta}_i \quad \dot{\eta}_i \quad \dot{\delta}_i]^T$ ,  $\dot{p}_i$  can divide in two parts: measurable vector  $\dot{p}_{mi}$  and immeasurable vector  $\dot{p}_{ni}$ , i.e.:

$$\dot{p}_i = K_{mi}\dot{p}_{mi} + K_{ni}\dot{p}_{ni} \quad (13)$$

As to wheel- $i$ , equation (13) deduced as:

$$\begin{bmatrix} \dot{\rho}_i \\ \dot{\psi}_i \\ \dot{\theta}_i \\ \dot{\xi}_i \\ \dot{\zeta}_i \\ \dot{\eta}_i \\ \dot{\delta}_i \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\rho}_i \\ \dot{\psi}_i \\ \dot{\theta}_i \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{\xi}_i \\ \dot{\zeta}_i \\ \dot{\eta}_i \\ \dot{\delta}_i \end{bmatrix} \quad (14)$$

In which,  $\rho_1 = \rho_2 = \rho$ ,  $\rho_3 = \rho_4 = -\rho$ ,  $\rho_5 = \tau$ . Substituting (13) into (10), equation below acquired:

$$\dot{u} = J_i K_{mi}\dot{p}_{mi} + J_i K_{ni}\dot{p}_{ni} = J_{mi}\dot{p}_{mi} + J_{ni}\dot{p}_{ni} \quad (15)$$

That is:

$$\dot{u} = [J_{mi} \quad J_{ni}] \begin{bmatrix} \dot{p}_{mi} \\ \dot{p}_{ni} \end{bmatrix} \quad (16)$$

Equation (16) explains the movement of reference point caused by a single wheel. Combine five of equations for each wheel with the form of equation (16), the movement of reference point determined by five wheels is with the form below:

$$\begin{bmatrix} I - J_{n1} & 0 & 0 & 0 & 0 \\ I & 0 & -J_{n2} & 0 & 0 \\ I & 0 & 0 & -J_{n3} & 0 \\ I & 0 & 0 & 0 & -J_{n4} \\ I & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{p}_{n1} \\ \dot{p}_{n2} \\ \dot{p}_{n3} \\ \dot{p}_{n4} \\ \dot{p}_{n5} \end{bmatrix} = \begin{bmatrix} J_{m1} & 0 & 0 & 0 & 0 \\ 0 & J_{m2} & 0 & 0 & 0 \\ 0 & 0 & J_{m3} & 0 & 0 \\ 0 & 0 & 0 & J_{m4} & 0 \\ 0 & 0 & 0 & 0 & J_{m5} \end{bmatrix} \begin{bmatrix} \dot{p}_{m1} \\ \dot{p}_{m2} \\ \dot{p}_{m3} \\ \dot{p}_{m4} \\ \dot{p}_{m5} \end{bmatrix} \quad (17)$$

briefly, it is:

$$Ax = J_m \dot{p}_m \quad (18)$$

Least square method was usually adopted to solve equation (18). But it need a lot of inverse computing, which is not suitable for real time solving. So artificial searching method could be used for solving equation (18). We adopt

$$f = (Ax - J_m \dot{p}_m)^T (Ax - J_m \dot{p}_m) \quad (19)$$

as the objective function, then, it could be solved numerically.

## 5 Inverse Kinematics

### 5.1 General Solution of Inverse Kinematics

The purpose of inverse kinematics is to calculate motion (steering and rotation) of each wheel based on pose and its rate (given, calculated or measured) of body reference point. Combine equations of each wheel in form of equation (16), formula below is obtained.

$$\begin{bmatrix} I \\ I \\ I \\ I \\ I \end{bmatrix} \dot{u} = \begin{bmatrix} J_{n1} & 0 & 0 & 0 & 0 & J_{m1} & 0 & 0 & 0 & 0 \\ 0 & J_{n2} & 0 & 0 & 0 & 0 & J_{m2} & 0 & 0 & 0 \\ 0 & 0 & J_{n3} & 0 & 0 & 0 & 0 & J_{m3} & 0 & 0 \\ 0 & 0 & 0 & J_{n4} & 0 & 0 & 0 & 0 & J_{m4} & 0 \\ 0 & 0 & 0 & 0 & J_{n5} & 0 & 0 & 0 & 0 & J_{m5} \end{bmatrix} \begin{bmatrix} \dot{p}_{n1} \\ \dot{p}_{n2} \\ \dot{p}_{n3} \\ \dot{p}_{n4} \\ \dot{p}_{n5} \\ \dot{p}_{m1} \\ \dot{p}_{m2} \\ \dot{p}_{m3} \\ \dot{p}_{m4} \\ \dot{p}_{m5} \end{bmatrix} \quad (20)$$

Briefly, it is:

$$E\dot{u} = B \begin{bmatrix} \dot{q}_n \\ \dot{q}_m \end{bmatrix} \quad (21)$$

Pseudo inverse is utilized to solve equation (21):

$$\begin{bmatrix} \dot{q}_n \\ \dot{q}_m \end{bmatrix} = (B^T B)^{-1} B^{-1} E\dot{u} \quad (22)$$

So,

$$\dot{q}_m = \begin{bmatrix} [J_{m1}^T \Delta(J_{m1}) J_{m1}]^{-1} J_{m1}^T \Delta(J_{n1}) \\ [J_{m2}^T \Delta(J_{m2}) J_{m2}]^{-1} J_{m2}^T \Delta(J_{n2}) \\ [J_{m3}^T \Delta(J_{m3}) J_{m3}]^{-1} J_{m3}^T \Delta(J_{n3}) \\ [J_{m4}^T \Delta(J_{m4}) J_{m4}]^{-1} J_{m4}^T \Delta(J_{n4}) \\ [J_{m5}^T \Delta(J_{m5}) J_{m5}]^{-1} J_{m5}^T \Delta(J_{n5}) \end{bmatrix} \dot{u} \quad (23)$$

Abstract rows about  $\dot{\psi}_i$  and  $\dot{\theta}_i$ , the actuate function is obtained.

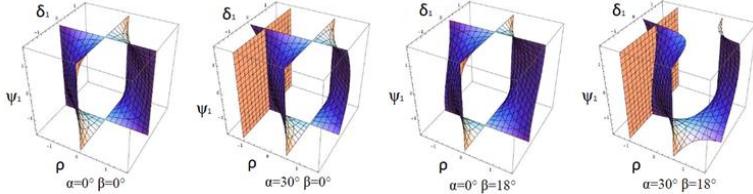
### 5.2 Applied Solution of Inverse Kinematics

However, the inverse kinematics analyzed upward have pseudo inverse involved, which is time consumptive for computing, so it will only used for theoretical analysis, rather than real time rover control. The real time actuation task is generalized as: solve  $\dot{\theta}_i$  and  $\dot{\rho}_i$  according to anticipant  $\dot{x}$  and  $\dot{\phi}_z$ . Combine the 1<sup>st</sup> and 6<sup>th</sup> equation in (II), with the assumption of slippage is zero, the real time actuation equation is as below:

$$\dot{\theta}_i = \frac{\dot{x} - J_{1,1}\dot{\rho} - J_{1,2}\dot{\psi}_i - J_{1,7}\dot{\delta}_i}{J_{1,3}} \quad (24)$$

$$\dot{\psi}_i = \frac{\dot{\phi}_z - J_{6,7}\dot{\delta}_i}{J_{6,2}} \quad (25)$$

So, if  $J_{1,3} = 0$  or  $J_{6,2} = 0$  there is no solution for (24) or (25), which means slippage is inevitable. Fig. 8 illustrates no solution space of different caster and camber.



**Fig. 8.** No solution space of different  $\alpha$  and  $\beta$

## 6 Impact of Caster and Camber to Kinematics

From the kinematics transformation prescribed, the transformation matrix from  $H_i$  to  $C_i$  is calculated as:

$$T_{H_i, C_i} = T_{H_i, S_i} T_{S_i, A_i} T_{A_i, C_i} \quad (26)$$

For wheel-1,  $T_{H_1, C_1}$  is:

$$T_{H_1, C_1} = T_{H_1, S_1} T_{S_1, A_1} T_{A_1, C_1} = \begin{bmatrix} F_{11} & F_{12} & F_{13} & x_1 \\ F_{21} & F_{22} & F_{23} & y_1 \\ F_{31} & F_{32} & F_{33} & z_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (27)$$

In which,

$$\begin{aligned} F_{11} &= c\alpha_1 c\delta_1 c\psi_1 + s\alpha_1 s\delta_1; \\ F_{21} &= c\alpha_1 s\beta_1 s\delta_1 + c\delta_1 (-c\psi_1 s\alpha_1 s\beta_1 + c\beta_1 s\psi_1); \\ F_{31} &= c\beta_1 (c\delta_1 c\psi_1 s\alpha_1 - c\alpha_1 s\delta_1) + c\delta_1 s\beta_1 s\psi_1; \\ F_{12} &= -c\alpha_1 s\psi_1; \\ F_{22} &= c\beta_1 c\psi_1 + s\alpha_1 s\beta_1 s\psi_1; \\ F_{32} &= c\psi_1 s\beta_1 - c\beta_1 s\alpha_1 s\psi_1; \\ F_{13} &= -c\delta_1 s\alpha_1 + c\alpha_1 c\psi_1 s\delta_1; \\ F_{23} &= -s\beta_1 (c\alpha_1 c\delta_1 + c\psi_1 s\alpha_1 s\delta_1) + c\beta_1 s\delta_1 s\psi_1; \\ F_{33} &= c\alpha_1 c\beta_1 c\delta_1 + s\delta_1 (c\beta_1 c\psi_1 s\alpha_1 + s\beta_1 s\psi_1); \\ x_1 &= -r c\alpha_1 c\psi_1 s\delta_1 + s\alpha_1 (r c\delta_1 + k_5); \\ y_1 &= r c\psi_1 s\alpha_1 s\beta_1 s\delta_1 - r c\beta_1 s\delta_1 s\psi_1 + c\alpha_1 s\beta_1 (r c\delta_1 + k_5); \\ z_1 &= -r c\beta_1 c\psi_1 s\alpha_1 s\delta_1 - r s\beta_1 s\delta_1 s\psi_1 - c\alpha_1 c\beta_1 (r c\delta_1 + k_5). \end{aligned}$$

### 6.1 Characteristics of $x_1$ , $y_1$ and $z_1$

$x_1$ ,  $y_1$ , and  $z_1$  illustrate the position of contact point  $C_1$  in frame  $H_1 - X_{H_1} - Y_{H_1} - Z_{H_1}$ . It is obvious that  $x_1$  is related to  $\alpha_1$ ,  $\psi_1$  and  $\delta_1$ , but  $\beta_1$  has no impact to  $x_1$ . However,  $\psi_1$ ,  $\delta_1$ ,  $\alpha_1$  and  $\beta_1$  all influence  $y_1$  and  $z_1$ . Fig.9 illustrates the impact of  $\alpha$  and  $\beta$  to  $x$  and  $y$ . Fig.10 illustrated the impact of  $\alpha$  and  $\beta$  to  $z$ .

Differentiate the 4<sup>th</sup> column of the matrix of equation (27), a equation reflects the influence to the rate of  $x$ ,  $y$  and  $z$  is as below:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{y}_1 \\ \dot{z}_1 \end{bmatrix} = \begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \\ M_{3,1} & M_{3,2} \end{bmatrix} \begin{bmatrix} \dot{\delta}_1 \\ \dot{\psi}_1 \end{bmatrix} \quad (28)$$

In which

$$M_{1,1} = -rca_1c\delta_1c\psi_1 - rsa_1s\delta_1;$$

$$M_{1,2} = rca_1s\delta_1s\psi_1;$$

$$M_{2,1} = rc\delta_1c\psi_1s\alpha_1s\beta_1 - rca_1s\beta_1s\delta_1 - rc\beta_1c\delta_1s\psi_1;$$

$$M_{2,2} = -rc\beta_1c\psi_1s\delta_1 - rsa_1s\beta_1s\delta_1s\psi_1;$$

$$M_{3,1} = -rc\beta_1c\delta_1c\psi_1s\alpha_1 + rca_1c\beta_1s\delta_1 - rc\delta_1s\beta_1s\psi_1;$$

$$M_{3,2} = -rc\psi_1s\beta_1s\delta_1 + rc\beta_1s\alpha_1s\delta_1s\psi_1.$$

This equation establishes the relationship between the rate of contact point and frame  $H$ , also, it shows the additional velocity brought by the impact of caster and camber.

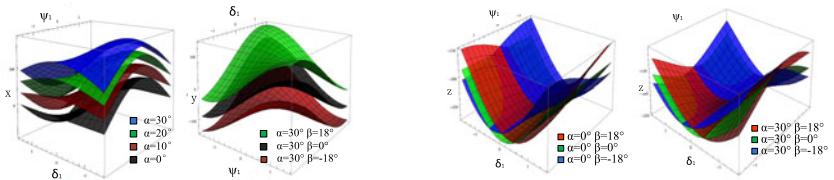


Fig. 9. The impact of  $\alpha$  and  $\beta$  to  $x$  and  $y$

Fig. 10. The impact of  $\alpha$  and  $\beta$  to  $z$

### 6.2 Equivalent Steer Angle

In the ground coordinate  $X - Y$ , the wheel with the caster and camber is equivalent to a wheel with equivalent eccentric distance, which rotates an equivalent angle  $\lambda_z$  in the  $X - Y$  plane. The equivalent steer angle is acquired as follows: the rotating components in (27) are equivalent to revolving around the axis  $y$ ,  $x$  and  $z$  at an angle of  $\lambda_y$ ,  $\lambda_x$ , and  $\lambda_z$  respectively:

$$Rot = \begin{bmatrix} c\lambda_z & -s\lambda_z & 0 \\ s\lambda_z & c\lambda_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c\lambda_x & -s\lambda_x \\ 0 & s\lambda_x & c\lambda_x \end{bmatrix} \begin{bmatrix} c\lambda_y & 0 & s\lambda_y \\ 0 & 1 & 0 \\ -s\lambda_y & 0 & c\lambda_y \end{bmatrix} \\ = \begin{bmatrix} c\lambda_y c\lambda_z - s\lambda_z s\lambda_y s\lambda_z & -c\lambda_x s\lambda_z & c\lambda_z s\lambda_y + c\lambda_y s\lambda_x s\lambda_z \\ c\lambda_z s\lambda_x s\lambda_y + c\lambda_y s\lambda_z & c\lambda_x c\lambda_z & -c\lambda_y c\lambda_z s\lambda_x + s\lambda_y s\lambda_z \\ -c\lambda_x s\lambda_y & s\lambda_x & c\lambda_x c\lambda_y \end{bmatrix} \quad (29)$$

The terms (2,1) and (2,2) in (27) and (29) could be used to obtain the equivalent steering angle  $\lambda_z$ :

$$\lambda_z = \arctan \frac{c\alpha_1 s\psi_1}{c\beta_1 c\psi_1 + s\alpha_1 s\beta_1 s\psi_1} \quad (30)$$

when  $\alpha_1 = 0$ ,  $\lambda_z = \arctan (sec\beta_1 tan\psi_1)$ ; when  $\beta_1 = 0$ ,  $\lambda_z = \arctan (cos\alpha_1 tan\psi_1)$ ; when  $\alpha_1 = 0$ ,  $\beta_1 = 0$ ,  $\lambda_z = \psi_1$ . So,  $\alpha$  will reduce the steering ability, but  $\beta$  will enhance it.

## 7 Conclusion

A methodology is presented for the kinematics modeling based on a castered-and-cambered 5WS-5WD ATR. Foward kinematics analysis make it possible to determine configuration-rate of the body reference. Inverse kinematics analysis gives a method to determine rotation speed and steering for individual wheel. Slippage of each wheel is abstracted and considered in analysis. In the end, the impact of caster and camber to the kinematics is discussed, especially, the impact to the displacement and velocity of the rocker end  $H_i$  is explained; besides, the equivalent steer angle of a single wheel is calculated.

## References

1. Hayati, S., et al.: The Rocky 7 rover: A Mars science craft prototype. In: Proc. IEEE Int. Conf. Robot. Autom., Albuquerque, NM, pp. 2458–2464 (1997)
2. Baerveldt, A.-J. (ed.): Agricultural Robotics. In: Autonomous Robots, vol. 13-1, Kluwer Academic Publishers, Dordrecht (2002)
3. Cunningham, J., Roberts, J., Corke, P., Durrant-Whyte, H.: Automation of underground Lhd and truck haulage. In: Proc. Australian IMM Conf., pp. 241–246 (1998)
4. [http://solarsystem.nasa.gov/news/display.cfm?News\\_ID=10263](http://solarsystem.nasa.gov/news/display.cfm?News_ID=10263)
5. Muir, P.F., Neuman, C.P.: Kinematic modeling of wheeled mobile robots. J. Robot. Syst. 4(2), 282–340 (1987)
6. Rajagopalan, R.: A Generic Kinematic Formulation for Wheeled Mobile Robots. Journal of Robotic Systems 14, 77–99 (1997)
7. Yi, B.-J., Kim, W.K.: The kinematics for redundancy actuated omnidirectional mobile robots. J. Robot. Syst. 19(6), 255–267 (2002)
8. Tarokh, M., McDermott, G., Hayati, S., Hung, J.: Kinematic modeling of a high mobility Mars rover. In: Proceedings of the IEEE International Conference on Robotics and Automation. Detroit, Michigan, pp. 992–998 (1999)
9. Tarokh, M., McDermott, G.: Kinematics Modeling and Analysis of Articulated Rovers. IEEE Trans. Robotics 21(4), 539–553 (2005)
10. Xu, H., Xue, K., et al.: 3-D Kinematics Modeling for Mobile Robot with Steering Castered-and-Cambered Wheels. In: Proceedings of the IEEE International conference on Mechatronics and Automation, Harbin, China, pp. 1345–1350 (2007)

# Experimental Investigation of a Prediction Algorithm for an Indoor SLAM Platform\*

Jung-Fu Hou<sup>1</sup>, Yu-Shin Chou<sup>2</sup>, Yau-Zen Chang<sup>1</sup>, and Jing-Sin Liu<sup>2,\*</sup>

<sup>1</sup> Department of Mechanical Engineering, Chung-Gung University, Taiwan

<sup>2</sup> Institute of Information Science, Academia Sinica, Taiwan

liu@iis.sinica.edu.tw

**Abstract.** This paper presents a scheme for the indoor simultaneous localization and mapping (SLAM) problem. The scheme is based on the scan matching method and is treated as an optimization problem solve by the Simplex method. The two-dimensional distance transform method is used to facilitate the cost value evaluation. In order to register scanned maps with built map through maximum overlap between the maps, a predictive algorithm is proposed. The algorithm can not only reduce search scope but also discard unexpected objects that may cause false match. The approach is investigated by an experimental platform with differential drives. The ICP-SLAM is also implemented for performance comparison. Experimental result shows that the prediction algorithm can improve accumulation error in the indoor environment.

**Keywords:** Simultaneous localization and mapping, Simplex method, Iterative closest point.

## 1 Introduction

In order to make mobile robot to have intelligent behavior such as path planning, obstacles avoiding. Mobile robot needs to know its position within environments. Self localization and map construct is an importance issue and need to be solved. This problem also calls a simultaneous localization and mapping (SLAM) [1][2][3][4]. Two-dimensional laser range finder (LRF) is one of the common used sensors to solve this problem. In laser scan matching method, the position and the orientation of the vehicle is estimated by rotating and translating the current scanned map and compare to the reference map until the max overlap is achieved. In modern SLAM approaches, some research use extended Kalman filter (EKF) to combine the information from laser scan and odometry data to solve SLAM problem. This form of SLAM is known as EKF-SLAM [5][6]. The main problem is laser scanned data need to be analyzed and to form a set of feature points. These feature point usually come from the geometry properties of scanned data such as corner, line curvature. Due to the lack of feature point detection, this may cause lower system performance and inaccuracy map mapping. Another kind of approach is used scan matching method. This method can be categorized into feature to feature, point to feature, point to point. In the

---

\* Corresponding author.

feature to feature approaches, the property of laser scanned data, such as line segments, corners are extracted then matched directly [7]. This kind of approach requires the extract features from laser scanned data precisely. In the point to feature approaches, the feature such as line is extracted from reference scanned data and forms a predefined map. Then current scanned data is matched to this predefined map. The point to point approaches do not need to analysis scanned data. The iterative closest point (ICP) is the most popular method to align two different point set [8][9]. In this method, the smallest Euclidean distance for every point in current scanned data to the reference scanned data is calculated to be the cost value to estimate the best overlap. The ICP method has better convergence performance in translation then rotation. The IMRP method [9][10] is then proposed to estimate the rotation.

In this paper, we present a scheme to solve SLAM problem. This scheme treat mapping problem as an optimization problem. In order to achieve the fast convergence, Simplex method is used and Occupy grid map method is used to reduce the size of computation data. Unlike the ICP or IMRP, Euclidean distances are pre-calculated and form a distance map before optimization process. A predict algorithm is used to find out the maximum overlap in two different maps.

The rest paper is organized as follows: distance transform is described in section 2.1. The concept of prediction algorithms is described in section 2.2. The Simplex method is described in section 2.3. Then the experiment is in section 3. Finally the conclusion is shown in section 4.

## 2 Method

### 2.1 Distance Transform

The concept of distance transform (DT) is to calculate the Euclidean distance from certain position on map to the closest obstacle. The example of DT is shown in Fig. 1. Before distance transform, the value of obstacle position is set to zero and other position is set to infinity (Fig1. (a)). After distance transform, the distance of every non-obstacle position to the closest obstacle is set, as shown in Fig 1. (b).

Inf									
Inf									
Inf	Inf	0	Inf	Inf	Inf	0	Inf	Inf	Inf
Inf	Inf	0	Inf	Inf	Inf	0	Inf	Inf	Inf
Inf	Inf	0	Inf	Inf	Inf	0	Inf	Inf	Inf
Inf	Inf	0	Inf	Inf	Inf	0	Inf	Inf	Inf
Inf	Inf	0	0	0	0	0	Inf	Inf	Inf
Inf									
Inf									
Inf									

(a)

**Fig. 1.** (a) map before distance transform. (b) result of distance transform

4	3	2	3	4	3	2	3	3	4
3	2	1	2	3	2	1	2	2	3
2	1	0	1	2	1	0	1	2	3
2	1	0	1	2	1	0	1	2	3
2	1	0	1	2	1	0	1	2	3
2	1	0	1	1	1	0	1	2	3
2	1	0	0	0	0	0	1	2	3
3	2	1	1	1	1	1	2	2	3
4	3	2	2	2	2	2	2	3	4
5	4	3	3	3	3	3	3	3	4

(b)

**Fig. 1.** (continued)

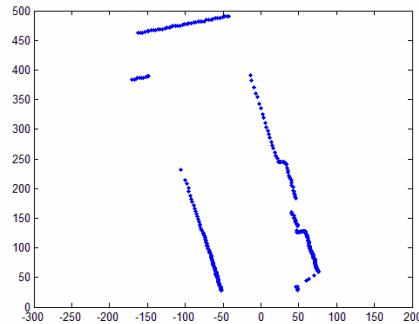
In this paper, data from laser range finder is formatted as the angle ( $\phi$ ) and relate distance (d) can be denoted as:

$$D = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 \dots \phi_n \\ d_1 & d_2 & d_3 \dots d_n \end{bmatrix} \quad (1)$$

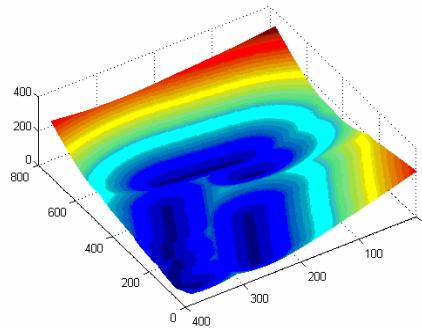
The  $D$  are then transform into Cartesian coordinate and expressed as:

$$D_c = \begin{bmatrix} x_1 & x_2 & x_3 \dots x_n \\ y_1 & y_2 & y_3 \dots y_n \end{bmatrix} \quad (2)$$

According to  $D_c$ , an obstacle map can be built, as shown in Fig. 2.

**Fig. 2.** Obstacle Map

Then the distance transform are then applied on obstacle map to form a DT Map. If we take the value of distance as Z-Axis, a three-dimensional figure can be form as shown in Fig. 3. In this figure, the darkest color (blue) indicates the position on the

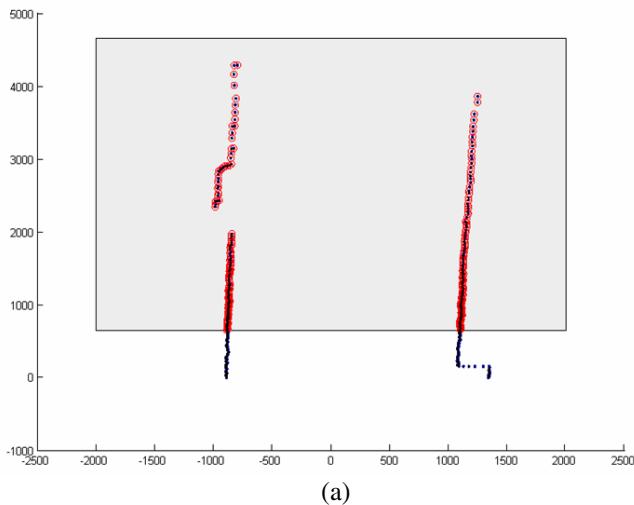


**Fig. 3.** The result of distance transform

map is closer than the bright color (red). This map will be used to generate cost value in the optimization process.

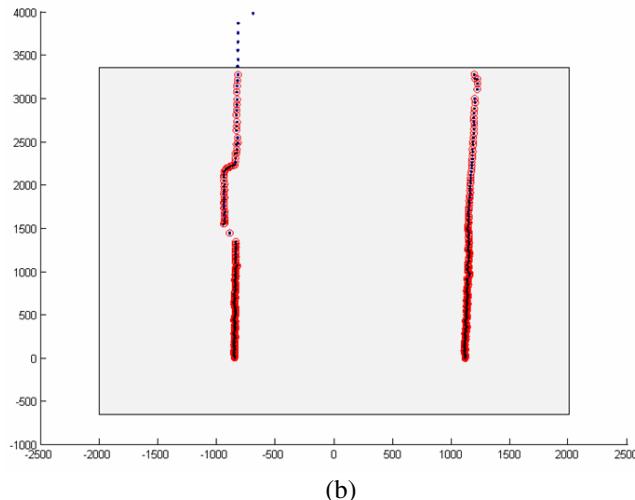
## 2.2 Map Prediction

For the two different time step scanned map ( $P(t)$ ,  $P(t+1)$ ), maps will change due to the motion of vehicle. Compare to these two maps, some points will disappear and some point will be discovered. These points will affect the mapping result. In order to find out the max overlap area, a prediction algorithm is proposed. According to the odometry of vehicle, a roughly position of vehicle can be estimated. By using this information, a virtual laser scanned area can be applied to predict the overlap points. The concept of map prediction is shown in Fig. 4; the rectangle indicated the virtual scanned area. The figure 4(a) shows the predict area from  $P(t)$  to  $P(t+1)$ . And the Fig.4(b) indicated the predict area from  $P(t+1)$  back to  $P(t)$  in the forward motion.



(a)

**Fig. 4.** (a) Prediction  $P(t)$  to  $P(t+1)$ . (b) Prediction  $P(t+1)$  back to  $P(t)$

**Fig. 4. (continued)**

In the experiment, the motion vehicle is separated into two modes: forward and rotation mode. The pseudo code is shown below:

```

INPUT: Pt (xn, Yn, θn), Pt+1 (xm, Ym, θm), T(x, y, θ)
OUTPUT: P_pret (xn, Yn, θn), P_pret+1 (xm, Ym, θm)

IF T(y) > 0 // forward mode
    FOR i from 1 to n
        IF Pt(yi) < y + min(Pt+1(y))
            Delete (Pt(xi, yi, θi));
        ENDIF
    END
    FOR i from 1 to m
        IF Pt+1(yi) > y + max(Pt(y))
            Delete (Pt+1(xi, yi, θi));
        ENDIF
    END
ELSE // rotation mode
    FOR i from 1 to n
        IF Pt(θi) < θ
            Delete (Pt(xi, yi, θi));
        ENDIF
    END
    FOR i from 1 to m
        IF Pt+1(θi) > 180-T(θ)
            Delete (Pt+1(xi, yi, θi));
        ENDIF
    END
END

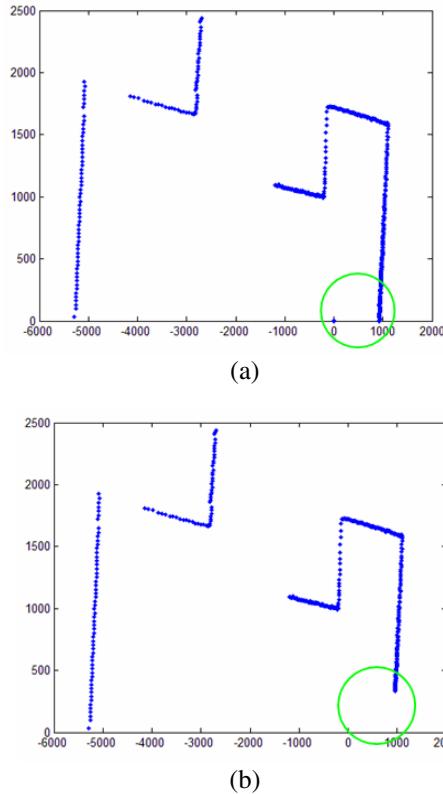
```

```

ENDIF
P_pret=Pt;           // output
P_pret+1=Pt+1;

```

From the above method, Fig. 5 shows two maps before and after prediction in rotation mode. The differences between two maps are shown in the green circle in Fig. 5(a) and (b). After the prediction process, the point that will disappear in next time step is discarded. And max overlap area of two maps is achieved to perform mapping.



**Fig. 5.** (a) Map before prediction. (b) map after prediction.

### 2.3 Simplex Method in Map Mapping

Simplex method is a kind of optimization algorithm to estimate the optimal value in the cost function. The purpose of simplex method is to estimate the relation between two different maps ( $P(t)$  and  $P(t+I)$ ). This relation can define by a transformation

matrix as below. Where  $x$ ,  $y$ ,  $\theta$  are defined as the x-axis, y-axis shift and rotation angle separately.

$${}_{t+1}^t T = T(x, y, \theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & x \\ \sin \theta & \cos \theta & 0 & y \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

For different time step maps such as  $P(t+n)$  can be simply multiply on transformation estimated before, as Eq(4).

$${}_{t+n}^t T = {}_{t+1}^t T \cdot {}_{t+2}^{t+1} T \cdots {}_{t+n}^{t+(n-1)} T = {}_{t+(n-1)}^t T \cdot {}_{t+n}^{t+(n-1)} T \quad (4)$$

In the definition of cost function, after distance transform is applied on reference map ( $P(t)$ ). The prediction algorithms are applied on current scanned point data ( $P(t+1)$ ) to discard unexpected point. And then transfer to reference to reference DT map according following equation.

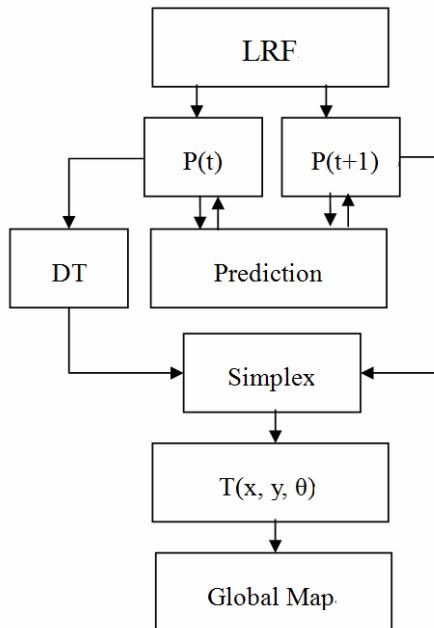
$$P_{trans} = {}_{t+1}^t T \cdot P_{current} \quad (5)$$

Then the cost value is defined as fig. 6. The red color grid on the distance map denoted the obstacle point on the reference map ( $P(t)$ ). And the yellow color grid is the transformed point data project on to the distance map, and the sum of distance ( $D_{sum}$ ) of yellow grid is used as cost value. The Simplex method is used to estimate the smallest  $D_{sum}$ .

22	14	10	14	22	28	36	42	50	57
20	10	0	10	14	22	28	36	42	50
22	14	10	0	10	14	22	28	36	42
28	22	14	10	0	10	14	22	28	36
36	28	22	14	10	0	10	14	22	32
42	36	28	22	14	10	0	10	20	30
36	28	22	14	10	0	10	14	22	32
32	22	14	10	0	10	14	22	28	36
30	20	10	0	10	14	22	28	36	42
32	22	14	10	14	22	28	36	42	50

**Fig. 6.** Example of cost calculation

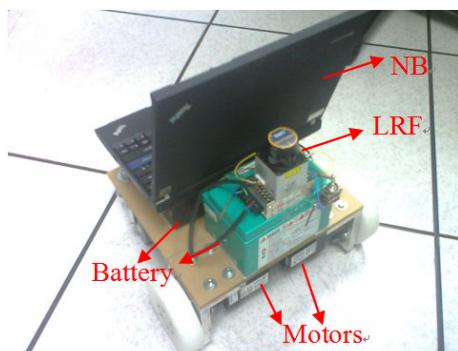
The complete flowchart of proposed scheme is shown in Fig. 7.



**Fig. 7.** Flow chart of proposed scheme

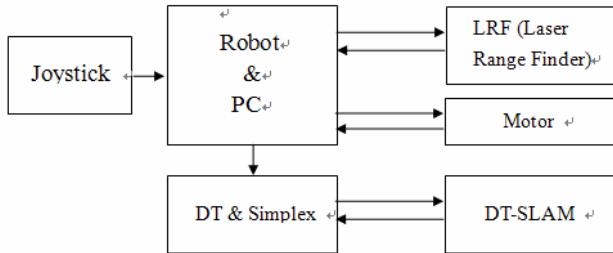
### 3 Experiment

In this paper, a vehicle with differential drive is used, as shown in Fig. 8. Laser range finder is mounted in the front of vehicle. A notebook is used to control the vehicle and perform the SLAM task. Two motor made by MONTROL is used and controlled through RS-232. The laser range finder is manufactured by HOKUYO URG-04LX. Scanned data is captured through USB.



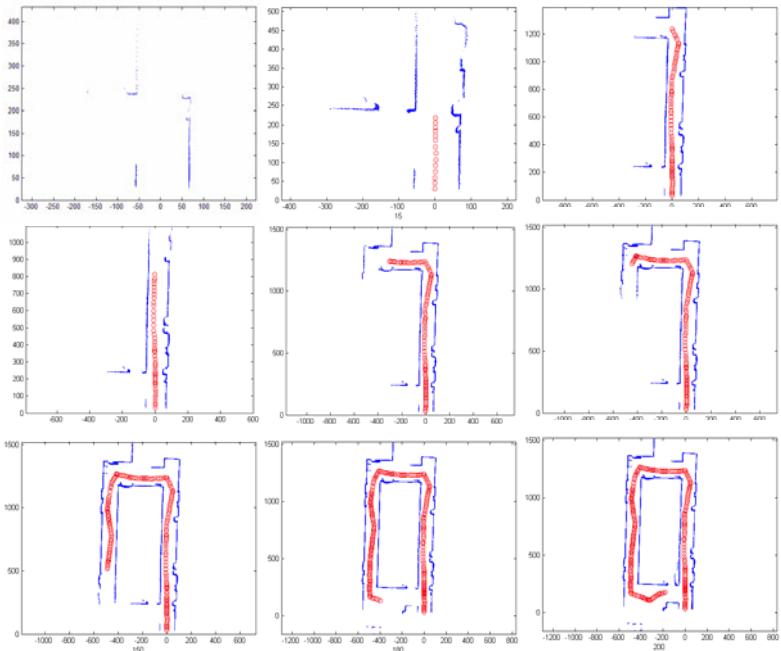
**Fig. 8.** Picture of experimental vehicle

The control structure of mobile robot is shown in Fig. 9. The robot is driven by two motor. Data from LRF is received by a notebook through RS-232, Sample rate is 2 seconds. The direction of robot move is controlled by a joystick.



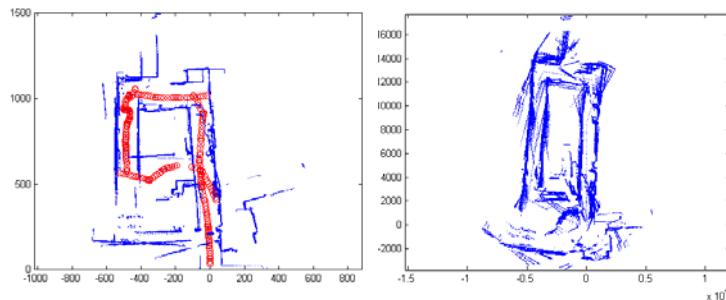
**Fig. 9.** Control structure of mobile robot

In the experiment, two methods are implemented to validate the performance of proposed scheme: odometry only, and ICP-SLAM. The experiment is in a 30mx5m indoor environment. In the first experiment, we drive vehicle to move around rectangular like environment. The robot is driven to move around this area to observe the accumulation error of proposed scheme. The result of experiment is shown in Figure 10.



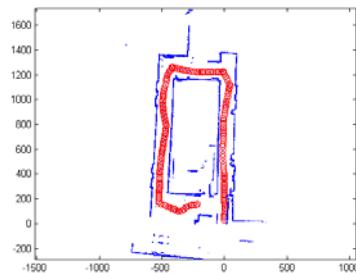
**Fig. 10.** Result for test environment

The experimental result of ICP and odometry only method is shown in Fig. 11. Compared with the ICP method, serious accumulation errors problem can be observed. These errors cause the SLAM map inaccuracy. The odometry only method also can find such problems.

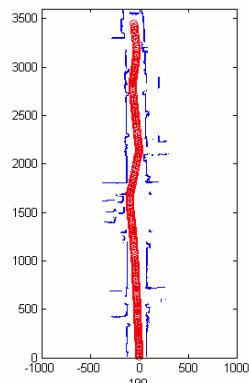


**Fig. 11.** (a) Result of ICP method (b) Result of odometry only

In order to test the performance of the prediction algorithm, the same scheme except prediction algorithm is performed. The result is shown in Fig. 12. In the green circle of Fig. 12, one can figure that the floor pillar doesn't converge well as prediction is applied.



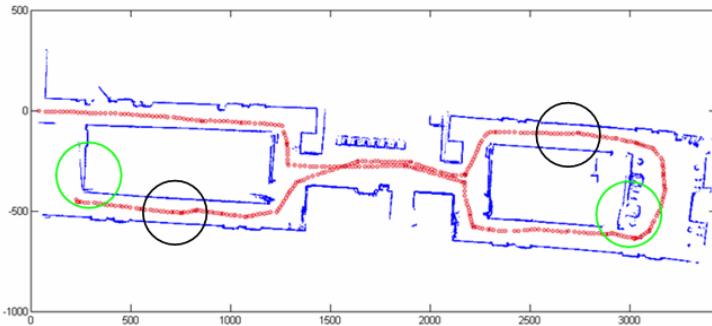
**Fig. 12.** Result of proposed scheme without prediction



**Fig. 13.** Result for long corridor environment

In another experiment, a long corridor environment is used to test the accumulation error. The result is shown in Fig. 13. The total distance of the corridor is 35m, and 72 scanned data is captured to construct the environment.

The last experiment is the largest scale, the whole environment of office. 450 data are used to construct the map. The result is shown in Fig. 14.



**Fig. 14.** Result for large scale indoor environment

In the green circle from above figure, the affect of accumulation error is apparent. This is caused by the poor point feature in the parallel wall scanned in the corridor (the black circle shown in Fig. 13.) To overcome this drawback, global optimization algorithms should be applied in the future.

## 4 Conclusion

In this paper, we proposed a scheme which integrates distance transform and the concepts of prediction into map mapping task. And the mapping problem is transform into a optimization problem. The experiment shows it could be continued in the surrounding environment. Even the maps have 450 maps to composition. The proposed scheme still gets efficiency to complete the SLAM task. However, there are some problems in the result and need to be overcome in the future. Although the integrated error has been decreased by prediction algorithm, but for a large scale environment it still exist. Global map optimizations algorithms need to be develop to overcome this drawback.

## References

1. Whyte, H.D., Bailey, T.: Simultaneous Localization and Mapping: Part I. IEEE Robot. Autom. Mag. 13(2), 99–110 (2006)
2. Dissanayake, M., et al.: A Solution to the Simultaneous Localization and Map Building (SLAM) Problem. IEEE Trans. Robotics, Automation 17(3), 229–241 (2001)
3. Robert, O., Kotaro, H.C.: A Comparison of SLAM Implementations for Indoor Mobile Robots. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, pp. 1479–1484 (2007)

4. Martinelli, A., et al.: A Relative Map Approach to SLAM Based on Shift and Rotation Invariants. *Robotics and Autonomous Systems* 55(1), 50–61 (2007)
5. Gutmann, J.-S.: Robuste Navigation autonomer mobiler Systeme. PhD thesis, Albert-Ludwigs-Universität Freiburg (2000)
6. Hähnel, D., Burgard, W., Fox, D., Thrun, S.: An efficient FastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. In: IROS 2003, vol. 1, pp. 206–211. IEEE, Los Alamitos (2003)
7. Rongchuan, S., Shugen, M., Bin, L., Yuechao, W.: Simultaneous Localization and Sampled Environment Mapping. In: 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, Shanghai, P.R. China, pp. 6684–6689 (2009)
8. Besl, P.J., McKay, N.D.: A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2), 239–256 (1992)
9. Lu, F.: Shape Registration Using Optimization for Mobile Robot Navigation. PhD thesis, University of Toronto (1995)
10. Lu, F., Milios, E.: Robot pose estimation in unknown environments by matching 2D range scans. *J. of Intelligent and Robotic Systems* 20, 249–275 (1997)
11. Nguyen, V., Gächter, S., Martinelli, A., Tomatis, N., Siegwart, R.: A comparison of line extraction algorithms using 2D range data for indoor mobile robotics. *Autonomous Robots* 23(2), 97–111 (2007)
12. McFetridge, L., Ibrahim, M.Y.: A new methodology of mobile robot navigation: The agoraphilic algorithm. In: *Robotics and Computer-Integrated Manufacturing*, pp. 545–551. Elsevier Press, Amsterdam (2009)
13. Castej, C., Blanco, D., Boada, B.L., Moreno, L.: Voronoi extraction of free-way areas in cluttered environments. In: The 11th international conference on Advanced Robotics. Coimbra, Portugal (2003)
14. Nidhi, K., Dave, F., Anthony, S.: Incremental Reconstruction of Generalized Voronoi Diagrams on Grids. In: International Conference on Intelligent Autonomous Systems(IAS), IOS Press, Amsterdam (2006)
15. Minguez, J.: The Obstacle-Restriction Method (ORM) for Robot Obstacle Avoidance in Difficult Environments. In: Intelligent Robots and Systems, pp. 3706–3712. IEEE press, Los Alamitos (2005)
16. Lars, H., Hanna, S.: An evaluation of measures for quantifying map information. *ISPRS Journal of Photogrammetry and Remote Sensing* (2007)
17. Felix, W., Frederic, M., Joaquin, S., Howie, C., Stephen, T., George, K.: Topological SLAM using Neighbourhood Information of Places. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 4937–4942. IEEE press, St. Louis (2009)

# Trajectory Estimation of a Skid-Steering Mobile Robot Incorporating Free Wheels

Tokuji Okada<sup>1</sup>, Abeer Mahmoud<sup>2</sup>, and Wagner Votelho<sup>1</sup>

<sup>1</sup> Niigata University, Niigata, Japan

okada@eng.niigata-u.ac.jp

<sup>2</sup> Ain Shams University, Cairo, Egypt

**Abstract.** This paper proposes a trajectory estimation method for a skid-steering mobile robot (SSMR) incorporating free wheels on uneven but smooth ground with different ground properties and distributed loads on wheel axes. We formulate a mathematical expression in terms of virtual work for evaluating the energy cost of all the wheel motions with skidding and slipping under the condition that each wheel is always in contact with the ground via a pivot suspension structure. Minimization of the expression for virtual work results in a unique solution for the new position of the robot after small movement of the wheels. Further, by iterating we calculate a continuous trajectory which is an important result for assigning powered wheel velocities for self-navigation. Trajectories demonstrated by a test robot confirm the validity of the method.

## 1 Introduction

Most studies conducted on wheeled mobile robots (WMRs) assume that their wheels satisfy a nonslip and nonskid conditions, however, in reality, tires deform while steering, thus violating both these assumptions. Therefore, implementing an exact model with a practical control scheme is important for the study of wheel behavior. A differential angular velocity between the two sides of a robot determines the steering of skid-steering mobile robots (SSMRs). In particular, a differential gearing system is beneficial for adjusting differential angular velocities without using steering mechanisms. A robot equipped with such a gearing system and free casters will move without sliding or skidding. A kinematic approach for analyzing tracked mobile robots has previously been proposed [1]. However, the SSMRs in that study are characterized by all skidding wheels using ground-wheels interaction and skidding effect. Meanwhile, stable motions of SSMRs are formulated in [2, 3]. A 4WD skid-steering EV has been developed in which the available torque is distributed to each wheel in order to execute the driver's commands [4]. Integrated estimation of WMR motion in the presence of wheel skidding and slipping has also been discussed [5]. In another study, a kinematic model applied to the control of an SSMR resulted in improved stability regulation and trajectory tracking in both position and orientation [6].

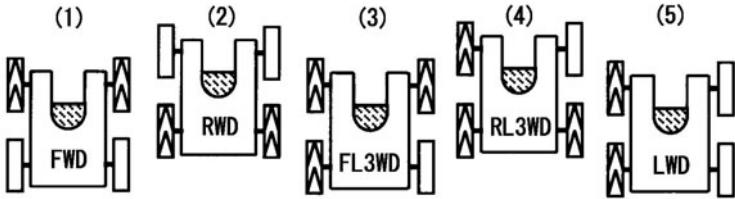
Skid-steer vehicles (SSVs) require increased power when cornering since turning requires an increased force to be exerted by the steering mechanism [7, 8].

In addition, SSVs show markedly higher tire wear compared with Ackermann-steered vehicles because of the relatively high values of friction encountered during skidding and slipping when steering. However, the absence of an explicit-steering assembly makes four wheel drive SSVs (4WDSSVs) mechanically robust and able to move on rough terrain with toughness and good maneuverability by controlling the wheels simply by varying their angular velocity, for example, as an all-terrain response vehicle (ATRV). Skid-steering based on dynamic models for real-time motion control of  $n$ -wheel driving SSMRs may be costly. Therefore, it is more practical and effective to control wheel pairs at opposite sides of the nWDSSMRs with differential velocity inputs.

This paper extends our previous trajectory estimation method [9] to apply not only to four-wheel-drive SSRs, but also to general SSMRs having decreased driving wheels either intentionally or unexpectedly. This paper focuses on various wheel driving arrangements, some of which include free wheels. Experimental data gathered from an actual robot to verify the method for estimating motion trajectories are presented.

## 2 Nomenclature

- ( $a, b$ ) :X and Y components of the robot center on the planes of  $\lambda_o$  and  $\lambda$ .  
 $(a_e, b_e)$  shows a small linear displacement
- $i$  :Subscript related to the wheel position ( $i = 1-5$ )
- $d_i$  :Translational wheel displacement, i.e.,  $\overline{P_i P'_i}$
- $\ell_i$  :Expected rolling distance of the wheel within  $\Delta t$
- $A_i$  :Wheel tread area (width;  $t_{ai}$ , length;  $t_{bi}$ )
- $L$  :Body size (body length;  $L_h$ , body width;  $L_w$ )
- $P_{i0}$  :Initial wheel position ( $x_{i0}, y_{i0}, z_{i0}$ ) on  $\zeta_0$
- $P'_i$  :Actual wheel position ( $x'_i, y'_i, z'_i$ ) on  $\zeta$  after a small wheel rotation
- $P_{si}$  :Expected wheel position ( $x_{si}, y_{si}, z_{si}$ ) on  $\zeta$  when the wheel rolls forward by a small angular displacement
- $P_c$  :Geometric center of the robot
- $P_g$  :Center of mass of the robot
- $P_5$  :Imaginary point where the resultant force at  $P_1$  and  $P_2$  operates
- $P_6$  :The point where the axis  $\eta$  crosses the vertical plane including  $P_1$  and  $P_2$
- $P_7$  :Vertical projection of  $P_g$  on the plane  $\zeta$
- $R_i$  :Wheel radius
- $U_i$  :Supposed virtual distance, i.e.,  $\overline{P'_i P_{si}}$
- $W_{i0}$  :Load at the wheel contact on  $\zeta_0$  ( $i = 1-5, 7$ )
- $W_i$  :Load at the wheel contact on  $\zeta$  (total load,  $W_t$ )
- $\alpha$  :Angular shift of a wheel from an initial direction.  $\alpha_e$  shows a small displacement
- $\eta$  :Axis of a pivot suspension; virtual body axis extending straight forward parallel from  $P_c$  internally around which the front body part is compliant and twist in order to have four wheels in contact with an irregular terrain
- $\phi$  :Twist angle of the front body part around  $\eta$
- $\theta_r$  :Wheel rotation angle (small increment,  $\Delta\theta_r$ )



**Fig. 1.** Top view of SSWMRs illustrating various driving wheel arrangements including free wheels

$\theta_s$  :Inclination angle of the ground surface

$\mu_i$  :Friction coefficient (static;  $\mu_{si}$ , dynamic;  $\mu_{di}$ )

$\nu_i$  :Linear speed of a wheel movement

$\omega_i$  :Angular velocity of  $i$ -th wheel (left;  $\omega_l$ , right;  $\omega_r$ )

$\lambda_\nu$  :Ratio of  $\omega_l$  to  $\omega_r$  ( $= \nu_r / \nu_l$ )

$\zeta_0$  : $X_0-Y_0$  plane; horizontal home ground of the robot before the coordinate transformation

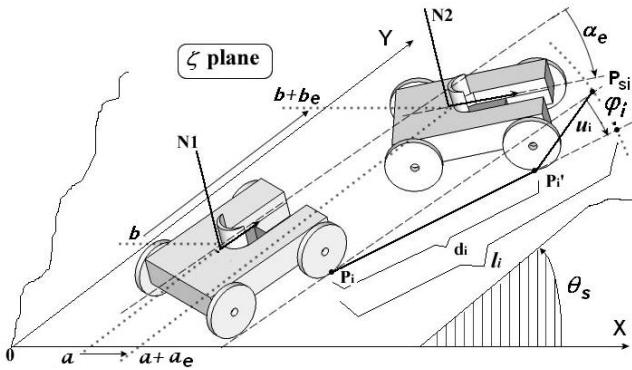
$\zeta$  : $X-Y$  plane; actual ground surface generated after the coordinate transformation of  $\Sigma_0$  with  $\theta_s$  around the  $X_0$  axis

$\Sigma_0$ :Coordinate system of the horizontal home ground

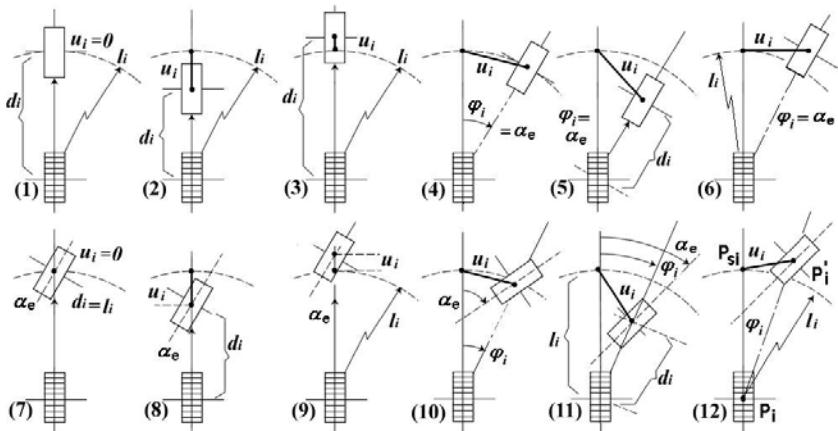
$\Sigma$  :Coordinate system of the ground inclined by  $\theta_s$ .

### 3 Wheel Driving Axel Arrangements

Wheeled robots move whenever there is at least one wheel driving wheel. Multi-wheeled vehicles may be assembled with wheels of different sizes, and some may be driving while others are not. Some wheels may roll freely without necessitating drive power by means of a link, chain, or belt. Steam locomotives, for example, make use of such a link. In skid-steering mobile robots and vehicles, arrangements where all the wheels are powered, e.g., 4WD are the most effective for generating large traction forces. However, unbalanced arrangements can occur in unexpected conditions. Various wheel arrangements (excluding the 4WD) are shown in Fig. 1. Here, rectangles with arrow-like hatches show powered wheels and their directions of movement, and rectangles without hatches show free wheels. Arrangements (1) and (2) are FWD (Front WD) and RWD (Rear WD), respectively. These balanced arrangements are popular when an explicit steering mechanism is used. The unbalanced arrangements (3) and (4) occur in unexpected conditions. We refer to them as *FL3WD* (Front-Left 3WD) and *RL3WD* (Rear-Left 3WD), respectively. Similarly, *FR3WD* (Front-right 3WD) and *RR3WD* (Rear-right 3WD) are also considered. These 3WD arrangements replace a driving wheel with a free wheel and operate better than FWD and RWD in steering performance. Arrangement (5) in Fig. 1, LWD (Left WD) and similarly RWD (Right WD), which is not shown in the figure will rarely happen. In this paper, we do not consider special arrangements such as a crawler or a robot in which the powered wheels are located diagonally.



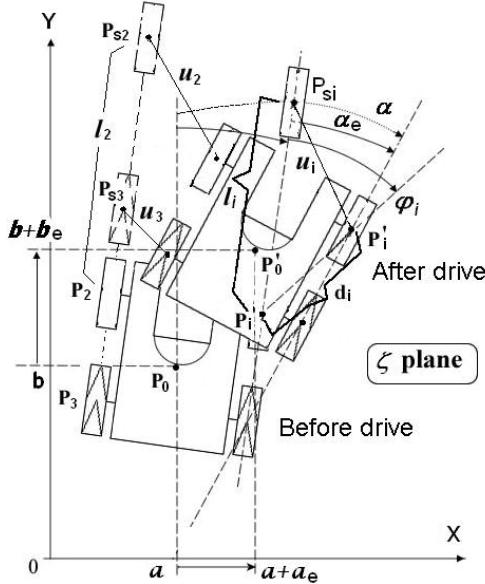
**Fig. 2.** Robot posture before and after small movement on a slope inclined at  $\theta_s$ .  $N_1$  and  $N_2$  are normal to the robot platform.



**Fig. 3.** Possible patterns of skidding and slipping behaviors as a mixture of twists and translations. Unhatched and hatched rectangles show wheel positions before and after small motion, respectively.

#### 4 Analysis of Skidding and Slipping Behaviors

Fig. 2 shows the movement of a 4WSSMR on the slope. Suppose that each wheel skids and slips while rolling. Then, the possible wheel behaviors can be categorized into eleven patterns shown in Fig. 3. Each pattern is not determined locally, because the motion is interactive among the four wheels (see Fig. 2). Since the robot trajectory cannot be expressed using a closed algebraic expression, we seek to solve for a continuous path by iteratively solving for incremental changes in the robot's position caused by small changes to the driving wheel angles.



**Fig. 4.** Illustration of virtual distance  $U_i$  connecting  $P_{si}$  and  $P'_i$

#### 4.1 Transformation of Wheel Tread Position

Suppose that the robot stands on the plane  $\zeta_0$  by aligning  $P_c$  on the  $Z_0$ -axis and having the robot point toward the  $Y_0$  direction. In Addition, imagine that each wheel rolls a distance  $\ell_i$  which is equal to  $R_i\omega_i\Delta t$  when drive power is supplied. Then,  $P_{si}(x_{si}, y_{si}, z_{si})$  in the system  $\Sigma$  is combined with  $P_{i0}(x_{i0}, y_{i0} + \ell_i, z_{i0})$  in the system  $\Sigma_0$  (Fig. 4). In fact,  $\ell_i$  is zero when the robot locks a wheel  $i$  from rolling. These relations are written in matrix form as given below:

$$\begin{bmatrix} x_{si} \\ y_{si} \\ z_{si} \\ 1 \end{bmatrix} = \text{Rot}(X, \theta_s) \text{Trans}(a, b, 0) \text{Rot}(Z, -\alpha) \begin{bmatrix} x_{i0} \\ y_{i0} + \ell_i \\ z_{i0} \\ 1 \end{bmatrix} \quad (1)$$

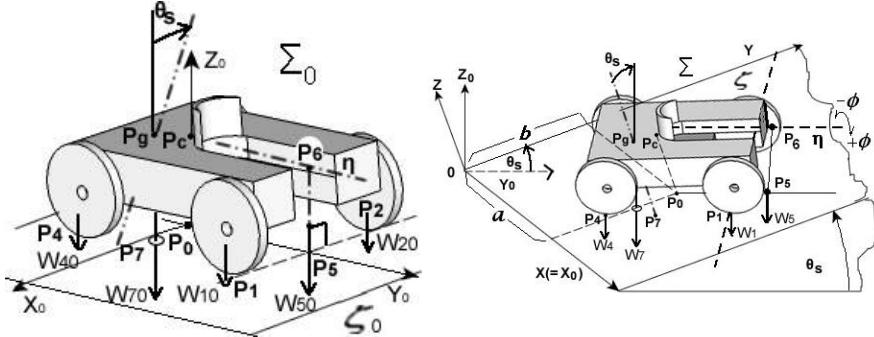
Similarly, after each wheel rolls independently on the plane  $\zeta$ ,  $P'_i(x'_i, y'_i, z'_i)$  is combined with  $P_{i0}(x_{i0}, y_{i0}, z_{i0})$  using the variables  $(a_e, b_e)$  and  $\alpha_e$ . oreover,  $z'_{i0} - z_{i0} = 0$  holds for the same plane  $\zeta_0$ . Accordingly,

$$x'_i = x_{i0} \cos(\alpha + \alpha_e) + (y_{i0} + \ell_i) \sin(\alpha + \alpha_e) + a + a_e \quad (2)$$

$$y'_i = \{-x_{i0} \sin(\alpha + \alpha_e) + (y_{i0} + \ell_i) \cos(\alpha + \alpha_e) + b + b_e\} \cos \theta_s - z_{i0} \sin \theta_s \quad (3)$$

$$z'_i = \{-x_{i0} \sin(\alpha + \alpha_e) + (y_{i0} + \ell_i) \cos(\alpha + \alpha_e) + b + b_e\} \sin \theta_s - z_{i0} \cos \theta_s \quad (4)$$

Concerning the weight distribution  $W_i$ , it is difficult to determine the load at each wheel tread uniquely because, in general, all four wheels are in contact with the ground, our robot model is compliant, resulting in four wheels always being



**Fig. 5.** Robot standing on the plane  $\zeta_0$  (left) is transferred to the position on the plane  $\zeta$  (right) after a coordinate transformation with the parameters  $(a, b)$  and  $\alpha$ .

in contact with the ground, we assume that the front and rear parts of the robot body are able to twist freely around the  $\eta$ -axis that extends straight forward from  $P_c$  in the robot body as a front pivot suspension (Fig. 5). We consider the line  $\overline{P_5P_6}$  as making a virtual wheel touch on the ground at  $P_5$  so that  $W_t$  is shared among the three points  $P_3$ ,  $P_4$  and  $P_5$ . Then, we distribute the weight  $W_5$  at  $P_5$  between  $W_1$  and  $W_2$  at  $P_1$  and  $P_2$ , respectively. This is our method for distributing the instantaneous weight to all four wheels in contact with the ground.

Actually, a static analysis introduces the following equilibrium relationships:

$$W_t = W_3 + W_4 + W_5 \quad (5)$$

$$W_5 = W_1 + W_2 \quad (6)$$

$$0 = \overrightarrow{P_gP_6} \times W_5 + \overrightarrow{P_gP_3} \times W_3 + \overrightarrow{P_gP_4} \times W_4 \quad (7)$$

$$0 = \chi \overrightarrow{P_6P_1} \times W_1 + (1 - \chi) \overrightarrow{P_6P_2} \times W_2 \quad (8)$$

where the parameter  $\chi$  is defined as a coefficient related to  $\pm\phi$  (see Fig. 5). This is important when the ground surface is not flat and the load balance between  $W_1$  and  $W_2$  varies.

In fact, the robot body compliance determines  $\pm\phi$  on uneven ground. Therefore, we write  $\chi = 0.5 + K_0\phi$ , where  $K_0$  is the twisting compliance factor of the front body. Note that  $\chi$  varies in the range  $0 < \chi < 1$ , and takes the value 0.5 when the body can twist freely, i.e., when  $K_0 = 0$ . By solving the independent equations (5)–(8) having 7 unknown variables, we can determine the position  $P_5$  and weights  $W_1 – W_5$  in the state described by  $(a, b)$  and  $\alpha$  on the plane  $\zeta$ .

## 4.2 Formulation of an Energy Cost Function Using a Virtual Work

Suppose that some wheels are driven independently. Then, skidding and slipping occur while the robot moves. Each wheel movement is composed of positional and angular displacements, which we call *linear sliding* and *twist sliding*. Drive

power is needed to accomplish sliding, which is a combination of skidding and slipping. In addition, energy is necessary when the robot rolls up a slope. When rolling down a slope, the robot moves downward by consuming its gravitational potential energy. Therefore, the virtual work for sliding appears in the form of energy for twist sliding, linear sliding and any change in gravitational potential energy [10]. Let us suppose that  $E_1$ ,  $E_2$  and  $E_3$  denote each of the three energy costs, respectively. Then these can be expressed by referring to Fig.4.

Suppose that each wheel in contact with the ground with a rectangular tread, and that the pressure in the area is unique, irrespective of the contact position. That is,  $W_i = t_{ai}t_{bi}\tau$ , where  $\tau$  is a pressure in units of pressure. Using  $\alpha_e$ , we obtain two expressions:

$$\begin{aligned} E_{1i} &= \int_0^{\alpha_i} \int_{-t_b/2}^{t_b/2} \int_{-t_a/2}^{t_a/2} \mu_{di}\tau (x^2 + y^2) dx dy d\alpha \\ &= \mu_{di} |\alpha_e| W_i \cos \theta_s (t_{ai}^2 + t_{bi}^2)/12 \end{aligned} \quad (9)$$

$$E_{2i} = \mu_{di} U_i c |\nu_i| W_i \cos \theta_s \quad (10)$$

It follows that

$$U_i c |\nu| = U_i c |U_i/\Delta t| = U_i^2 c / \Delta t \equiv K U_i^2, \quad (11)$$

where

$$\begin{aligned} U_i^2 &= (x'_i - x_{si})^2 + (y'_i - y_{si})^2 + (z'_i - z_{si})^2 \\ &= a_e^2 + C_{1i}a_e \cos \alpha_e + C_{2i}a_e \sin \alpha_e + C_{3i}a_e + b_e^2 + C_{4i}b_e \cos \alpha_e \\ &\quad + C_{5i}b_e \sin \alpha_e + C_{6i}b_e + C_{7i} \cos \alpha_e + C_{8i} \sin \alpha_e + C_{9i} \end{aligned} \quad (12)$$

The constants  $C_{1i} - C_{9i}$  are defined as

$$\begin{array}{ll} C_{1i} = 2(x_{i0} \cos \alpha + y_{i0} \sin \alpha) & C_{2i} = -2(x_{i0} \sin \alpha - y_{i0} \cos \alpha) \\ C_{3i} = -2 \{ x_{i0} \cos \alpha + (y_{i0} + \ell_i) \sin \alpha \} & C_{4i} = C_{2i} \quad C_{5i} = -C_{1i} \\ C_{6i} = -C_{2i} - 2\ell_i \cos \alpha & C_{7i} = (C_{1i}C_{3i} + C_{2i}C_{6i})/2 \\ C_{8i} = 2x_{i0}\ell_i & C_{9i} = 2(x_{i0}^2 + y_{i0}^2) + \ell_i(2y_{i0} + \ell_i) \end{array}$$

Using another constant  $K$  simplifies expression (10) as follows:

$$E_{2i} = W_i \mu_{di} K U_i^2 \cos \theta_s \quad (13)$$

So far in our analysis,  $U_i^2$  is assumed to be zero when the  $i$ -th wheel has no power. This free wheel is supposed to roll a distance  $\ell_j$  ( $j \neq i$ ) that is generated by a neighboring powered wheel since the robot maintains its body frame nearly rigid except while twisting. However, there is some force feedback caused by the rotational friction around each wheel axis in an actual machine. Therefore, we assume that  $U_i^2$  has a certain value proportional to the friction force, and express  $U_i^2$  as a small value by defining a coefficient  $\delta_i$  that is nearly equal to zero. For

a powered wheel,  $\delta_i$  is equal to 1. This consideration generalizes equation (13) as follows:

$$E_{2i} = \delta_i W_i \mu_{di} K U_i^2 \cos \theta_s \quad (14)$$

The exact value of  $\mu_{di}$  is difficult to predict because it is subject to constant change in an actual environment. However, assuming there is a certain constant value, say  $e$ , then  $\mu_{di}$  in equations (9), (10), (13) and (14) can be expressed as in (11), such that

$$\mu_{di} = (1 - e \nu_i) \mu_{si} \quad (15)$$

In particular,  $\nu_i$  as well as  $d_i$  are both small quantities in our analysis. Thus,  $\mu_{di}$  is assumed to be a smaller constant than that of  $\mu_{si}$ . This implies that a faster motion will be estimated using the small value of  $\mu_i$ . Henceforth we use  $\mu_i$  in place of  $\mu_{di}$  without the risk of underestimating motion.

Concerning the gravitational potential energy, it is evaluated by the following expression:

$$\begin{aligned} E_{3i} &= W_i(z'_i - z_i) \\ &= W_i \sin \theta_s (C_{2i} \cos \alpha_e - C_{1i} \sin \alpha_e + 2b_e - C_{2i})/2 \end{aligned} \quad (16)$$

Finally, we can sum each energy cost to yield an expression for the total cost.

$$\begin{aligned} E_t &= \sum_{i=1}^4 \{E_{1i} + E_{2i} + E_{3i}\} \\ &= \sum \left\{ \left\{ (t_{ai}^2 + t_{bi}^2) / 12 \mu_i |\alpha_e| \cos \theta_s + \mu_i \delta_i K U_i^2 \cos \theta_s + (z'_i - z_i) \right\} W_i \right\} \\ &= a_e (\lambda_1 a_e + \lambda_2 \cos \alpha_e + \lambda_3 \sin \alpha_e + \lambda_4) + b_e (\lambda_5 b_e + \lambda_6 \cos \alpha_e \\ &\quad + \lambda_7 \sin \alpha_e + \lambda_8) + \lambda_9 \cos \alpha_e + \lambda_{10} \sin \alpha_e + \lambda_{11} + \lambda_{12} |\alpha_e| \end{aligned} \quad (17)$$

where

$$\begin{aligned} \lambda_1 &= K \cos \theta_s \sum_{i=1}^4 \{\delta_i W_i \mu_i\}, & \lambda_2 &= K \cos \theta_s \sum \{\delta_i W_i \mu_i C_{1i}\} \\ \lambda_3 &= K \cos \theta_s \sum \{\delta_i W_i \mu_i C_{2i}\}, & \lambda_4 &= K \cos \theta_s \sum \{\delta_i W_i \mu_i C_{3i}\} \\ \lambda_5 &= \lambda_1, & \lambda_6 &= \lambda_3, & \lambda_7 &= -\lambda_2 \\ \lambda_8 &= \sum [W_i \{\delta_i K \mu_i \cos \theta_s C_{6i} + \sin \theta_s\}] \\ \lambda_9 &= \sum [W_i \{\delta_i K \mu_i \cos \theta_s (C_{1i} C_{3i} + C_{2i} C_{6i}) + C_{2i} \sin \theta_s\}/2] \\ \lambda_{10} &= \sum [W_i \{\delta_i K \mu_i \cos \theta_s C_{8i} - C_{1i} \sin \theta_s/2\}] \\ \lambda_{11} &= \sum [W_i \{\delta_i K \mu_i \cos \theta_s C_{9i} + C_{6i} \sin \theta_s/2\}] \\ \lambda_{12} &= \sum \{W_i \mu_i \cos \theta_s (t_{ai}^2 + t_{bi}^2)/12\} \end{aligned}$$

### 4.3 Minimization of Virtual Work

The robot movement can be predicted by solving for the unknown parameters ( $a_e$ ,  $b_e$ ) and  $\alpha_e$ . Thus, we differentiate (17) by each of these parameters since,

based on the principle of virtual work, (17) is minimized when there is no energy loss except for skidding and slipping.

$$dE_t/da_e = 2\lambda_1 a_e + \lambda_2 \cos \alpha_e + \lambda_3 \sin \alpha_e + \lambda_4 = 0 \quad (18)$$

$$dE_t/db_e = 2\lambda_5 b_e + \lambda_6 \cos \alpha_e + \lambda_7 \sin \alpha_e + \lambda_8 = 0 \quad (19)$$

$$\begin{aligned} dE_t/d\alpha_e = & -(\lambda_2 \sin \alpha_e - \lambda_3 \cos \alpha_e)a_e - (\lambda_6 \sin \alpha_e - \lambda_7 \cos \alpha_e)b_e \\ & -(\lambda_9 \sin \alpha_e - \lambda_{10} \cos \alpha_e - \lambda_{12}) = 0 \end{aligned} \quad (20)$$

Substituting  $a_e$  in (18) and  $b_e$  in (19) into (20) generates a single equation with only one unknown parameter  $\alpha_e$ . However, it is difficult to solve the resulting fourth order equation for  $\alpha_e$ , algebraically. Therefore, we assume that  $\alpha_e \approx 0$  so that  $\sin \alpha_e \simeq \alpha_e$ ,  $\cos \alpha_e \simeq 1$  and  $\alpha_e^2 \simeq 0$ . This reasonable approximation simplifies the solution such that

$$a_e = -(\lambda_2 \cos \alpha_e + \lambda_3 \sin \alpha_e + \lambda_4)/(2\lambda_1) \quad (21)$$

$$b_e = -(\lambda_6 \cos \alpha_e + \lambda_7 \sin \alpha_e + \lambda_8)/(2\lambda_5) \quad (22)$$

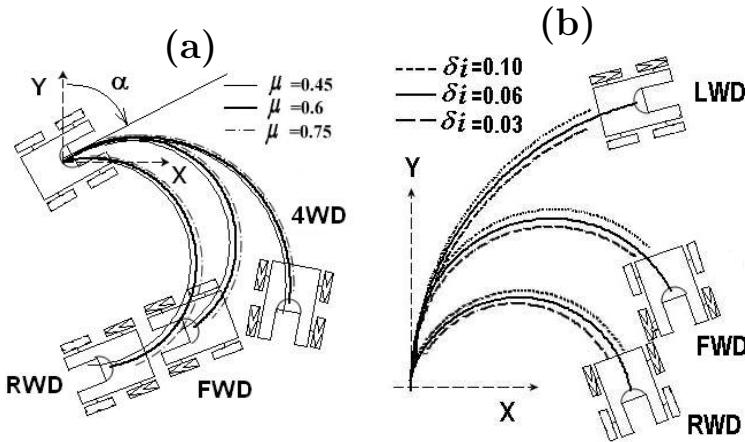
$$\alpha_e = \frac{\lambda_1 \{\lambda_7 \lambda_{68} - 2\lambda_5 (\lambda_{10} + \lambda_{12})\} + \lambda_3 \lambda_5 \lambda_{24}}{\lambda_1 \{\lambda_6 \lambda_{68} - 2\lambda_5 \lambda_9 - \lambda_7^2\} + \lambda_2 \lambda_5 \lambda_{24} - \lambda_3^2 \lambda_5} \quad (23)$$

where  $\lambda_{24}$  and  $\lambda_{68}$  denote  $\lambda_2 + \lambda_4$  and  $\lambda_6 + \lambda_8$ , respectively. Note that  $\lambda_{11}$  is a constant and disappears from (21)–(23). The next robot position is introduced repeatedly by replacing the old parameters ( $a$ ,  $b$ ) and  $\alpha$  with the new ones ( $a+a_e$ ,  $b+b_e$ ) and  $\alpha + \alpha_e$  to connect position histories in order to form a continuous trajectory.

## 5 Algorithm for Estimating a Robot's Trajectory

The robot's motion is simulated by evaluating each step in the following procedure.

1. Assign the robot's dimensional specifications  $L_h$ ,  $L_w$ ,  $R_i$ ,  $W_t$ ,  $A_i$ ,  $\mu_i$ ,  $\eta$ ,  $\theta_s$ ,  $\omega_i$  and  $\Delta\theta_r$ . Values for each of these parameters are given in the experiment discussed in Section 6. In normal conditions, the constants  $K$  and  $\chi$  are 1.0 and 0.5, respectively, however, they are not limited to these values depending on the robot's specifications and unevenness of the ground.
2. Locate the robot on the home ground to assign  $P_{i0}(i=1-4)$  on the plane  $\zeta_0$ . Then assign  $\ell_i(i=1-4)$  and apply the coordinate transformation using the initial parameters  $a$ ,  $b$ ,  $\alpha$  and  $\ell_i(i=1-4)$  to calculate  $P_{si}(i=1-4)$ .
3. Similarly, locate the robot on the home ground again and apply the coordinate transformation using the parameters  $(a+a_e, b+b_e)$  and  $\alpha + \alpha_e$  to calculate the next incremental position  $P_{si}(i=1-4)$  on the plane  $\zeta$ . If the  $i$ -th wheel is free, then assign a value for  $\delta_i$  by considering the wheel resistance, and also assign a value for  $\ell_i$  equal to that of the driving wheel of the same side on the front or rear. If all the wheels on one side are free, then use the value of  $\ell_i$  on the other side for the front or rear. Set each  $\ell_i$  to be the same value as that of the powered wheel if only one wheel is powered.



**Fig. 6.** (a) Right turn trajectories influenced by  $\mu_i$  on horizontal ground within  $\alpha = 60^\circ$ . (b) Changes in trajectory depending on the free-wheel resistance  $\delta_i$  for FWD, RWD and LWD configurations on horizontal ground.

4. Calculate the wheel load  $W_i(i = 1-4)$  through the weight  $W_5$  which is the intrinsic weight operating at the virtual wheel tread at  $P_5$ .
5. Calculate the parameters  $C_{1i} - C_{9i}$  to determine the square of the virtual distance  $U_i^2$ .
6. Calculate the parameters  $\lambda_1 - \lambda_{10}$ ,  $\lambda_{12}$ ,  $\lambda_{24}$  and  $\lambda_{68}$  to evaluate  $E_t$ .
7. Solve the equation to obtain  $(a_e, b_e)$  and  $\alpha_e$ , and determine the position  $P'_i(i = 1-4)$  on the plane  $\zeta$  for the next desired position after incrementing  $\theta_r$  with  $\Delta\theta_r$ .
8. Save the incremental position as a history along the trajectory. Then repeat the procedure from step 2. by replacing the parameters  $(a, b)$  and  $\alpha$ , with  $(a+a_e, b+b_e)$  and  $\alpha + \alpha_e$ , respectively.

## 6 Motion Simulation Using Generated Trajectories

Traversing a straight path, steering by turning and pinning are demonstrated and compared while studying the effects of the following factors with each of the powered wheel arrangements shown in Fig. II

- Ground friction coefficient
- Powered wheel arrangement
- Free – wheel resistance
- Wheel tread area
- Shifting of robot's center of mass

Let us suppose that the robot turns with  $0.45 \leq \mu_i \leq 0.75$  and  $\lambda_v = 2.0$ . The coefficients for assigning the resistance in FWD and RWD are  $\delta_3 = \delta_4 = 0$  and  $\delta_1 = \delta_2 = 0$ , respectively. Then the circular trajectories for 4WD, FWD, and RWD configurations are shown in Fig. 6(a). The starting direction  $\alpha$  is  $60^\circ$  and  $0^\circ$ , in (a) and (b), respectively. In each case,  $\mu_i$  and tread are 0.6 and

**Table 1.** Specifications of the GAIA-1a mobile robot

Symbol	Property value
$L_h$	53cm(overall), 48cm(between axes)
$L_w$	36cm(overall), 31cm(between wheels)
Wheel	front(solid); radius = 8.7cm, rear(nylon); radius = 10.8cm
Tread	front; $t_a = 7.5\text{cm}$ , $t_b = 3.0\text{cm}$ , rear; $t_a = 8.0\text{cm}$ , $t_b = 4.5\text{cm}$
$\mu_i$	$p - \text{tile}$ ; 0.44(front), 0.51(rear). $\text{wood}$ ; 0.56(front), 0.79(rear)
Driver	Vesta board, Battery(24Ah, 12V), 100ppr, Max. speed; 2.8km/h
Sensor	Encoder; 100 ppr, Marker position; 32.5cm( $P_{mf}$ ), - 28cm( $P_{mb}$ )

$(t_{ai}, t_{bi}) = (1\text{cm}, 2\text{cm})$ , respectively. The scale is 1/44.2. For clarity, only the final state of the robot is sketched. It is clear that the direction of the robot shifts away from starting direction in each case.

To study how the trajectories are influenced by the free wheel resistance, we simulate the motion for FWD ( $\lambda_v = 2$ ), RWD ( $\lambda_v = 2$ ) and LWD by assigning different values for  $\delta_i$  in (17). In particular,  $\delta_3 = \delta_4$  in FWD,  $\delta_1 = \delta_2$  in RWD and  $\delta_1 = \delta_4$  in LWD are assigned to be 0.03, 0.06 and 0.10, respectively. Fig.6(b) shows the simulated trajectories. All the left-side-powered angular velocities are equal in the three arrangements. Other conditions are maintained constant while turning, that is, for instance,  $\mu_i = 0.6$ ,  $(t_{ai}, t_{bi}) = (1\text{cm}, 2\text{cm})$  and  $\chi = 0.5$  are maintained. The results clearly show that each trajectory has a circular shape with a curvature that decreases when the resistive force increases. These results seem reasonable based on the common sense that a turn becomes more difficult with increasing resistance.

## 7 Experimental Verification

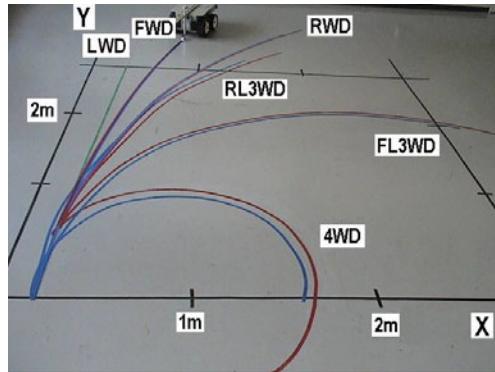
### 7.1 Experimental Setup

We prepared three test grounds for demonstrating robot motion. The first ground surface is a flat p-tile office floor. The second test ground is a flat wooden floor that is constructed by laying pieces of plywood of size 180 cm  $\times$  90 cm  $\times$  1.2 cm tightly on top of the p-tile floor. The third test ground is an inclined floor made of wood with a flat area of  $2.7 \times 3.6 \text{ m}^2$  inclined by  $\theta_s = 6^\circ$ . We refer to the three test floors as the *p-tile floor*, *wooden floor* and *wooden slope*. We affixed 2-cm wide black tape to make 1m  $\times$  1m grids on each floor.

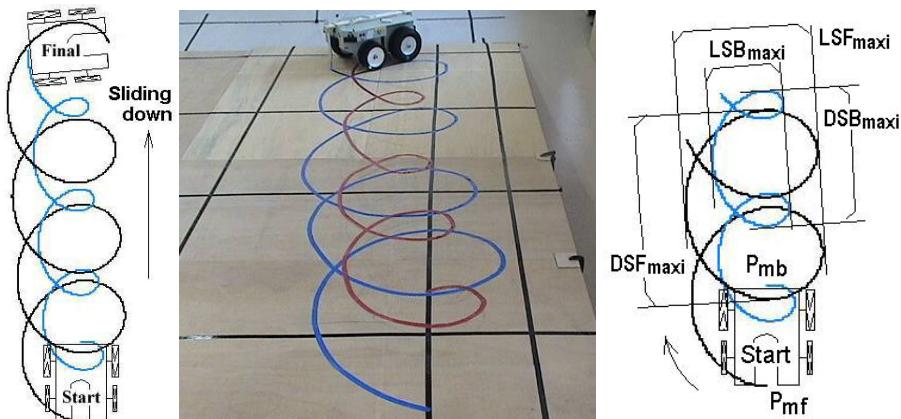
For these experiments, we used a commercially available robot, the GAIA-1a whose specifications are presented in Table 1. Note that the friction coefficients are dependent on the combination of the material of the floor and that of each wheel.

### 7.2 Turning

Drawn as the sets of  $P_{mf}$  and  $P_{mb}$ , Fig.7 shows the trajectories produced using 4WD, 3WD and 2WD powered driving arrangements on the p-tile floor. While turning right, the robot faces the centrifugal direction in RWD and RL3WD,



**Fig. 7.** Trajectories of  $P_{mf}$  and  $P_{mb}$  traced by the GAIA-1a robot during turning under 4WD, FWD, RWD, FL3WD, RL3WD and LWD configurations with  $\alpha = 0^\circ$



**Fig. 8.** Simulated and experimental trajectories of spin motion from the high-leveled bottom to low-leveled top by the GAIA-1a. Simulation results, experimental results and definition of measurement variables are shown on the left, center, and right, respectively.

but the centripetal direction in 4WD, FWD and FL3WD. This results in the fact that the trajectories following  $P_{mf}$  appear inside of those following  $P_{mb}$  trajectories. By comparing the trajectories in Fig. 7 we confirm that 4WD is capable of making the tightest turn. This is also confirmed by the simulation results. However, page limitation eliminates showing the results in this paper.

### 7.3 Spinning

Spinning motion was studied by having the GAIA-1a robot move gradually down the wooden slope by 4WD. Fig. 8 shows the experimental trajectories produced by  $P_{mf}$  and  $P_{mb}$  after the motion was complete along with the simulated trajectory

**Table 2.** Comparison of simulated and experimental trajectories for the GAIA-1a robot spinning down a wooden slope.

Length(mm)	Simulated	Experimental				Error(mm)
		First	Second	Third	Average	
LSF <sub>max</sub>	816.8	760.0	843.4	805.6	803.0	+13.8
DSF <sub>max</sub>	1172.7	1191.5	1243.2	1269.2	1234.6	-61.9
LSB <sub>max</sub>	449.6	460.0	441.4	439.0	446.8	+2.8
DSB <sub>max</sub>	812.9	819.7	791.7	1100.0	903.8	-90.9

under the aforementioned specifications and the supposition that  $\chi = 0.5$ . Graphically comparing the simulation results with the experimental results indicates a very close mapping between the two. Any error between the simulation and the experimental results may be due to the surface imperfections of the wood on the slope.

To evaluate the correctness of the simulation quantitatively, however, the simulation results are compared with the experimental results by defining several variables of interest to measure as shown on the right-hand side of Fig. 8. These variables include DSF<sub>max</sub>; the maximum span within the front marker trajectory during decent, and LSF<sub>max</sub>; the maximum lateral span within the front marker trajectory. DSB<sub>max</sub> and LSB<sub>max</sub> are similarly defined for the back marker trajectory. These four measurements are taken three times consecutively in the experiment. Table 2 shows the simulated and experimental data for these measurements. From the table, it is clear that descending span is larger than the lateral span in both the simulated and experimental results. Positional error between the experimental and simulation data is less than 10 cm and it appears as stretching the descending span in the experimental data as compared to a smaller simulated span. This discrepancy may be caused by a decrease in the value of  $\mu_i$  in the direction of decent, which would tend to stretch the span. In other words, the robot gains linear velocity as it traverse down the slope due to acceleration of the robot that is not unaccounted in our simulation model because the value of  $\mu_i$  is constant.

## 8 Conclusion

We proposed a trajectory estimation method for four-wheeled SSMRs and SSVs configured with a combination of powered and non-powered wheels. A pivot suspension structure made it possible for all the wheels to be in contact with uneven ground by partially twisting the robot body. We formulated an energy cost function for expressing the effect of skidding and slipping including the change in gravitational potential energy. We specifically focused on different wheel arrangements, the resistive force of non-powered (i.e., free) wheels, the influence of the friction coefficient of the ground surface, the instantaneous load distribution to each wheel, the wheel tread area, and the slope of the ground. Minimizing the sum of the energy costs using the principle of virtual work resulted in an iterative solution for a continuous trajectory. We successfully depicted continuous trajectories as

simulated graphical plots for cases in which a robot moves straight forward/backward, turns and spins.

To verify the simulation results, we conducted experiments with GAIA-1a robot. Measured trajectories from having the robot execute motions on three different types of floors showed excellent agreement with the simulated trajectories both size and shape.

In future work, we intend to develop a self-navigation system that determines the exact value of  $\lambda_\nu$ , automatically by considering environmental ground properties using sensor data (for instance, slope angle as determined by a two-axis gravity sensor) so that a driver can maneuver using a joystick in an outdoor environment with the same feeling as in an indoor environment.

## References

- [1] Martinez, J.L., Mandow, A., Morales, J., Pedraza, S., Garcia-Cerezo, A.: Approximating kinematics for tracked mobile robots. *Int. J. of Robotics Research* 24(10), 867–878 (2005)
- [2] Kozlowski, K., Pazderski, D.: Practical stabilization of a skid-steering mobile robot – A kinematic-based approach. In: Proc. IEEE 3rd Int. Conf. on Mechatronics, pp. 519–524 (2006)
- [3] Wang, D., Low, C.B.: Modeling and analysis of skidding and slipping in wheeled mobile robots: Control design perspective. *J. IEEE Trans. on Automatic Control* 24(3), 676–687 (2008)
- [4] Shuang, G., Cheung, N., Cheng, E., Lei, D., Xiaozhong, L.: Skid steering in 4WD EV. In: Proc. of 4th Int. Workshop on Robot and Motion Control, pp. 175–180 (2004)
- [5] Low, C.B., Wang, D.: Integrated estimation for wheeled mobile robot posture, velocities, and wheel skidding perturbations. In: IROS 2008 Workshop on Modeling, Estimation, Path Planning and Control of All Terrain Mobile Robots, pp. 20–30 (2008)
- [6] Pazderski, D., Kozlowski, K.: Trajectory tracking control of Skid-Steering Robot - experimental validation. In: Proc. of 17th World Congress, IFAC, pp. 5377–5382 (2008)
- [7] Maclaurin, B.: Comparing the steering performances of skid- and Ackermann-steered vehicles. *J. of Automobile Engineering* 222(5), 739–756 (2008)
- [8] Shamah, B.: Experimental comparison of skid steering vs. explicit steering for a wheeled mobile robot, Report of the Robotics Institute, Carnegie Mellon Univ, CMU-RI-TR-99-06, pp. 1–54 (1999)
- [9] Okada, T., Mahmoud, A., Botelho, W., Shimizu, T.: "Trajectory analysis of an independently driven wheeled robot and its experimental verification. In: Proc. 12th Int. Conf. on CLAWAR, pp. 781–790 (2009)
- [10] Hadwich, V., Pfeiffer, F.: The principle of virtual work in mechanical nad electromechanical systems. *Archive of Applied Mechanics* 65, 390–400 (1995)
- [11] Sakai, H.: Friction and wear of tire. *Tire Science and Technology (TSTCA)* 25(1) (1997)

# Kinematics Modeling and Analyses of an Articulated Robot

Zhang Yang, Liu Xinhui, and Wang Tongjian

The Institute of Mechanical Science and Engineering, Jilin University,

Changchun, China

zy2659@163.com, liuxh@jlu.edu.cn, wtj@jlu.edu.cn

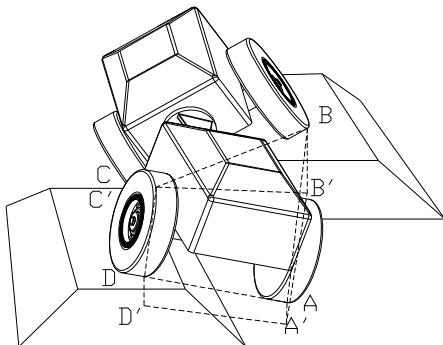
**Abstract.** In this paper, we propose a mathematical model for the kinematics of an articulated robot, called 2DWAR. The wheeled robot has two bodies joined by a 2 DOF hinge assembly. This robot can reshape its articulated frame to perform the excellent mobile stability when it runs in the rough-terrain. For the unique 2 DOF articulated frame, we propose the new concept of correlative stability. Based on the proposed model, the correlative stability angle of the robot is developed using multi-body kinematics and dynamic energy stability method. Equations are implemented in an algorithm used to study the kinematics of the articulated robot.

**Keywords:** 2 DOF articulated frame, correlative stability, multi-body kinematics, dynamic energy stability.

## 1 Introduction

Mobile robots mostly run in the rough terrain. The rough environment will require the mobile robots to be more flexible, environment adaptable and responsive. However, flexible mobility in rough terrain is limited to the locomotion concepts. There are two concepts, passive and active locomotion, in the classification of the locomotion concepts. A passive locomotion robot is based on passive frames which follow the terrain automatically without any sensors or actuators. An active robot needs a closed loop control to keep its stability. The latter is more complex and needs extended control than the former. In this paper, we present an innovative passive frame for rough terrain. The mobile robot named 2DWAR has the 2 DOF articulated frame, four-wheel drive and articulated steering. With reshaping its articulated frame, 2DWAR could adapt the terrain automatically, achieve the excellent rough-terrain mobility eventually.

Mobile stability becomes the most important factor to demonstrate the rough-terrain mobility. 2DWAR could reposition its centre of mass to improve stability on the rough terrain. For example, when traveling on the terrain 2DWAR can rotate one body of the frame to increase the tire-road contact points, get much more stability margin, see Fig. 1.

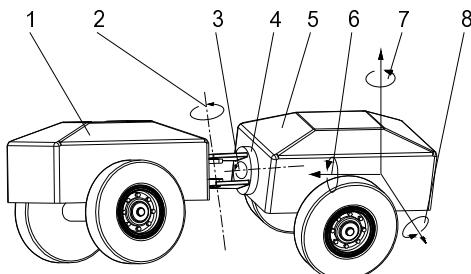


**Fig. 1.** 2DWAR travels on the rough terrain

However, all the tire-road contact points of 2DWAR do not appear on the same plane when running in the terrain. The motion traces of bodies are different. The entire frame stability has been changing all the time. This complex situation happens because of the unique frame and the bodies' connection that 2DWAR owns. The 2DWAR frame (fig.2) consists of two bodies B and F, which joined by a 2 DOF hinge assembly Q. The robot could rotate along two different directions. The steer angle  $\theta_2$  driven by cylinders indicates the angle between two bodies. The rotary angle  $\theta_1$  follows the terrain to obtain more contact points.

Even if one of the bodies loses its stability, its lost doesn't bring huge influence to the entire frame stability. The instability body keeps rotating separately until it knocks on the block of the hinge assembly. And then, all the bodies move together at the same state. Obviously, the tip over stability of the entire frame is determined by the state of the stable one before they rotate together.

For this, a new concept, called correlative stability, is proposed to describe the special situation of the unique 2 DOF articulated frame. It indicates that, the stability of single body and the hinge state between the bodies, both work on the entire frame stability.



**Fig. 2.** 2DWAR with 1.front body F 2.steering angle  $\theta_2$  3. rotary angle  $\theta_1$  4. hinge assembly Q 5.back body B 6. roll angle  $\gamma$  7. yaw angle  $\psi$  8. pitch angle  $\varphi$ .

## 2 Stability Analysis

According to the definition, the stability angle of entire frame  $\lambda_s$ , the front one  $\lambda_f$  and the back one  $\lambda_b$  are affected by a few factors, such as body pitch angle  $\varphi$ , yaw angle  $\psi$ , roll angle  $\gamma$ , rotary angle  $\theta_1$ , steering angle  $\theta_2$  between the two bodies, the mobile velocity, angular velocity and the gravitational force. Here, the variable factors will bring different results which will be analyzed as follows.

### 2.1 Theoretical Background

In this work, we assume that the state of rigid bodies of the robot system can be described by a generalized position. The back body coordinate is the reference coordinate. And it has a slope angle of  $\sigma_b^t$  at the beginning. The front body tips over meanwhile the back one is still in stable state. The angles are in ascending order in an anticlockwise manner when viewed from back. The front body rotates along the rotation axis of the hinge assembly when it is turning over. The rotary angle  $\theta_1$  is limited in  $|\pm\theta_1| \leq \theta_0$ ,  $\max\{\lambda_f, \lambda_b\} \leq \theta_0$ . The maximum stability angle of front body is defined as  $\lambda_f^t$ , so is the back one  $\lambda_b^t$ .

### 2.2 Analysis for Tip-Over Stability

The rotary angle between front body and back body is  $\theta$  at time  $t$ . The front body could rotate clockwise or opposite with the outside force (torque) at time  $t + \Delta t$ .

Case 1: The front body rotates clockwise, (fig. 3a).

1) If  $\lambda_f^t > \sigma_b^t + \theta$ , the front body is at the maximum stability angle when it has rotated  $\lambda_f^t - \sigma_b^t - \theta$ , but the entire frame keeps stable at that moment. The front body would knock on the block of the hinge assembly and be fixed join with the back one after it rotates another  $\theta_0 - \lambda_f^t + \sigma_b^t$ . And then all of the bodies rotate  $\Delta\lambda^t$  continuously at the same speed till they lose the entire stability.

2) If  $\lambda_f^t \leq \sigma_b^t + \theta$ , the front body has lost its stability for a while. It has rotated  $\sigma_b^t + \theta - \lambda_f^t$  since it lost. All of the bodies would be in move-together state after the front rotates another  $\theta_0 - \theta$ . They rotate  $\Delta\lambda^t$  continuously till entire stability loses.

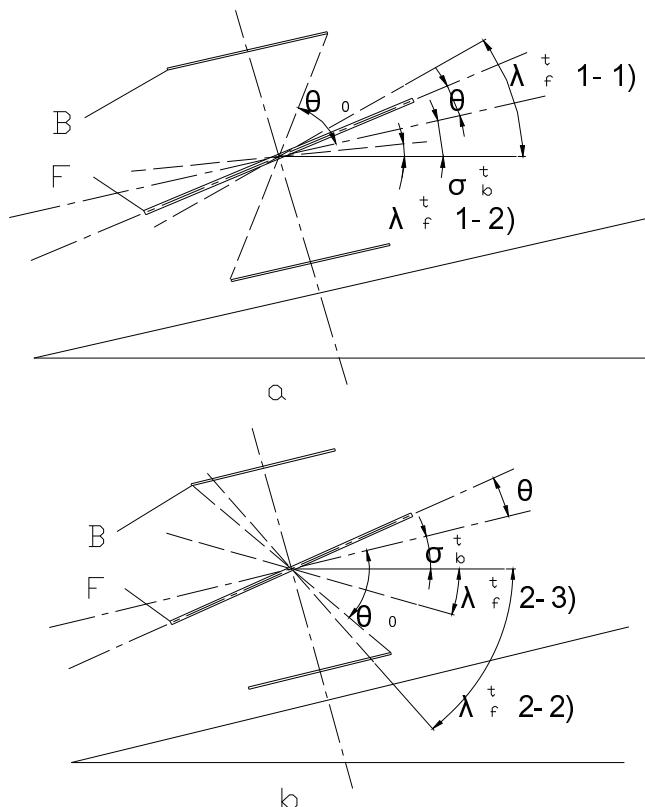
Case 2: The front body rotates anticlockwise, (fig. 3b).

It shows three different results because the front body and the back body rotate in different direction.

1) The entire frame is to keep stable.

2) If  $\lambda_f^t + \sigma_b^t > \theta_0$ , all of the bodies would be in move-together state after the front rotates  $-(\theta_0 + \theta)$ . When the entire frame loses their stability after they rotate  $\Delta\lambda^t$ , the front body hasn't lost its stability yet.

3) If  $\lambda_f^t + \sigma_b^t \leq \theta_0$ , the front body is at the maximum stability angle when it rotates  $-(\sigma_b^t + \theta + \lambda_f^t)$ . The front body would be fixed join with the back one after the front rotates another  $-(\theta_0 - \lambda_f^t - \sigma_b^t)$ . They rotate  $\Delta\lambda^t$  continuously till they lose the entire stability.



**Fig. 3.** Stability diagram

Case 3: The back body tips over meanwhile the front one is still in stable state. The former has an angle of  $\sigma_b^t$  at the beginning. The discussion about the state of the robot system by this position includes two situation considered.

1) If the back body rotates anticlockwise and  $\lambda_b^t \leq \lambda_f^t$  (the entire stability loses immediately in case of  $\lambda_b^t > \lambda_f^t$ ), we couldn't get an answer in this case. The concept of correlative stability we proposed above is based on the process from one body lost single stability to the entire lost. The back body has been instable already. The entire stability would lose as soon as the front loses.

2) If the back body rotates clockwise, the same as case 2.

Case 4: Analysis of the situations with reference based on the front body coordinate is the same as case 1 and case 2.

The stability analysis results indicate that: if one of the robot bodies loses its stability, this body could rotate continuously until it knocks on the block of the hinge assembly. After that, all the bodies would be in move-together state till they lose the entire stability. The tip-over stability of the entire frame is determined by the state of the stable one before they move together.

We define the correlative stability for the articulated robot, to demonstrate the influence that the stability of single body and the hinge state between the bodies work on the entire frame stability.

The correlative stability angle is to describe the correlative stability:

Articulated robot is moving on the rough terrain at time  $t$ . The slope angle of the entire frame refers to the maximum stability angle before the entire frame lose, which is caused by one body's stability loss. The slope angle of this body refers to its maximum stability angle. The correlative stability angle is defined as the minus between the former and the latter.

$$\varepsilon_{is}^t = (\pm\sigma_j^t) + \theta_0 + \Delta\lambda_i^t - \lambda_i^t \quad (i, j) = \begin{cases} 1, 2 \\ 2, 1 \end{cases} \quad (1)$$

Where  $\sigma_j^t$  is taken positive when all the bodies tip over in the same direction, negative for the opposite direction.  $\lambda_i^t$  denotes its maximum stability angle in case of the state  $\theta_1 = 0, \varphi, \psi, \gamma, \theta_2$  fixed.

As we know, being on the slope angle of  $\pm\sigma^t$ , the maximum stability angle of the unit frame is  $\pm\sigma^t + \Delta\lambda^t$ . There is an added quantic  $\theta_0 - \lambda^t \geq 0$  compared the correlative stability angle of the articulated frame with the unit one from (1). Since that, the articulated frame robot has more powerful ability of anti-tip over.

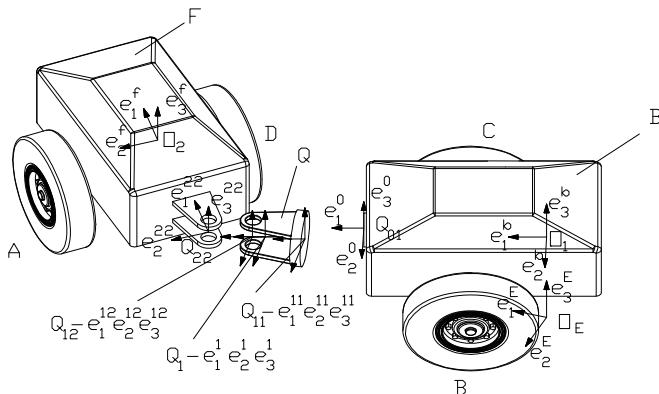
### 3 Modeling for the Correlative Stability

In this section, the kinematics properties of the 2DWAR with steered are described, when it is running on the rough terrain. The equations are based on the multi-body kinematics. We assume that the terrain, obstacles and the wheels are rigid without any

flexible. There is no skid between the wheels and the ground. We calculate in the case of the front body tips over and the back one keeps stable at the beginning.

The model consists of back body mass, back body, joints, front body and front body mass. We define the global coordinate as frame  $\underline{e}^E$ . To facilitate the location of each body, we affix local coordinate frames to each body (frame  $i$  is attached to body  $i$ ), where:

- 1) The axis  $e_3^i$  of the frame  $i$  is aligned with the revolute axis of body  $i$ .
- 2) The  $e_1^i$  axis is directed along the line from body  $i$  to next body.
- 3)  $i = b, 01, 11, q, 12, 22, f$



**Fig. 4.** Global coordinates and local coordinates

The vector of coordinates that define the global position and orientation of these bodies is:

$$\underline{q} = [\varphi \ \psi \ \gamma \ \theta_1 \ \theta_2]^T \quad (2)$$

The homogeneous transformation matrix  $A$  can be formed as the product of rotational and translational matrices:

$$\begin{aligned} \underline{A}^{Eb} = & \\ & \begin{bmatrix} \cos\varphi\cos\psi & -\cos\varphi\sin\psi\cos\gamma & -\sin\varphi\sin\gamma & \cos\varphi\sin\psi\sin\gamma & -\sin\varphi\cos\gamma \\ \sin\psi & \cos\psi\cos\gamma & 0 & -\cos\psi\sin\gamma & 0 \\ \sin\varphi\cos\psi & -\sin\varphi\sin\psi\cos\gamma & \cos\varphi\sin\gamma & \sin\varphi\sin\psi\sin\gamma & \cos\varphi\cos\gamma \end{bmatrix} \end{aligned} \quad (3)$$

$$\underline{A}^{E1} = \underline{A}^{Eb} \underline{A}^{b0} \underline{A}^{0(11)} \underline{A}^{(11)} \quad (4)$$

$$\underline{A}^{Ef} = \underline{A}^{E1} \underline{A}^{1(12)} \underline{A}^{(12)(22)} \underline{A}^{(22)f} \quad (5)$$

Where

$$\underline{A}^{0(11)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_1 & -\sin \theta_1 \\ 0 & \sin \theta_1 & \cos \theta_1 \end{bmatrix}, \underline{A}^{(12)(22)} = \begin{bmatrix} \cos \theta_2 & -\sin \theta_2 & 0 \\ \sin \theta_2 & \cos \theta_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \underline{A}^{b0} = \underline{A}^{(11)1} = \underline{A}^{1(12)} = \underline{A}^{(22)f} = I$$

The vector of the centre of mass on the bodies

$$\left. \begin{aligned} r_b &= O_E O_1 \\ r_q &= r_b + O_1 Q_{01} + Q_{11} Q_1 \\ r_f &= r_q + Q_1 Q_{12} + Q_{22} O_2 \end{aligned} \right\} \quad (6)$$

The vector of the tire-road contact points A~D , joint point Q<sub>01</sub> connect the back body and the hinge assembly

$$\left. \begin{aligned} r_A &= r_f + O_2 A \\ r_B &= r_b + O_1 B \\ r_C &= r_b + O_1 C \\ r_D &= r_f + O_2 D \\ r_{Q01} &= r_b + O_1 Q_{01} \end{aligned} \right\} \quad (7)$$

The global position of the center mass of bodies are obtained also from the product of individual transformation matrices

$${}^E \underline{r}_b = [0 \ 0 \ h_1]^T \quad (8)$$

$${}^E \underline{r}_q = {}^E \underline{r}_b + \underline{A}^{Eb} {}^b \underline{O}_1 Q_{01} + \underline{A}^{E1} {}^1 \underline{Q}_{11} Q_1 \quad (9)$$

$${}^E \underline{r}_f = {}^E \underline{r}_q + \underline{A}^{E1} {}^1 \underline{Q}_1 Q_{12} + \underline{A}^{Ef} {}^f \underline{Q}_{22} O_2 \quad (10)$$

$$M = m_b + m_q + m_f \quad (11)$$

According to the define of the center of mass,

$$Mr_s = m_b r_b + m_q r_q + m_f r_f \quad (12)$$

Next, using (7) ~ (12) , we obtain  $r_s$  the vector of the entire frame.

The global position of the local points

$$\left. \begin{aligned} {}^E\underline{r}_A &= {}^E\underline{r}_f + \underline{A}^{Ef} {}^f\underline{O}_2 A \\ {}^E\underline{r}_B &= {}^E\underline{r}_b + \underline{A}^{Eb} {}^b\underline{O}_1 B \\ {}^E\underline{r}_C &= {}^E\underline{r}_b + \underline{A}^{Eb} {}^b\underline{O}_1 C \\ {}^E\underline{r}_D &= {}^E\underline{r}_f + \underline{A}^{Ef} {}^f\underline{O}_2 D \\ {}^E\underline{r}_{Q01} &= {}^E\underline{r}_b + \underline{A}^{Eb} {}^b\underline{O}_1 Q_{01} \end{aligned} \right\} \quad (13)$$

The vector array for the global position of the center mass of bodies

$$\underline{x} = \left[ \left( {}^E\underline{r}_b \right)^T \quad \left( {}^E\underline{r}_q \right)^T \quad \left( {}^E\underline{r}_f \right)^T \right]^T \quad (14)$$

The Jacobian matrix of (14) can be calculated as

$$\underline{J}^X = \left[ \left( \underline{J}^{X1} \right)^T \quad \left( \underline{J}^{X2} \right)^T \quad \left( \underline{J}^{X3} \right)^T \right]^T \quad (15)$$

Differentiation of both sides of relations(15)

$$\dot{\underline{J}}^X = \left[ \left( \dot{\underline{J}}^{X1} \right)^T \quad \left( \dot{\underline{J}}^{X2} \right)^T \quad \left( \dot{\underline{J}}^{X3} \right)^T \right]^T \quad (16)$$

$$\dot{\underline{x}} = \underline{J}^X \dot{\underline{q}} + \partial \underline{x} / \partial t \quad (17)$$

The matrix  $\underline{\beta}$  can be formed as the product of rotational and translational matrices:

$$\underline{\beta} = \left[ \left( \underline{\beta}_b \right)^T \quad \left( \underline{\beta}_q \right)^T \quad \left( \underline{\beta}_f \right)^T \right]^T \quad (18)$$

$$\underline{\beta}_b = \left[ A_{12}^{Eb} \quad A_{22}^{Eb} \quad A_{32}^{Eb} \quad A_{13}^{Eb} \quad A_{23}^{Eb} \quad A_{33}^{Eb} \quad A_{11}^{Eb} \quad A_{21}^{Eb} \quad A_{31}^{Eb} \right]^T$$

$$\underline{\beta}_q = \left[ A_{12}^{E1} \quad A_{22}^{E1} \quad A_{32}^{E1} \quad A_{13}^{E1} \quad A_{23}^{E1} \quad A_{33}^{E1} \quad A_{11}^{E1} \quad A_{21}^{E1} \quad A_{31}^{E1} \right]^T$$

$$\underline{\beta}_f = \left[ A_{12}^{Ef} \quad A_{22}^{Ef} \quad A_{32}^{Ef} \quad A_{13}^{Ef} \quad A_{23}^{Ef} \quad A_{33}^{Ef} \quad A_{11}^{Ef} \quad A_{21}^{Ef} \quad A_{31}^{Ef} \right]^T$$

The matrix  $\underline{\alpha}$  can be formed as the product of rotational and translational matrices:

$$\underline{\alpha} = \text{diag} \left[ \underline{\alpha}_1 \quad \underline{\alpha}_2 \quad \underline{\alpha}_3 \right] \quad (19)$$

$$\begin{aligned}
\underline{\alpha}_b &= \begin{bmatrix} A_{13}^{Eb} & A_{23}^{Eb} & A_{33}^{Eb} \\ & A_{11}^{Eb} & A_{21}^{Eb} & A_{31}^{Eb} \\ & & A_{12}^{Eb} & A_{22}^{Eb} & A_{32}^{Eb} \end{bmatrix} \\
\underline{\alpha}_q &= \begin{bmatrix} A_{13}^{E1} & A_{23}^{E1} & A_{33}^{E1} \\ & A_{11}^{E1} & A_{21}^{E1} & A_{31}^{E1} \\ & & A_{12}^{E1} & A_{22}^{E1} & A_{32}^{E1} \end{bmatrix} \\
\underline{\alpha}_f &= \begin{bmatrix} A_{13}^{Ef} & A_{23}^{Ef} & A_{33}^{Ef} \\ & A_{11}^{Ef} & A_{21}^{Ef} & A_{31}^{Ef} \\ & & A_{12}^{Ef} & A_{22}^{Ef} & A_{32}^{Ef} \end{bmatrix} \\
\underline{\omega} &= \underline{\alpha} \dot{\underline{\beta}} \tag{20}
\end{aligned}$$

The mass and inertia matrix of entire frame

$$\underline{m} = diag \begin{bmatrix} \underline{m}_b & \underline{m}_q & \underline{m}_f \end{bmatrix} \tag{21}$$

where

$$\begin{aligned}
\underline{m}_b &= diag \begin{bmatrix} \underline{m}_b & \underline{m}_b & \underline{m}_b \end{bmatrix}; \underline{m}_q = diag \begin{bmatrix} \underline{m}_q & \underline{m}_q & \underline{m}_q \end{bmatrix}; \underline{m}_f = diag \begin{bmatrix} \underline{m}_f & \underline{m}_f & \underline{m}_f \end{bmatrix} \\
\underline{I} &= diag \begin{bmatrix} {}^b I^b & {}^1 I^q & {}^f I^f \end{bmatrix} \tag{22}
\end{aligned}$$

where

$$\begin{aligned}
{}^b \underline{I}^b &= diag \begin{bmatrix} I_1^b & I_2^b & I_3^b \end{bmatrix}; {}^1 \underline{I}^q = diag \begin{bmatrix} I_1^q & I_2^q & I_3^q \end{bmatrix}; {}^f \underline{I}^f = diag \begin{bmatrix} I_1^f & I_2^f & I_3^f \end{bmatrix} \\
\underline{M} &= diag \begin{bmatrix} \underline{m} & \underline{I} \end{bmatrix} \tag{23}
\end{aligned}$$

According to the definition, the dynamic energy stability of entire frame can be calculated as

$$W_s = \min \{W_{AB} - W_{BC} - W_{CD} - W_{DA}\} - \frac{1}{2} \underline{m} \dot{x}^2 - \frac{1}{2} \underline{I} \underline{\omega}^2 \tag{24}$$

Where  $W_{AB}, W_{BC}, W_{CD}, W_{DA}$  denote the gravitational increment of the robot, the lines join as the tipover axis.

The robot could keep in stable state with

$$W_s > \frac{1}{2} \underline{m} \dot{x}^2 + \frac{1}{2} \underline{I} \underline{\omega}^2$$

In the case of  $\theta_1 = 0, \varphi, \psi, \gamma, \theta_2$  fixed, the subpoint of front body mass on the plant  $O_E - e_1^E e_2^E$  is  $O_2^0$ , the subpoints of the tire-road contact points A,D and joint point

$Q_{01}$  are  $A^0, D^0, Q_{01}^0$ . When the minimum distance reduces to zero from the point  $O_2^0$  to any side of  $\Delta A^0 Q_{01}^0 D^0$ , the gravitational increment of the front body is equal to the dynamic energy, the rotary angle  $\theta_1$  is the maximum stability angle  $\lambda_f^t$ .

The subpoint of entire frame mass on the plant  $O_E - e_1^E e_2^E$  is  $O_s'$ , the subpoints of the local points are  $A', B', C', D'$ . The front body is losing its stability when the angle  $\theta_1$  rotates up to  $\lambda_f^t$  at time  $t_1$ . The rotary angle between two bodies gets maximum  $\theta_0$  at time  $t_2$ . And then, all of the bodies rotate  $\Delta\lambda^t$  continuously at the same speed. The minimum distance reduces to zero from the point  $O_s'$  to any side of quadrangle  $A' B' C' D'$ , meanwhile the gravitational increment of the entire frame is equal to the dynamic energy, the robot loses its entire stability.

Using (1) and the  $\lambda_f^t$ ,  $\Delta\lambda^t$  obtained above, we get the correlative stability angle of front body relative to the entire frame

$$\varepsilon_{fs}^t = (\pm \sigma_b^t) + \theta_0 - \lambda_f^t + \Delta\lambda^t \quad (25)$$

We could get the correlative stability angle of back body relative to the entire frame  $\varepsilon_{bs}^t$  as well.

## 4 Conclusion

This paper presents the study of the kinematics of an articulated robot, 2DWAR. This robot has two bodies joined by a 2 DOF hinge assembly. The new concepts, called correlative stability and correlative stability angle, have been presented for the unique 2 DOF articulated frame. The analyses for correlative stability and correlative stability angle describe the tip-over motion of 2DWAR. The kinematics model based on the multi-body kinematics and dynamic energy stability method is thus applicable to high mobility robots running in rough terrain. There are a number of the proposed kinematics models and analyses that we are currently investigating. There are also dynamics effects that can influence the robot mobile stability. Nevertheless, the proposed models and analyses in this paper could provide useful tools for studying behaviors of articulated robots running in the rough terrain.

**Acknowledgments.** This work is supported by the National High-Technology Research and Development Program ("863" Program) of China under contract number 2007AA04Z208.

## References

1. Song, X., Wang, Y., Tan, D., Wu, Z.: Study on Modeling and Control of All-terrain Mobile robots. *Robot* 29, 509 (2007)
2. Hong, J.: Computational Dynamics of Multibody Systems. Chinese Higher Education Press (1999)

# A Tracking Robot Based on Wireless Beacon

Zhou Fang<sup>1</sup>, Lei Deng<sup>1</sup>, Yongsheng Ou<sup>1,2</sup>, and Xinyu Wu<sup>1,2</sup>

<sup>1</sup> Shenzhen Institutes of Advanced Technology, Chinese Academy Sciences, China

<sup>2</sup> The Chinese University of Hongkong, Hong Kong, China

zhou.fang@siat.ac.cn, lei.deng@siat.ac.cn, ys.ou@sub.siat.ac.cn,  
xy.wu@siat.ac.cn

**Abstract.** This paper presents a tracking robot, which can chase its owner and go where the owner is. It adds a wireless beacon to the target to build a Radio Frequency (RF) wireless system. It uses correlative phase interferometry method based on wireless location of single receiving station to get the orientation of the target, and uses a sonar system to get distance information. Then it tracks the target according to orientation and distance information. The advantages of the system are (1) tracking in short distance; (2) robustness in target loss; (3) the simple structure. The prototype was constructed and a number of experiments were carried out, demonstrating the feasibility of the system.

## 1 Introduction

In the modern world, robots have become an integrated part of our society. Tracking robots, which have a wide potential scope of application, involve a series of problems such as path planning, map generation, automatic obstacle avoidance, tracking, localization, etc. So how to combine the target tracking method with mobile robot to achieve a complete and reliable system is also quite important.

In the past 20 years, target tracking theory and method have become a hot research area. But how to apply it to mobile robot is still a new research direction. As the mobile robot is in a dynamic and unknown environment, the background is variable and complex. It puts a higher requirement for the tracking stability. The uncertainty of the target location requires higher real-time performance.

Currently the tracking and localization methods are mostly based on image and video processing. The main methods are frames difference method, optical flow segmentation method[1] and image matching method[2][3]. However, these methods have their limitations. These methods are potential to lose targets. For example, if the target move out of the camera it is hard to find it back. These methods are very sensitive to the background as well. These methods are difficult to identify multiply targets. And the computational complexity of image and video processing is very high, which will increase the response time of these systems.

Because of such shortcomings discussed above, in this paper we use Radio Frequency (RF) wireless signal to realize the target localization and tracking, instead of a image based method. First we need a RF wireless signal, but the target doesn't emit the RF signal itself. So we add a transmitter to the target, which means the transmitter

and the target is in the same location. We just need to get the location of the transmitter. In this way we build a RF wireless system, and then we can use the wireless location methods to realize the target localization and tracking. In this paper we will present how to integrate the wireless location methods to the mobile robot to realize the tracking function.

The following sections are organized as follows. Section 2 introduces the wireless location methods and analyzes their advantages and disadvantages. In section 3, the algorithm of the system is described. The system design and the prototype are introduced in Section 4. The experiment results are summarized in Section 5. Finally, conclusions and future works are presented in Section 6.

## 2 Wireless Location Methods

The parameters of a wireless signal are amplitude, frequency and phase. According to the parameters measured the wireless location methods can be divided into based on signal strength of arrival (SSOA), based on time of arrival (TOA), based on time difference of arrival (TDOA) and based on angle of arrival (AOA).

SSOA measures the received signal strength then uses the signal attenuation model and the transmit signal strength to estimate the distance from the transmitter to the receiver. TOA is similar with SSOA. It measures the time of signal from transmitter to multiple receivers at the same time, then calculates the distance. TDOA measures the time difference of the signal reaches two different receivers. It can reduce the time synchronization requirement rather than measure the absolute time of arrival. The target is in the hyperbola with the two receivers as focus. So we can use multiple receivers to estimate its location. AOA measures the incident angle of the signal to form one orientation line from the receiver to the target. The intersection of multiple orientation lines is the target position.

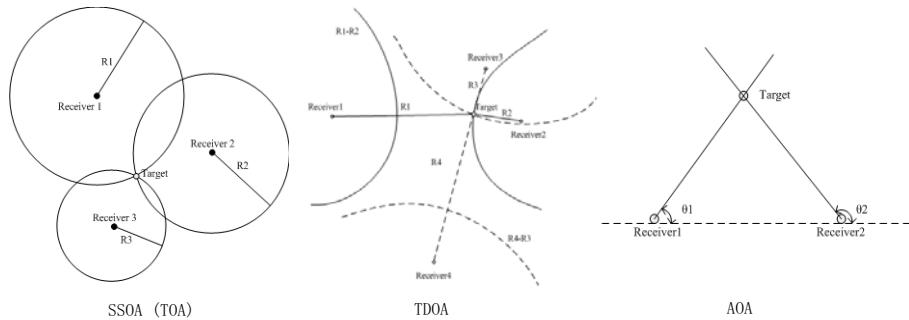
The principles of methods discussed above are illustrated in Fig. 1, respectively.

In order to determine the location of the target, the above methods need multiple receivers to receive the signal. Currently most wireless location systems are based on wireless sensor network. In order to improve the accuracy of the localization, these systems need to set a great number of wireless sensors in the localization space and integrate the signal received by the sensors to determine the location of the target. So it is generally used at indoor area and can be only used in a specific place.

However, if only one receiver is used to realize the target localization and tracking, then the overall complexity of system could be decreased. Currently some methods achieved this goal, including the amplitude method, the ray tracking method, the phase interferometry method, etc.

The amplitude method uses the antenna directivity. When the signal reaches in different angles, the signal amplitude received by the antenna will have different value. We can determine the orientation of the target by the different output of signal strength.

The ray tracking method is based on the polyhedron surface model in the location region, and tracks all rays from the transmitter to the receiver. First we theoretically calculate the received signal when the signal source at different position according to



**Fig. 1.** The Principle of SSOA(TOA), TDOA, AOA, respectively

the surrounding environment information, and then compare the actual received signal to the calculated value to determine the location of the signal source.

The phase interferometry method measures the phase difference of signal received by different antennas in the antenna array, and then identifies orientation of the target from the phase difference.

The advantages of single receiving station compare to wireless sensor network are simple structure and wide application, but the disadvantage is the location accuracy is not as good as the wireless sensor network especially in the distance measurement. So the location methods based on single receiving station is generally used in the target orientation measurement system.

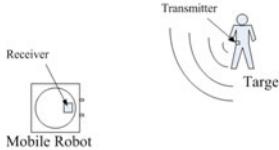
The goal of the tracking robot is to achieve stable tracking of target. Although the target location is the prerequisite of the target tracking, we only need the location accuracy to meet the requirement of tracking. So the simple structure and wide application of single station orientation is perfect for our requirement of tracking, and shortcomings of low location accuracy is acceptable. And we can use another means to reduce the distance measurement errors, in the actual design we use the ultrasonic ranging system to assist the distance measurement.

Amplitude method uses the signal strength and the directive characteristics of the antenna, So need a good directive of the antenna. The signal strength is easy to influence by the around objects which lead the effect is not very good. Ray tracking method requires prior knowledge of the surrounding environment and the computational is very complexity. The phase interference method only needs a receiver array to determine the orientation of the signal source, so it is more appropriate for the tracking robot.

### 3 Algorithm of the System

The tacking robot is based on single receiving station location method of RF wireless signal to realize the target location and tracking. The overall design of the tracking robot is shown in Fig. 2. It can be divided into two parts: the location system and the mobile robot. The location system determines the position of the tracked target. The mobile robot is motion mechanism and it tracks the target according to the position

information provided by the location system. The location system also has two parts: transmitter and receiver. Target takes the transmitter which transmits RF signal to help the receiver locate the position of target. Receiver module installed in the mobile robot, determines the location of the target through the received signal and then control the robot to track the target according the location information.



**Fig. 2.** Overall Design of the Tracking Robot

The system uses the phase interference method to identify the orientation of the target. It measures the phase difference of the signal received by different antennas in the antenna array. From the phase difference of the signal received by the two antennas in the antenna array we can calculate the distance difference from the transmitter to the two antennas according to (1).

$$\Delta R = \left( \frac{\varphi}{2\pi} * T \right) * c \quad (1)$$

Where  $\Delta R$  is the distance difference of the signal reaches the two antennas,  $\varphi$  is the phase difference of signal received by the two antennas,  $T$  is the period of the signal,  $c$  is the speed of light.

From the distance difference from the transmitter to the two antennas we can know the target is in the hyperbola with the two antennas as focus. It is similar with TDOA method described above as shown in Fig. 1.

However, TDOA method is used in the situation when the distance between two receivers is very long. As the distance of two antennas in the antenna array is much shorter than the distance from the transmitter to the receiver, we can think the signal received by the two antennas is parallel as shown in Fig. 3. The algorithm can be simplified as follow.

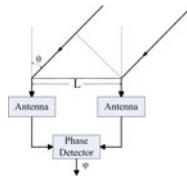
The incident angle of the signal is  $\theta$ . The phase difference of signal received by the two antennas is  $\varphi$ . The relationship between the incident angle and the phase difference can be expressed as (2).

$$\varphi = 2\pi \frac{L \sin \theta}{\lambda} \quad (2)$$

where  $L$  is the distance of two antennas,  $\lambda$  is the wavelength of the signal. So the incident angle of the signal is:

$$\theta = \arcsin \frac{\lambda \varphi}{2\pi L} \quad (3)$$

According to (3), in order to calculate  $\theta$  we need to know  $L$ ,  $\varphi$  and  $\lambda$ .  $L$  is determined by the position of the Antennas.  $\varphi$  can be detected by the phase detector.  $\lambda$  is decided



**Fig. 3.** Two antennas array

by the design of the transmitter circuit. Then the incident angle of the signal can be calculated according to (3).

The analysis above is based on two antennas. However, only using two antennas will have some shortcomings such as ambiguity problem, low measurement accuracy, etc. Therefore, the actual measurement system uses multiple receiving antennas to form the antenna array. The location principle is the same as two antennas. We can select any two antennas in the antenna array to compare the phase difference of the signal received by them. We can get multiple phase difference from different pairs of antennas. So we can get more information about the target location to make the measurement more accuracy.

We can calculate the orientation of the target after get the phase information. The direct way is to calculate the orientation of the target according to the relationship between the incident angle and the phase difference. From the antenna array we can get multiple incident angles based on different baseline, and we can integrate those incident angles to get the exact orientation of the target. Using the direct way to calculate the orientation of the target requires the signal transmission and reception are in ideal conditions. However, in the actual environment there are many factors will affect the measurement result such as noise, channel matching, antenna array mutual coupling, etc. So if we want to get accurate result, we need to add result correction to eliminate the error caused by the factors. However, it will increase the complexity of the algorithm and the error is difficult to be completely eliminated.

On the other way, we can use correlative phase interferometry method to calculate the orientation of the target. The algorithm of correlative phase interferometry is as follow. For a signal from a fixed incident angle we can get a set of phase difference from the antenna array. In the  $360^\circ$  whole range we select n equally spaced directions expressed as  $\theta_i$  ( $i=1, \dots, n$ ). In each direction we measure all the phase difference in advance expressed as  $\varphi_j$  ( $j=1, \dots, m$ ), m is the number of phase difference. All the phase difference values make up of the original phase samples. We write the original phase samples into vector form expressed as  $\vec{\theta}_i = (\varphi_{1i}, \varphi_{2i}, \dots, \varphi_{mi})$ ,  $i=1, 2, \dots, n$ . For an actual measured signal all the measured phase difference values can be written into vector form  $\vec{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_m)$ . Then we calculate the correlation between the measured values and the original phase samples according to (4). The incident angle in the original phase samples with the maximum correlation coefficient is the orientation of the target.

$$\rho_i = \frac{\vec{\theta}_i \bullet \vec{\varphi}^T}{(\vec{\theta}_i \bullet \vec{\theta}^T)^{1/2} (\vec{\varphi} \bullet \vec{\varphi}^T)^{1/2}} \quad (4)$$

The correlative phase interferometry method can weaken the errors caused by mutual coupling, carrier and so on. Although these errors are still exist, as long as they are stable, distortion caused by these factors has been stored into the original samples. So in the same incident angle the distortion on the sample values and the measurement values should be the same. We only need to find the closest sample value to the measurement value, and don't need to consider the errors. In this way we can weaken the effect on the measurement accuracy cased by the system errors. On the other hand the original samples contain the error of the circuit, so using the correlative phase interferometry method can reduce the difficulty of the hardware design.

## 4 System Design

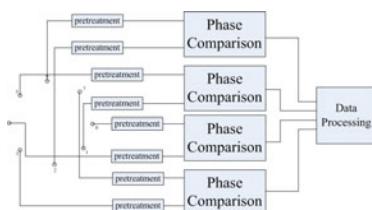
### 4.1 Transmitter Design

The transmitter is shown in Fig. 6. As the transmitter is carried by the target, the transmitter should be small, light in weight and easy to carry. So the structure of the transmitter is simple. It uses a battery for power supply. It composed by a crystal oscillator, an amplifier and an antenna.

The crystal oscillator is used to generate the appropriate signal. The frequency of the signal is decided by the crystal oscillator. The amplifier amplifies the signal to the appropriate power. The transmit power of the transmitter is decided by the performance of the amplifier. The antenna is used to transmit the signal. In order to ensure that the receiver can receive the signal the antenna should be omnidirectional antenna.

### 4.2 Receiver Design

The framework of the receiver is shown in Fig. 4. It composed by an antenna array, pretreatment modules, phase comparison modules, and a data processing module.



**Fig. 4.** Framework of Receiver

The antenna array is used to receive the signal from the transmitter. We use eight antennas to make up the antenna array, and two antennas at the opposite position make up a pair of antennas to compare the phase difference of the signal received by them. So from four pairs of antennas we can get four phase difference values.

Pretreatment module makes the signal received by the antenna fit for the phase comparison module. It includes the amplifier and the filter. The amplifier is used to amplify

the weak signal received by the antenna, and the performance of the amplifier is very important to the receiver. The filter is used to eliminate the unwanted signal received by the antenna. As each signal received by the antenna need to be amplified and filtered, there are eight pretreatment modules.

The phase comparison module is used to compare the phase difference of two signals. We use a pair of matched logarithmic amplifiers to realize the measurement. Logarithmic amplifiers provide a logarithmic compression function that converts a large range of input signal levels to a compact decibel-scaled output. The general mathematical form is:

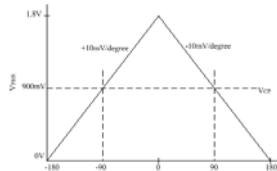
$$V_{out} = V_{slp} \log \frac{V_{in}}{V_Z} \quad (5)$$

Where  $V_{in}$  is the input voltage,  $V_Z$  is called the intercept voltage, and  $V_{slp}$  is called the slope voltage.

We use  $V_{inA}$  and  $V_{inB}$  to replace  $V_{in}$  and  $V_Z$  during the measurement, and then we can get the phase output:

$$V_{PHS} = V_\varphi [\varphi(V_{inA}) - \varphi(V_{inB})] \quad (6)$$

According to (6) we can get the phase difference from the value of output voltage. The phase transfer characteristic of the phase comparison is shown in Fig. 5. As there are four pairs of antennas need to compare the phase difference, so we have four phase comparison modules.



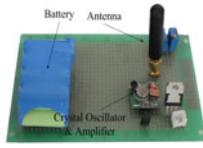
**Fig. 5.** Phase Transfer Characteristic

The data processing module is used to calculate the orientation of the target according to the correlative phase interferometry method. First we need to collect the voltage value from the phase comparison modules. There are four voltage values need to be collected, so we need a multiple channel ADC conversion. From the voltage value we can calculate the phase difference, and then we can calculate the orientation of the target according to the correlative phase interferometry method.

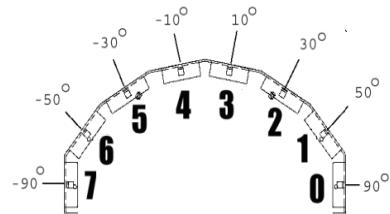
The prototype of the receiver is shown in Fig. 7.

#### 4.3 Mobile Robot

The mobile robot is shown in Fig. 8. We use Pioneer3 robot as the mobile platform. The movement of the mobile platform can be control by the program. So we can realize the tracking on the mobile platform according the algorithm.

**Fig. 6.** Prototype of Transmitter**Fig. 7.** Prototype of Receiver

There are eight sonar sensors on Pioneer3 as shown in Fig. 9. The sonar sensors use ultrasound to measure the distance from the robot to the obstacle around. Because we use the wireless location method base on single receiving station which is not good at the distance measurement. We can use the sonar system to realize the distance measurement and obstacle avoidance.

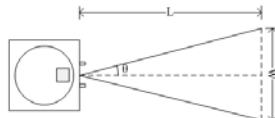
**Fig. 8.** Mobile Robot Pioneer3**Fig. 9.** Sonar System

From the location system we can get the orientation of the target. Then we get the distance data at this orientation from the sonar system. As the tracking robot follows the target at a close distance, there is no object between the robot and the target in general. So we can think the distance data at the orientation of the target is the distance between the robot and the target.

From the distance and orientation information we can control the robot to track the target. First we make the robot turn to the target according to the orientation information. Then we control the speed of robot according to the distance information. If the robot is too far to the target, we increase the speed, and if it is too near the target, we decrease the speed.

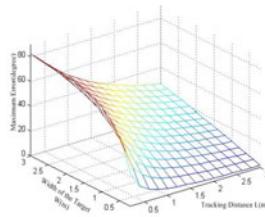
## 5 Experiments and Results

In order to evaluate the performance of the system, experiments are implemented. First we analyze the accuracy requirement of the orientation measurement for the tracking robot. The tracking robot maintains a certain distance to the target as shown in Fig. 10.



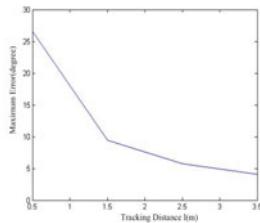
**Fig. 10.** Tracking Sketch Map

Suppose the width of the target is  $W$  and the tracking distance is  $L$ , we can get the maximum error of the orientation measurement should be less than  $\theta \leq \arctan \frac{W}{2L}$ . In different width of the target and tracking distance the maximum error of the orientation measurement is shown in Fig. 11.



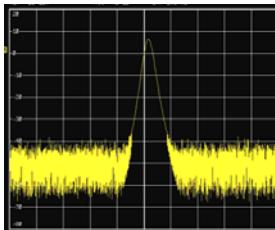
**Fig. 11.** Maximum Errors in Different Target Width and Tracking Distance

Take the human tracking as an example. The tracking distance is about 1 meter to 3 meter. The width of an adult is about half a meter. We can see that the error of the orientation measurement under  $5^\circ$  will meet the requirement of the human tracking.



**Fig. 12.** Maximum Errors in Human Tracking

We measure the transmit power of the transmitter. First we connect the spectrum analyzer to the transmitter directly. Then we measure the signal power in the space about 3 meters away from the transmitter. The display of the spectrum analyzer is shown in Fig. 13. Result shows that the transmit power is about 10dBm and the Space attenuation is about 40dB.

**Fig. 13.** Transmit Power

Then the power gain of the receiver circuit is measured. We use signal generator to generates input signal and spectrum analyzer to measure the output signal, and than calculate the power gain of the circuit. We measure the power gain of filter module, amplifier module and the whole channel in eight channels. The result is shown in table [II](#)

**Table 1.** Experiment result

	<i>Amplifier</i>	<i>Filter</i>	<i>WholeChannel</i>
Channel 1	13dB	-3dB	9dB
Channel 2	23dB	-3dB	20dB
Channel 3	12dB	-4dB	7dB
Channel 4	10dB	-2dB	8dB
Channel 5	12dB	-3dB	7dB
Channel 6	11dB	-2dB	8dB
Channel 7	8dB	-2dB	6dB
Channel 8	9dB	-2dB	7dB

At last we measure the accuracy of the location system. In the experiments the transmitter is placed in different distance and angle, and then the result is get from the receiver. The average result of repeated measurements is shown in table [II](#)

From the experiments results, we have the following conclusions:

- i. The transmit power meets the design requirement.
- ii. The power gain of the amplifier is lower then the standard value of the component. It caused a certain impact to the performance of the system.
- iii. The location system can identify the orientation of the target.
- iv. According to  $\Delta\theta_{RMS} = \sqrt{\frac{\sum(\Delta\theta_i)^2}{n}}$ , the root-mean-square error of the result is  $4.99^\circ$  which is under  $5^\circ$ . So it can be used on the tracking robot.

**Table 2.** Experiment result

	$-90^\circ$	$-60^\circ$	$-30^\circ$	$0^\circ$	$30^\circ$	$60^\circ$	$90^\circ$
0.8m	$-86.4^\circ$	$-62.6^\circ$	$-32.2^\circ$	$1.4^\circ$	$31.5^\circ$	$58.6^\circ$	$85.2^\circ$
1.6m	$-84.5^\circ$	$-64.3^\circ$	$-27.3^\circ$	$3.7^\circ$	$28.7^\circ$	$57.9^\circ$	$83.6^\circ$
2.4m	$-82.7^\circ$	$-65.2^\circ$	$-33.9^\circ$	$-4.6^\circ$	$27.6^\circ$	$56.6^\circ$	$81.7^\circ$
3.2m	$-80.6^\circ$	$-67.1^\circ$	$-34.2^\circ$	$-5.3^\circ$	$34.5^\circ$	$54.7^\circ$	$80.4^\circ$

## 6 Conclusions and Future Works

In this paper, a tracking robot is presented. It uses the wireless location method base on single receiving station to get the orientation of the target and uses the sonar system to get the distance information. The system principle and design are discussed. Experiments are carried out. Results show that the system can achieve the tracking function.

In order to simplify the system design we use 433MHz of the pure carrier signal to locate and track, so it is very easy to be jammed by the signal with the similar frequency around environment. In the future we can add the modulation and demodulation module to increase the system anti-jamming capability.

In the actual environment there are various objects will reflect the signal. So the signal received by the antennas is the combination of direct wave and reflected wave. That is multipath effect. It will affect the performance of the system. How to reduce the errors caused by the multipath effect is a part of our future work.

## Acknowledgment

The work described in this paper is partially supported by the grant from Key Laboratory of Robotics and Intelligent System, Guangdong Province(2009A060800016), and by the grant from Shenzhen-Hong Kong, Innovation Circle Special Subsidy Scheme, and by the Knowledge Innovation Program of the Chinese Academy of Sciences Grant No. KGCX2-YW-156.

## References

1. Horn, B.K.P., Schunck, B.G.: Determining optical flow. *Artificial Intelligence* 17, 185–204 (1981)
2. Badenas, J., Bober, M., Pla, F.: Motion and intensity-based segmentation and its application to traffic monitoring. In: The International Conference on Image Analysis and Proceedings, Florence, Italy, pp. 502–509 (1997)
3. Kim, W., Lee, C.Y., Lee, J.J.: Tracking moving object using Snake's jump based on image flow. *Mechatronics* 11, 119–216 (2001)

4. Liu, H., Darabi, H., Banerjee, P., Liu, J.: Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Transaction on Systems, Man, and Cybernetics* 37(6), 1067–1077 (2007)
5. Manesis, T., Avouris, N.: Survey of Position Location Techniques in Mobile Systems, PMobileHCI, Salzburg (2005)
6. Li, B., Dempster, A., Rizos, C., Barnes, J.: Hybrid method for localization using WLAN. In: *Spatial Science Conference*, Melbourne, Australia, pp. 341–350 (2005)
7. Chan, H.K., Ye, W., Lam, T.L., Ou, Y., Xu, Y.: Sensor System for a Human-Following Robot. In: *Automation Control and Applications* (2005)
8. Jin, L., Yin, Q.Y., Jiang, B.F.: Direction Finding of Wideband Signals via Spatial-Temporal Processing in Wireless Communications. In: *IEEE International Symposium on Circuits and Systems*, Geneva, Switzerland (2000)
9. Manapure, S., Darabi, H., Patel, V., Banerjee, P.: A comparative study of radio frequency-based indoor location systems. In: *IEEE International Conference on Networking, Sensing and Control*, vol. 2, pp. 1265–1270 (2004)
10. Sidenbladh, H., Krägic, D., Christensen, H.I.: A person following behavior for a mobile robot. In: *Robotics and Automation*, pp. 1321–1326 (1998)
11. Manesis, T., Avouris, N.: Survey of Position Location Techniques in Mobile Systems. In: *Proceedings of the 7th international conference on Human computer interaction with mobile devices and services*, Salzburg, Austria (2005)
12. Hashima, M., Hasegawa, F., Kanda, S., Maruyama, T., Uchiyama, T.: Localization and Obstacle Detection for a Robot for Carrying Food Trays. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 345–351 (1997)
13. Nagumo, Y., Ohya, A.: Human following behavior of an autonomous mobile robot using light-emitting device. In: *10th IEEE International Workshop on Robot and Human Communication*, pp. 225–230 (2001)
14. Wren, C.R., Azarbayejani, A., Darrell, T., Pentland, A.P.: Pfnder: Real-Time Tracking of the Human Body. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(7) (1997)
15. Gwon, Y., Jain, R., Kawahara, T.: Robust indoor location estimation of stationary and mobile users. In: *IEEE INFOCOM*, vol. 2, pp. 1032–1043 (2004)
16. Yuta, T., Ohya, A., Yuta, S.: A Fast and Accurate Sonarring Sensor for a Mobile Robot. In: *IEEE International Conference on Robotics and Automation*, pp. 630–636 (1999)
17. Kleeman, L.: Fast and accurate sonar trackers using double pulse coding. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1185–1190 (1999)

# A Class of Biped Mechanisms with Three Links and Two Joints

Chang-Huan Liu, Yao-bin Tian, Xi-ming Wang, and Yan-an Yao

School of Mechanical and Electronic Control Engineering

Beijing Jiaotong University, 100044, Beijing, China

{06116265, 07121772, 07121773, yayao}@bjtu.edu.cn

**Abstract.** In this paper, a class of novel biped mechanisms with three links and two planar lower joints (prismatic joint and revolute joint) are presented. Based on the topological structure of mechanisms, they are classified into three subgroups: 2R, 2P and 1P1R biped mechanisms. Each subgroup contains several mechanisms with the different locomotion manners which result from different layouts of joints. Their locomotion manners are simulated with the kinematic and dynamic simulation software. Prototypes of some selected walking mechanisms are built to demonstrate their mobility. These mechanisms offer options for designing simple and robust biped robots.

**Keywords:** Biped robot; legged locomotion; legged robots; mechanism design; underactuated robots.

## 1 Introduction

Numerous biped robots have been developed successfully for use in industrial, space, medical or military applications. These robots have static walk models or fast dynamic walk models. Some of them are simple and some are very complex.

One of the simplest biped models containing only two straight legs linked by a revolute joint was demonstrated to walk down a shallow slope, powered only by gravity [1]. It was the simplest case of the passive dynamic models fabricated by McGeer [2]. Based on this simplest model, a prototype, the dynamic walking toy named Elvis, was fabricated to walk down a shallow slope all by itself [3]. A three-link biped model containing two kneeless legs connected through two activated revolution joints to the torso was proposed and studied theoretically to research the nonlinear control stratagem [4,5]. A single rigid body and two legs of variable length constituted a biped model [6]. This model was driven by three actuators to walk dynamically and validate the potential energy conserving orbit. The biped walker composed of a body and two telescopic legs was explored to walk dynamically driven by only two actuators [7]. The algorithm based on the tracking of the commanded path was implemented by the prototype. HRP-2 and HRP-3 [8,9] were the humanoid robots which could cope with uneven surface, walk at two third level of human speed, walk on a narrow path, and get up by itself if it tipped over. RunBot [10] was a biped robot

which could walk with a high speed, self-adapting to minor disturbances, and reacting in a robust way to abruptly induced gait changes. The more sophisticated full-size humanoid robots were developed by companies, like Honda's ASIMO [11,12,13] and Toyota's running robot [14]. They were the most representative examples of highly complex humanoid biped robots which could achieve a variety of gaits.

From the view of robusticity, the simpler the structure is, the more indefectible the biped robot is. In this paper, a class of simple biped mechanisms containing three links and two planar lower pairs (prismatic joint and revolute joint) is proposed and analyzed. The 2R subgroup is proposed first. The biped mechanisms in this subgroup have the same kinematic chain with the model in [4,5], but they are underactuated driven by only one actuator, not two. They can achieve straight walking, straight crawling and circular crawling due to the different layouts of the two joints. According to the theory of topological structure of mechanisms, we replace 2R with 2P to obtain the 2P subgroup biped mechanisms. The difference between the mechanisms in the 2P subgroup is the layouts of the axes of 2P. We replace one of the two revolute joints with one prismatic joint to obtain 1P1R subgroup biped mechanisms. Different layouts of the two joints cause the models' different locomotion manners.

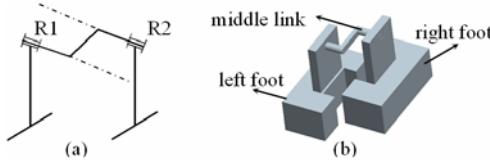
This paper is arranged as follows. In section 2, three kneeless 2R biped mechanisms actuated by one motor are proposed. Their walking principles are analyzed and verified by simulations and experiments. In section 3, two 2P biped mechanisms are presented. Their walking gaits are expressed by the simulations. In section 4, the 1P1R biped mechanisms are proposed and their walking principles are illustrated by simulations. One of the models is fabricated to testify its movability.

## 2 Biped Mechanisms with Two Revolute Joints (2R)

The biped mechanisms in this subgroup are composed of three links and two revolute joints. Either of the two joints is active. This is the significant difference between these models and the one possessing two active revolute joints in [4,5]. These 2R biped mechanisms are much simpler because only one joint is needed to be driven to complete the locomotion. The three models in this section have different locomotion manners, straight walking, straight crawling and circular crawling. We can introduce extra degree of freedoms and redesign the links to combine their characters required for us.

### 2.1 The Straight Walking 2R Biped Mechanism

The straight walking 2R biped mechanism is symmetrical. The axes of the two revolute joints are parallel as displayed in Fig. 1(a). In the 3D (three dimension) model in Fig. 1(b), the three links are named as left foot, right foot and middle link. Two feet are designed to embed mutually, so the ZMP (zero moment point) is always in the support polygon formed by the stance foot. Therefore, the mutual embedded feet prevent the biped mechanism from dumping when the swing foot is lifted up. In the following, the



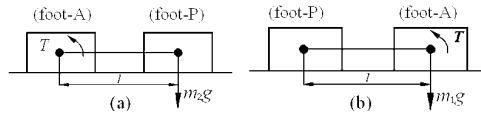
**Fig. 1.** The straight walking 2R biped mechanism: (a) the sketch; (b) the 3D model

foot on which a motor is mounted is called foot-A (Active) and the other is called foot-P (Passive).

As depicted in Fig. 2, in order to realize the straight walking, the motor torque should satisfy the following conditions:

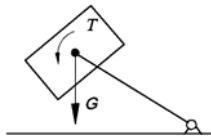
$$T \geq m_2 g l, \quad T \geq m_1 g l, \quad (1)$$

where  $T$  is the motor torque,  $m_1$  the mass of foot-A,  $m_2$  the mass of foot-P,  $g$  the acceleration of gravity,  $l$  the distance between the axes of the two revolute joints.



**Fig. 2.** The forces analysis of the straight walking 2R biped mechanism: (a) Foot-A needs to be lifted; (b) Foot-P needs to be lifted

If the center of the revolute joint is located at the mass center of foot-A, the robot can not keep the dynamic balance. As depicted in Fig. 3,  $G$  is the weight of foot-A.

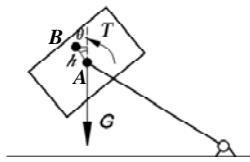


**Fig. 3.** Foot-A can not keep the dynamic balance

To keep the dynamic balance, the relation between the motor anti-torque acting on foot-A and the position of the revolute joint should be:

$$G \cdot h \cdot \sin \theta = T, \quad (2)$$

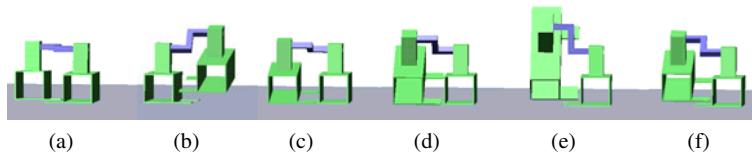
where  $\theta$  is the acute angle between the vertical line and line  $AB$ , seeing Fig. 4. Point  $A$  is the gravity center of foot-A and point  $B$  is the center of the axis of the revolute joint.  $h$  is the distance between point  $A$  and  $B$ .



**Fig. 4.** Foot-A can keep the dynamic balance

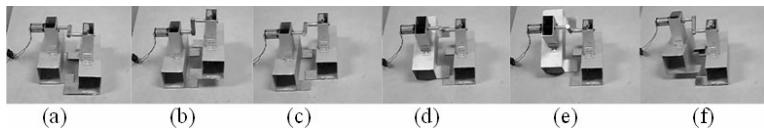
A virtual prototype of the biped mechanism is modeled. The phases of the straight walking gait are simulated and shown in Fig. 5:

(a) The mechanism is in its initial posture and position. (b) Right foot is lifted up and put forward. (c) Right foot is put down. (d) Left foot is lifted up. (e) Left foot is put forward. (f) Left foot is put down.



**Fig. 5.** The phases of the straight walking gait of the model

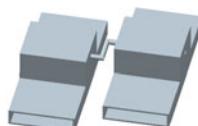
The prototype of the straight walking 2R biped mechanism is constructed. The phases of the gait in the experiment are shown in Fig. 6, corresponding to the simulation in Fig. 5. From the experiment, the stride of the robot equals the distance between the axes of two parallel revolute joints.



**Fig. 6.** The phases of the straight walking gait of the prototype

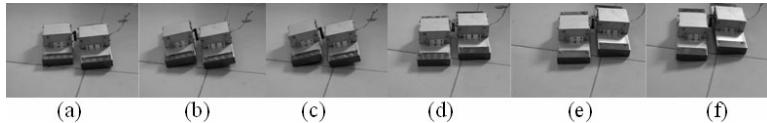
## 2.2 The Straight Crawling 2R Biped Mechanism

If the double feet of the straight walking 2R biped model are not mutual embedded, as shown in the 3D model in Fig. 7, the biped mechanism can perform the crawling gait with its feet always touching the ground.



**Fig. 7.** 3D model of the straight crawling 2R biped mechanism

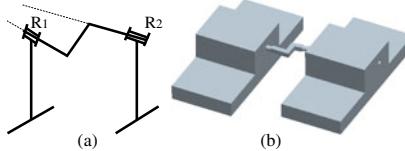
The prototype of the straight crawling 2R biped mechanism is fabricated. Its locomotion ability is experimented as displayed in Fig. 8.



**Fig. 8.** The phases of the straight crawling gait

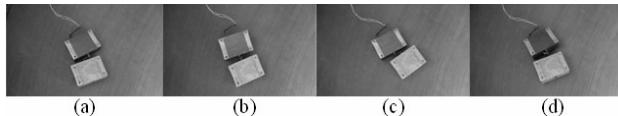
### 2.3 The Circular Crawling 2R Biped Mechanism

If the axes of the two revolute joints of the straight crawling biped mechanism are not parallel, but intersect with a certain angle as shown in Fig. 9(a), the biped mechanism can not move along a straight line, but along a circular line. The 3D model is displayed in Fig. 9(b).



**Fig. 9.** The circular crawling 2R biped mechanism: (a) the sketch; (b) the 3D model

A prototype is developed to testify its locomotion ability. The phases of the circular crawling gait are displayed in Fig. 10.



**Fig. 10.** The phases of the circular crawling gait

## 3 Biped Mechanisms with Two Prismatic Joints (2P)

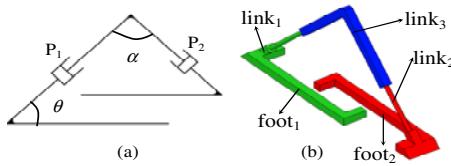
If the three links are redesigned and connected serially by two prismatic joints, the 2P biped mechanisms are obtained. There are two models in this subgroup due to the different layouts of the two prismatic joints. They are all driven by two actuators. We can introduce extra degree of freedoms and redesign the links to combine their characters required for us.

### 3.1 The Triangle 2P Biped Mechanism

As shown in Fig. 11, three links of the mechanism are labeled as link<sub>1</sub>, link<sub>2</sub> and link<sub>3</sub> respectively. Link<sub>1</sub> and link<sub>3</sub> are connected by the prismatic joint P<sub>1</sub>. Link<sub>2</sub> and link<sub>3</sub>

are connected by the other prismatic joint  $P_2$ . Neither of the two axes of the prismatic joints is parallel or vertical to the ground. Link<sub>1</sub> and link<sub>2</sub> are used as feet and legs. The mechanism touches the ground in turn with foot<sub>1</sub> and foot<sub>2</sub> which are parts of link<sub>1</sub> and link<sub>2</sub> respectively. It is necessary to design appropriate foot shapes to ensure that ZMP is always in the support polygon.

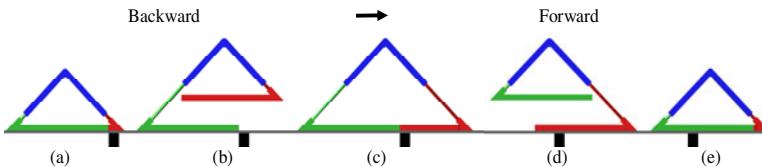
In the plane of the triangle, the acute angle between the axis of  $P_1$  and the longitudinal axis of foot<sub>1</sub> is labeled as  $\theta$ . The feet should be designed with special shapes to avoid interferences when the mechanism moves. The axis of  $P_1$  intersects the axis of  $P_2$  with an angle  $\alpha$ . The angle between the axis of  $P_2$  and the longitudinal axis of foot<sub>2</sub> can be calculated by  $180 - \theta - \alpha$ . The 3D model is built as shown in Fig. 11(b). The U-shape feet solve the interference problem.



**Fig. 11.** The triangle 2P biped mechanism: (a) the sketch; (b) the 3D model

This biped mechanism realizes the straight walking gait through the following five phases simulated by the dynamic simulation software ADAMS, as shown in Fig. 12:

(a) The biped mechanism is in its initial posture and position when  $P_1$  and  $P_2$  are contracted entirely. (b) Expand  $P_1$  to make the other two links move along the axis of  $P_1$  while link<sub>1</sub> is the stance foot. (c) Stop expanding  $P_1$  and then expand  $P_2$  until link<sub>2</sub> touches the ground. (d) Contract  $P_1$  while link<sub>2</sub> is the stance foot. (e) Contract  $P_2$  until link<sub>1</sub> touches the ground. The mechanism completes a step of forward movement after the five phases. Due to the symmetrical structure, the mechanism can move backward.



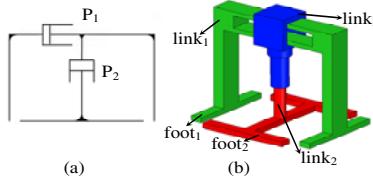
**Fig. 12.** The phases of the forward movement (triangle 2P biped mechanism)

From the simulation, we can find that it is not optional to expand or contract the two prismatic joints because the later expanding or contracting must ensure that the swing foot touches the ground. Therefore, if the displacement of  $P_1$  is supposed as  $L$ ,  $P_1$  should expand or contract  $L\sin\theta/\sin(\theta+\alpha)$ . The stride of the robot is  $L\sin\alpha/\sin(\theta+\alpha)$ .

### 3.2 The Gantry Crane 2P Biped Mechanism

For the triangle 2P biped mechanism, if the axis of  $P_1$  is changed to be parallel to the ground and the axis of  $P_2$  vertical, as shown in Fig. 13(a), the gantry crane 2P biped

mechanism is obtained. Inspired by the motions of the gantry crane, if link<sub>1</sub> is the gantry mounting, link<sub>2</sub> can be regarded as an object which is lifted. The object can be moved in a plane. If link<sub>2</sub> is considered as a foot and the gantry mounting as the other foot, the 2P biped mechanism could walk. In order to walk steadily by keeping ZMP in the support polygon, it is necessary to design special feet shapes for link<sub>1</sub> and link<sub>2</sub>. To avoid interference, foot<sub>1</sub> is T-shaped and foot<sub>2</sub> is I-shaped. The entire structure is modeled as shown in Fig. 13(b). When the mechanism moves, foot<sub>1</sub> and foot<sub>2</sub> touch the ground in turn by expanding or contracting the prismatic joints.

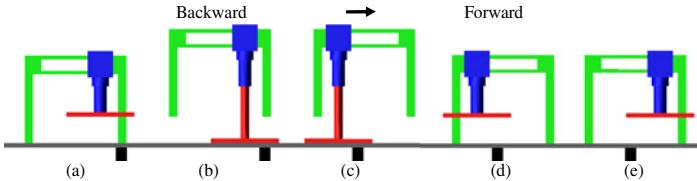


**Fig. 13.** The gantry crane 2P biped mechanism: (a) the sketch; (b) the 3D model

The walking motions are simulated as shown in Fig. 14:

(a) The biped mechanism is in its initial posture and position with link<sub>1</sub> being the stance foot and other two links being set at the extreme right location of link<sub>1</sub>. (b) Expand P<sub>2</sub> until foot<sub>2</sub> supports the other two links. (c) Actuate P<sub>1</sub> until link<sub>1</sub> slips right extremely. (d) Contract P<sub>2</sub> until link<sub>1</sub> supports the other two links. (e) Actuate P<sub>1</sub> until link<sub>2</sub> and link<sub>3</sub> slip right extremely.

The biped mechanism completes a step of forward movement and returns to the initial posture after the five phases. It can also move backward.



**Fig. 14.** The phases of the forward movement (gantry crane 2P biped mechanism)

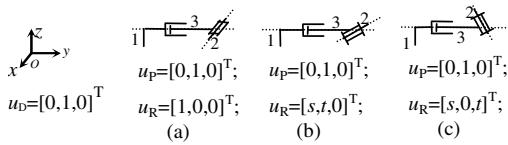
From the simulation, we can find that the expanding or contracting of the two prismatic joints is optional, unlike the triangle 2P biped mechanism. Consequently, the stride of the biped robot is the displacement of P<sub>1</sub>.

#### 4 Biped Mechanisms with One Prismatic Joint and One Revolute Joint (1P1R)

The former biped mechanisms are composed of three links connected serially by two prismatic joints or two revolute joints. According to the theory of topological structure of mechanisms, we replace one prismatic joint of the 2P subgroup with a revolute joint to obtain the 1P1R subgroup. Inspired by the actions of the climber who

throws a rope with a noose at one end to catch a hold to pull himself up, the movable 1P1R biped mechanisms are developed. The revolute joint enables the mechanisms to imitate the action of throwing and the prismatic joint enables the mechanisms to imitate the action of pulling up. In this section, we introduce the 1P1R biped mechanisms.

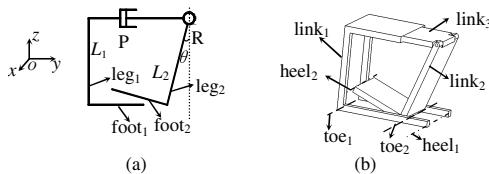
The models with different layouts of the prismatic joint and the revolute joint are found out to have different locomotion manners. The vector of the axes of the prismatic joint is  $\mathbf{u}_P$  and that of the revolute joint is  $\mathbf{u}_R$ . Suppose that the vector of the movement direction is  $\mathbf{u}_D=[0,1,0]^T$ . In the following, we give three examples with different layouts of  $\mathbf{u}_P$  and  $\mathbf{u}_R$ , as shown in Fig. 15. Certainly, there are some other different  $\mathbf{u}_P$  and  $\mathbf{u}_R$  combinations for different locomotion manners.



**Fig. 15.** The different layouts of  $\mathbf{u}_P$  and  $\mathbf{u}_R$ : 1, 2 and 3 represent the three links, link<sub>1</sub>, link<sub>2</sub> and link<sub>3</sub>;  $\sqrt{s^2 + t^2} = 1$ ,  $s \neq 0$ ,  $t \neq 0$

#### 4.1 The 1P1R (a) Biped Mechanism

The vector of the axes of the prismatic joint is  $\mathbf{u}_P=[0,1,0]^T$  and that of the revolute joint is  $\mathbf{u}_R=[1,0,0]^T$ , as shown in Fig. 16(a). In the 3D model shown in Fig. 16(b), link<sub>1</sub> and link<sub>3</sub> are connected by the prismatic joint and link<sub>3</sub> and link<sub>2</sub> are connected by the revolute joint. The dashed lines represent the lines of toes (toe<sub>1</sub> and toe<sub>2</sub>) and the dashed-and-dotted lines represent the lines of heels (heel<sub>1</sub> and heel<sub>2</sub>). In Fig. 16(a), link<sub>1</sub> and link<sub>2</sub> are used as not only inflexible leg<sub>1</sub> and leg<sub>2</sub> but also foot<sub>1</sub> and foot<sub>2</sub>. The length of leg<sub>1</sub> on link<sub>1</sub> is  $L_1$  and that of leg<sub>2</sub> on link<sub>2</sub> is  $L_2$ . There is an angle  $\theta$  between the vertical line and leg<sub>2</sub>. So  $L_1$  is less than  $L_2$ .

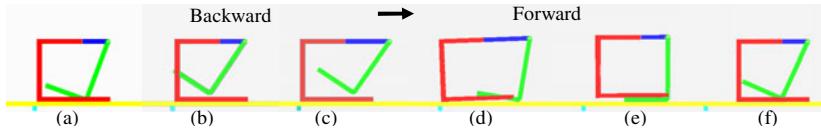


**Fig. 16.** The 1P1R (a) biped mechanism: (a) the sketch; (b) the 3D model

The initial posture is that the sole of link<sub>1</sub> and toe<sub>2</sub> of link<sub>2</sub> (a plane and a line), touch the ground at the same time as shown in the 3D model. The process of the movement of this 1P1R (a) biped mechanism is simulated using the ADAMS software as shown in Fig. 17:

(a) The mechanism is in its initial position and posture. (b) Link<sub>2</sub> is rotated clockwise to leave the ground. (c) Link<sub>2</sub> is moved forward by the prismatic joint. (d) Link<sub>2</sub> is rotated counterclockwise until the plane (the sole of link<sub>2</sub>) touches the ground. At

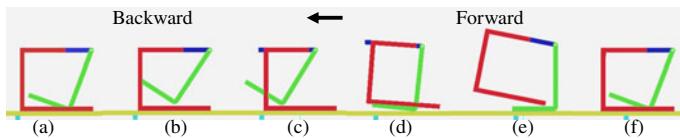
this time, only a line on link<sub>1</sub>, toe<sub>1</sub>, touches the ground. (e) Link<sub>1</sub> is moved forward by the prismatic joint because only a line on link<sub>1</sub> touches the ground, but a plane on link<sub>2</sub> touches the ground. Then, link<sub>1</sub> is lift up because leg<sub>2</sub> is longer than leg<sub>1</sub>. (f) Link<sub>1</sub> and link<sub>3</sub> are rotated counterclockwise until the sole of link<sub>1</sub> and toe<sub>2</sub> of link<sub>2</sub> touch the ground. At this time, the biped mechanism returns to its initial posture.



**Fig. 17.** The phases of the forward movement of the 1P1R (a) biped mechanism

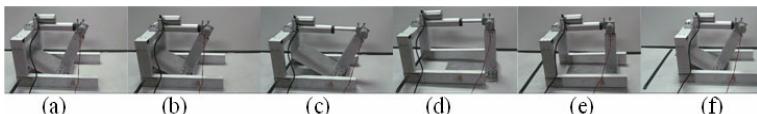
This biped mechanism can also walk backward with the phases shown by the simulation results in Fig. 18:

(a) The mechanism is in its initial position and posture. (b) Link<sub>2</sub> is rotated clockwise to leave the ground. (c) Link<sub>2</sub> is moved backward by the prismatic joint. (d) Link<sub>2</sub> is rotated counterclockwise until the plane (the sole of link<sub>2</sub>) touches the ground. At this time, only a line on link<sub>1</sub>, heel<sub>1</sub>, touches the ground. (e) Link<sub>1</sub> is moved backward by the prismatic joint. (f) Link<sub>1</sub> and link<sub>3</sub> are rotated counterclockwise until the sole of link<sub>1</sub> and toe<sub>2</sub> of link<sub>2</sub> touch the ground. At this time, the biped mechanism returns to its initial posture.



**Fig. 18.** The phases of the backward movement of 1P1R (a) biped mechanism

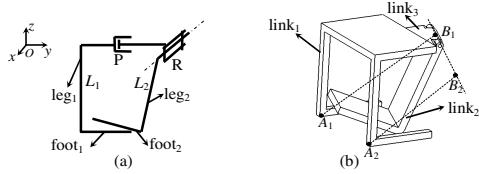
Corresponding to the simulation results in Fig. 17, a prototype is built to do the forward movement experiment. The experimental results are displayed in Fig. 19.



**Fig. 19.** The phases of forward movement of the 1P1R (a) biped mechanism

#### 4.2 The 1P1R (b) Biped Mechanism

If the vector of the axis of the prismatic joint is still  $\mathbf{u}_P=[0,1,0]^T$  and that of the revolute joint is changed to  $\mathbf{u}_R=[s,t,0]^T$ , as shown in Fig. 20(a), the biped mechanism does not walk straight. For the 3D model in Fig. 20(b),  $A_1$  is the point which connects leg<sub>1</sub> and foot<sub>1</sub>.  $A_2$  is the point which connects leg<sub>2</sub> and foot<sub>2</sub>. The distances from  $A_1$  and  $A_2$



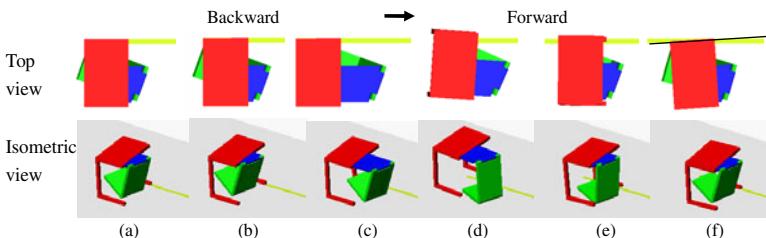
**Fig. 20.** The 1P1R (b) biped mechanism: (a) the sketch; (b) the 3D model

to the axis of the revolute joint are  $|A_1B_1|$  and  $|A_2B_2|$ . Because the vector of the axis of the revolute joint is  $\mathbf{u}_R = [s, t, 0]^T$ ,  $|A_1B_1| > |A_2B_2|$ .

The process of the forward movement of this 1P1R (b) biped mechanism is simulated as shown in Fig. 21. The motor acts are the same with the 1P1R (a) biped mechanism. The difference is the effect of the rotation in the procedure (f). At the beginning of procedure (f), link<sub>2</sub> touches the ground as the stance foot. Link<sub>1</sub> and link<sub>3</sub> are rotated around the axis of the revolute joint until the sole on link<sub>1</sub> touches the ground. Because  $|A_1B_1| > |A_2B_2|$ , the side of  $A_1B_1$  will not touch the ground at the same time with the side of  $A_2B_2$ . Therefore, point  $A_1$  touches the ground firstly and the rest parts of the mechanism rotate around it until point  $A_2$  touches the ground. Namely, the deflection is produced, as shown in Fig. 21(f).

All the phases are shown in Fig. 21:

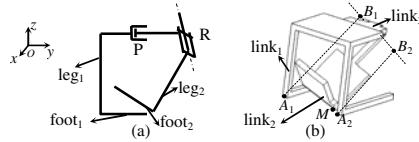
- (a) The mechanism is in the initial position and posture.
- (b) Link<sub>2</sub> is rotated clockwise.
- (c) Link<sub>2</sub> is moved forward.
- (d) Link<sub>2</sub> is rotated counterclockwise.
- (e) Link<sub>1</sub> and link<sub>3</sub> are moved forward.
- (f) Link<sub>1</sub> and link<sub>3</sub> are rotated counterclockwise.



**Fig. 21.** The phases of the forward movement of the 1P1R (b) biped mechanism

### 4.3 The 1P1R (c) Biped Mechanism

If the vector of the axis of the prismatic joint is still  $\mathbf{u}_P = [0, 1, 0]^T$  and that of the revolute joint is changed to  $\mathbf{u}_R = [s, 0, t]^T$ , as shown in Fig. 22(a), the biped mechanism does not move straight either. For the 3D model in Fig. 22(b),  $A_1$  is the point which connects leg<sub>1</sub> and foot<sub>1</sub>.  $A_2$  is the point which connects leg<sub>2</sub> and foot<sub>2</sub>. The distances from  $A_1$  and  $A_2$  to the axis of the revolute joint are  $|A_1B_1|$  and  $|A_2B_2|$ . For the 1P1R (c) biped mechanism,  $|A_1B_1| > |A_2B_2|$ , because the vector of the axis of the revolute joint is  $\mathbf{u}_R = [s, 0, t]^T$ .



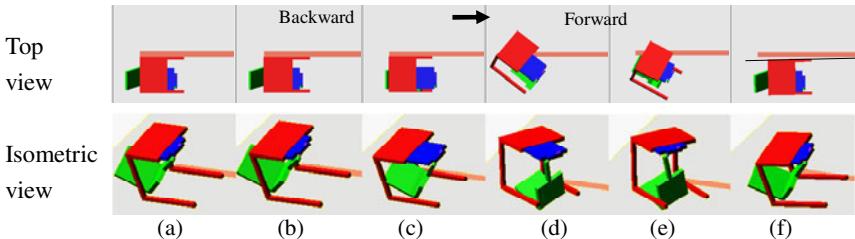
**Fig. 22.** The 1P1R (c) biped mechanism: (a) the sketch; (b) the 3D model

The process of the forward movement of this 1P1R (c) biped mechanism is simulated as shown in Fig. 23. The motor acts are also the same with the 1P1R (a) biped mechanism.

The initial posture is different from that of the former two models. The sole on link<sub>1</sub> and a lowest point *M* on toe<sub>2</sub> touch the ground because the vector of the revolute joint is  $\mathbf{u}_R=[s,0,t]^T$ . Namely, a plane and a point instead of a line touch the ground. This causes the different effects of the rotations in procedure (d) and (f). At the beginning of procedure (d), link<sub>1</sub> is the stance foot. Link<sub>2</sub> is rotated counterclockwise. During the rotation, the point *M* on link<sub>2</sub> touches the ground firstly since *M* is the lowest point on link<sub>2</sub>. Therefore, the biped mechanism rotates around point *M* as shown in Fig. 23(d). At the beginning of procedure (f), link<sub>2</sub> is the stance foot with its sole touching the ground. Link<sub>1</sub> is rotated counterclockwise. Since  $|A_1B_1|>|A_2B_2|$ , the point *A*<sub>1</sub> on link<sub>2</sub> touches the ground firstly. So the biped mechanism rotates around *A*<sub>1</sub>*M*. The deflection arises as shown in Fig. 23(f).

All the phases are shown in Fig. 23:

(a) The mechanism is in the initial position and posture. (b) Link<sub>2</sub> is rotated clockwise. (c) Link<sub>2</sub> is moved forward. (d) Link<sub>2</sub> is rotated counterclockwise. (e) Link<sub>1</sub> is moved forward. (f) Link<sub>1</sub> is rotated counterclockwise.



**Fig. 23.** The phases of the forward movement of the 1P1R (c) biped mechanism

## 5 Conclusions

A class of biped mechanisms with three links and two joints is proposed in this paper. According to the theory of topological structure of mechanisms, these mechanisms are classified into three subgroups as 2R, 2P and 1P1R biped mechanisms. In each subgroup, the structures with different locomotion manners are presented by changing the layouts of the joints and the shapes of the links. Their walking principles are explained under the help of simulations. Some prototypes of these biped mechanisms are fabricated and testify their locomotion ability successfully.

These mobile mechanisms with novel and simple structures are derived from mechanism theory. They afford several new members for the family of mobile robots. In addition, redesigning the parts shapes and introducing some other degree of freedoms to change the layouts of the joints can get more flexible robot mechanisms. Further work is to conduct in-depth assessment of their locomotion abilities, according to specific applications and their own characteristics.

**Acknowledgments.** This work was supported by Program for New Century Excellent Talents in University (NCET-07-0063), National Natural Science Foundation of China (50875018) and Beijing Natural Science Foundation (3093025).

## References

1. Garcia, M., Chatterjee, A., Ruina, A., Coleman, M.: The Simplest Walking Model: Stability, Complexity and Scaling. *ASME Journal of Biomechanical Engineering* 120(2), 281–288 (1998)
2. McGeer, T.: Passive Dynamic Walking. *The International Journal of Robotics Research* 9(2), 68–82 (1990)
3. DYNAMIC WALKING (2008),  
<http://www.dynamicwalking.org/dw2008/?q=node/46>
4. Kiefer, J., Bale, R.: Walking Viability and Gait Synthesis for a Novel Class of Dynamically Simple Bipeds. *Informatica* 17, 145–155 (1993)
5. Spong, M.W., Lozano, R., Mahony, R.: An Almost Linear Biped. In: Proc. the 39th IEEE Conf. Decision and Control, vol. 5, pp. 4803–4808 (2000)
6. Kajita, S., Tani, K., Kobayashit, A.: Dynamic Walking Control of a Biped Robot along a Potential Energy Conserving Orbit. *IEEE Trans. Robotics and Automation* 8(4), 431–438 (1992)
7. Grishin, A.A., Formal'sky, A.M., Lensky, A.V., Zhitomirsky, S.V.: Dynamic Walking of a Vehicle with Two Telescopic Legs Controlled by Two Drives. *The International Journal of Robotics Research* 13(2), 137–147 (1994)
8. Kaneko, K., Kanehiro, F., Kajita, S., Hirukawa, H., Kawasaki, T., Hirata, M., Akachi, K., Isozumi, T.: Humanoid robot HRP-2. In: Proc. of the IEEE Int. Conf. Robotics and Automation (ICRA), pp. 1083–1090 (2004)
9. Kaneko, K., Harada, K., Kanehiro, F., Miyamori, G., Akachi, K.: Humanoid robot HRP-3. In: Proc. of IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), pp. 2471–2478 (2008)
10. Manoonpong, P., Geng, T., Kulvicius, T., Porr, B., Worgotter, F.: Adaptive, Fast Walking in a Biped Robot under Neuronal Control and Learning. *PLoS Computational Biology* 3(7), 1305–1320 (2007)
11. Chestnutt, J., Lau, M., Cheung, G., Kuffner, J., Hodgins, J., Kanade, T.: Footstep planning for the honda asimo humanoid. In: Proc. IEEE Int. Conf. Robotics and Automation, pp. 629–634 (2005)
12. Mutlu, B., Osman, S., Forlizzi, J., Hodgins, J., Kiesler, S.: Perceptions of ASIMO: an exploration on co-operation and competition with humans and humanoid robots. In: Proc. of 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction, Salt Lake City, Utah, USA, March 02-03 (2006)
13. Hirose, M., Ogawa, K.: Honda humanoid robots development. *Philosophical Transactions of the Royal Society London, Series A* 365(1850), 11–19 (2007)
14. Tajima, T., Honda, D., Suga, K.: A fast running experiment involving a humanoid robot. In: Proc. The RSJ Annual Conference (2008)

# Time Efficient Strategy to Explore Unknown Indoor Environments by Mobile Ground Robots

Mirco Alpen, Stephan Stuedemann, and Joachim Horn

Helmut-Schmidt-University / University of the Federal Armed Forces Hamburg,  
Department of Electrical Engineering, Institute for Control Engineering,  
P.O. Box 700822, D-22008 Hamburg, Germany

**Abstract.** Here we present an algorithm for exploration of an unknown indoor environment for mobile ground robots. Due to the combination of local and global navigation strategies this algorithm is robust in relation to dead end floors. We present an AFA (**A**rtificial **F**orce **A**pproach), based on the APF (**A**rtificial **P**otential **F**ields) method. A combination of AFA and the A\*-algorithm leads to a robust and reasonable exploration algorithm of indoor environments based on an orthogonal SLAM (**S**imultaneous **L**ocalisation **A**nd **M**apping) algorithm.

All methods and approaches of this paper are validated by simulation and real time experiments with a ground robot named Scorpion.

## 1 Introduction

Autonomous robots are used wherever hazardous environments are too dangerous or unacceptable for human. Mobile robots may be reliable tools especially in exploring unknown environments. Therefore autonomous movement and mapping in unknown areas are demanding tasks to achieve.

Navigation techniques are generally differentiated in local and global techniques [1]. In local navigation techniques actions are based on current sensor data, which allows low response times. Therefore they are very useful for collision avoidance and reaching goals nearby. Because of the limited information processed in local techniques they are vulnerable to moving objects and dead ends. They also may not find the optimal path to a goal position. Examples for local navigation techniques are rule based navigation [2], vector field histogram techniques [3], the artificial potential field method [4] or dynamic window approach [5].

In this paper we use a simplified potential filed method called AFA (**A**rtificial **F**orce **A**pproach). This method has a very low computing time and thus it is dedicated for real time applications.

In order to compensate the disadvantages of local techniques, global navigation techniques are used. With comprehensive information they are able to avoid dead ends and find optimal paths for even distant goals. As many problems of engineering this task can be related to the general problem of finding an optimal path through a graph [6]. So the known environment must be divided in order to

create a graph. There are different techniques known to subdivide the area like cell decomposition [7], road map by visibility graph [8] or Voronoi diagram. Now the graph can be evaluated and search algorithms like Dijkstra or A\* [6] can be used to find an optimal path regarding the evaluation. Here we are focusing on the visibility graph because the SLAM algorithm [9] generates a feature based map of the environment.

To evaluate the graph we use the A\*-algorithm [6]. This algorithm belongs to the class of informed search algorithms. By using heuristics the A\*-algorithm has a low computing time and always leads to the shortest path between the two current points. Therefore, assuming the robot is moving with a constant velocity, this algorithm leads to the time optimal path.

It becomes clear, that global navigation techniques are useful to cover huge distances but they have disadvantages in obstacle avoidance. Therefore a combination of both, global and local techniques may build a reliable and flexible navigation algorithm to face various locations.

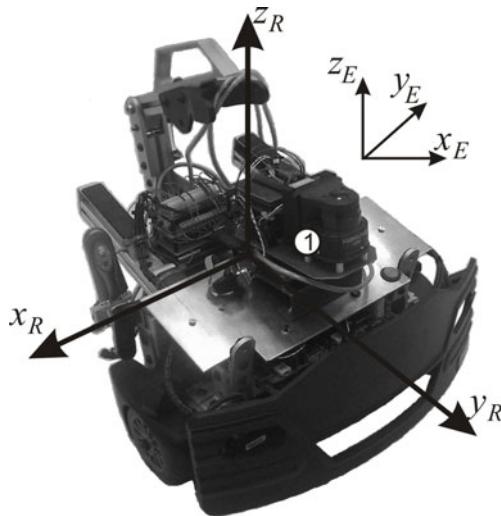
Exploration of unknown environments needs more than a navigation technique. This wide and complex task can be broke down into four subtasks [10]. These are the identification of the robots position and the position of the objects in its environment, the collision avoidance, the path planning and the action selection to resolve conflicts between collision avoidance and path planning. This paper covers the realisation of different local and global navigation techniques, an orthogonal SLAM algorithm and a goal setting and evaluation algorithm [11] in order to provide the ability of exploring and mapping unknown indoor environments. The strategy of exploration was to set goals and evaluate them regarding the mapped features of the environment.

All methods and approaches described in section 3 up to section 5 of this paper are validated by simulation and real time experiments with the robot 'Scorpion' which is described in section 2. Some of the results are shown in section 6.

## 2 Introductive System Description

Robots are complex systems that are specially designed to fulfil their task and to operate reliable in their application area. In line with this paper the application area can be defined as typical static indoor area. The features of that area are easy to identify as walls and optionally the corresponding corners. These features allow using a line feature based SLAM algorithm as base for exploration and navigation strategies. The robot Scorpion, used in this work is shown in figure 1. It is structured in sensors, actors, control and power supply.

The robots sensors include 20 SHARP® GP2D12 infrared sensors for close range obstacle detection and a HOKUYO® UGB-04LX-F01 laser scanner with an effective reach up to 4 meters. The laser range finder marked with a '1' in figure 1 is the primary sensor and continuously provides the data for the SLAM algorithm.



**Fig. 1.** The robot Scorpion

The robot is equipped with two separate DC motors that allow turning the robot on the spot. Due to the robots rotation, we are dealing with two reference frames. The body fixed coordinates are marked with the index 'R' and the earth coordinates with an 'E'. In the literature several methods to obtain the angle between both reference frames are mentioned. Here this is done by evaluation of the odometric data.

An on-board computer based on a mini-ITX VIA® main board with a 1200 MHz VIA CPU is used to control and monitor the operation. With its multiple serial interfaces the on-board computer also allows an easy communication with the different micro controllers and other hardware devices. It is also used to run the Matlab® code.

### 3 Local Navigation

The main task of the local navigation is reaching the given goal  $g$  including the obstacle avoidance. The generation of possible goals  $g_i$  will be described in section 5. The local navigation is done by the two steps described in this section. During the first step the robots heading is computed by the AFA (**A**rtificial **F**orce **A**pproach). The second step contains the obstacle avoidance by a rule based navigation.

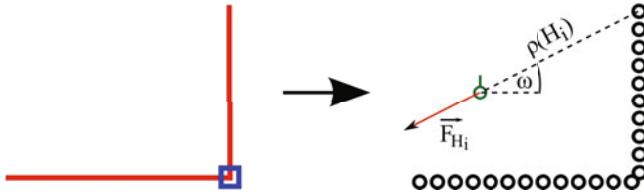
To use the AFA, features of the local map build by the SLAM algorithm are screened as shown in figure 2. As can be seen in figure 2 the lines are represented by equally spaced points dedicated with a repulsive force  $F_{H_i}$  depending on the distance between robot and feature. The  $x$  and  $y$  components of this force are computed by

$$Fx_{H_i} = \frac{1}{k_1} \cos(\omega) \frac{1}{\rho(H_i)^3} \quad (1)$$

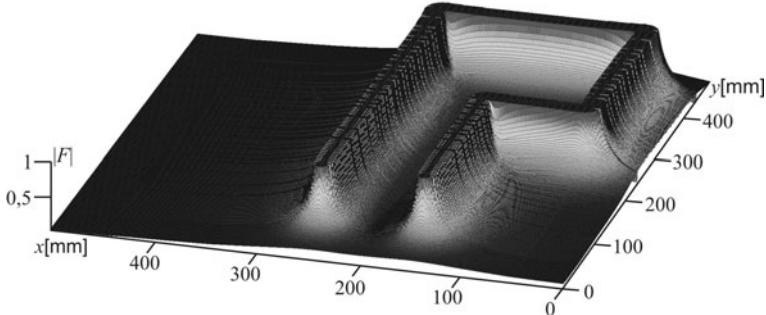
and

$$Fy_{H_i} = \frac{1}{k_1} \sin(\omega) \frac{1}{\rho(H_i)^3}. \quad (2)$$

The function  $\rho(H_i)$  describes the Euclidean distance between robot and feature. Here it is a cubic function to ensure higher forces near by the obstacles and comparatively low forces with growing distances.  $k_1$  is a tuning parameter depending on the environment.



**Fig. 2.** Screening of features and resulting forces



**Fig. 3.** Distribution of forces

Figure 3 shows an example of the distribution of forces. The  $z$ -axes of this plot gives information about the magnitude of forces caused by the features of a given map. Very close to the walls the forces have their maximum value.

The given goal is dedicated with an attractive force  $F_Z$ . The  $x$  and  $y$  components of this force are computed by

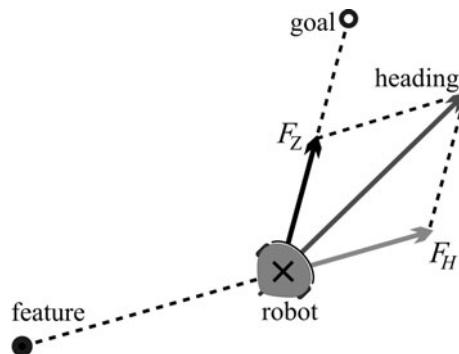
$$Fx_Z = k_2 \cos(\alpha) \quad (3)$$

and

$$Fy_Z = k_2 \sin(\alpha), \quad (4)$$

where  $\alpha$  denotes the angle of the imaginary line between the waypoint and the current robots position and  $k_2$  is a tuning parameter depending on the environment. The magnitude of this attractive force  $F_Z$  is constant to ensure an adequate influence on the robot for the whole environment.

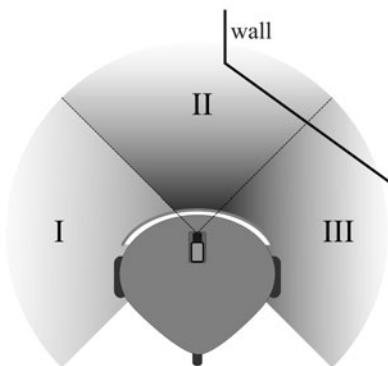
The robots heading can now be computed by a superposition of all repulsive forces  $F_{H_i}$  and the attractive force  $F_Z$ . Figure 4 shows a simple example with only one feature and the cooresponding force  $F_H$  for this procedure.



**Fig. 4.** Schematic computation of the robots heading

Before the robot starts his movement in the computed direction, potential obstacles have to be detected. Due to the SLAM algorithm only orthogonal lines are saved in the map. Therefore it is necessary to consider the current measurement data of the laser range finder and the infrared sensors.

The operating range of the sensors is divided into three regions as can be seen in figure 5. If there is an obstacle in one of theses regions it is closed for the robots movement. In the given example we are dealing with two closed sectors.

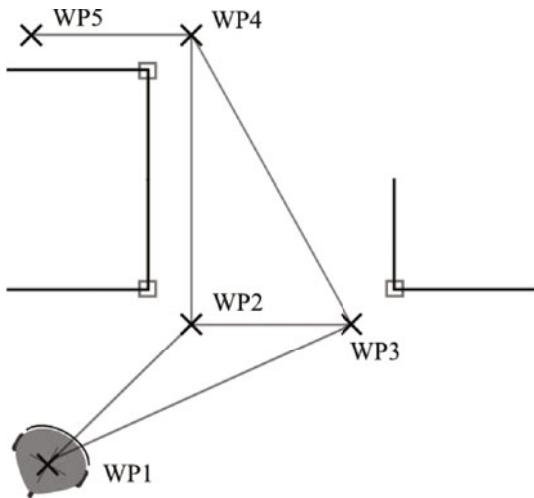


**Fig. 5.** Three regions of the operating range

If the computed heading is located within a closed sector, the course is set to a nonclosed region. If all three regions are closed the robot has to make a turn of 180 degree.

## 4 Global Navigation

The global navigation used in this paper is based on a road map approach. A road map represents all possible paths between the current position and the goal in a network of curves or lines. Here this road map is given by a visibility graph. Figure 6 shows such a graph for a small example.



**Fig. 6.** Visibility graph with five waypoints

In this example waypoint number 1 (WP1) denotes the robots current position and waypoint number 5 (WP5) denotes the goal. All waypoints are connected with lines. The generation of the different waypoints is based on the obstacle growing method.

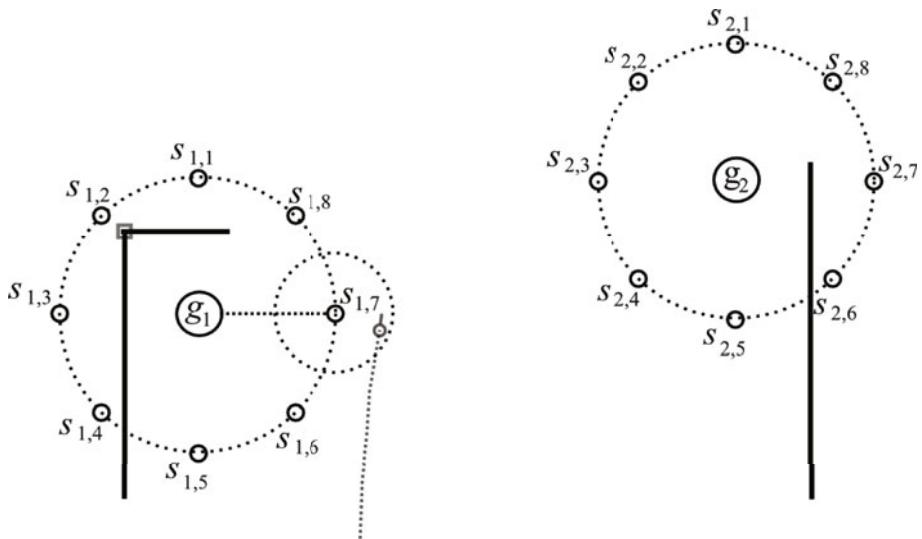
After computing the visibility graph the A\*-algorithm is used to find an optimal path between the robots current position and the goal. This leads to the shortest way between both points. Due to the fact dealing with even and coplanate undergrounds as well as a constant velocity of the robot, the result is the time optimal solution, too. The application flow of the A\*-algorithm is sufficiently described in the literature [6] and therefore not part of this paper. For the small example given in figure 6 the resulting time optimal path would include the waypoints WP2 and WP4.

## 5 Strategy of Exploration

In our case the exploration of an unknown environment is done by a feature based approach. As mentioned before, the exploration is based on an orthogonal SLAM. Therefore the resulting features extracted from the measurement data are walls replaced by lines and the corresponding corners. Our strategy of exploration can be summarized in four rules. The order of the rules corresponds to the order of listing.

1. Exploration of environments with low density of features.
2. Avoidance of areas that are already known.
3. Exploration of near and easy to reach environments.
4. Exploration of solitary areas if the nearer environment is already known

The exploration starts by generating goals  $g_i$  based on the features of the map and the evaluation algorithm presented in [11]. The potential goals  $g_i$  are put near by every open end of the current features. A small example with two possible goals is shown in figure 7.



**Fig. 7.** Example of feature exploration with two possible goals

8 samples named  $s_{i,j}$  are put orbital around each goal. The distance between sample and goal depends on the resolution of the map. The potential goals are weighted based on these samples. Here each sample has a weight of 1. The weight of the current goal is computed by dividing the sum of weight of the valid samples and the sum of weight of all samples. Valid samples are those which are not blinded by features from the view of  $g_i$  and not placed near by the robots current or a past position. Therefore, for the goal  $g_1$  of figure 7 only the samples

$s_{1,8}$ ,  $s_{1,6}$  and  $s_{1,5}$  are valid. Sample  $s_{1,7}$  is located near by the robots position and the other ones are blinded by features. Therefore the weight  $e(g_1)$  is computed by

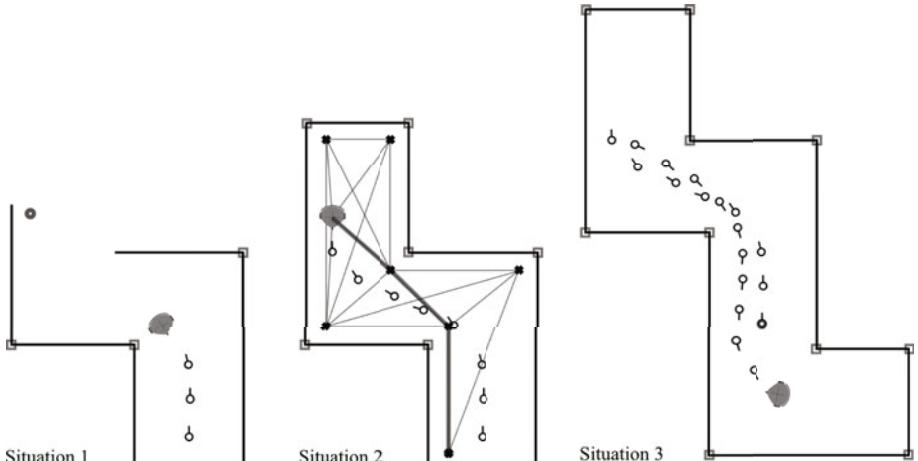
$$e(g_1) = \frac{e(s_{1,8}) + e(s_{1,6}) + e(s_{1,5})}{\sum_{j=1}^{j=8} e(s_{1,j})} = \frac{3}{8}. \quad (5)$$

The resulting weight of goal  $g_2$  is computed in the same way. Here we get

$$e(g_2) = \frac{e(s_{2,1}) + e(s_{2,2}) + e(s_{2,3}) + e(s_{2,4}) + e(s_{2,5}) + e(s_{2,8})}{\sum_{j=1}^{j=8} e(s_{2,j})} = \frac{6}{8}. \quad (6)$$

In this example the robot would choose goal  $g_2$  because of its higher value.

After choosing its goal, the robot starts to move towards it by using the local navigation described in section 3. If the distance between goal and current position is above a certain threshold value the robot would use the global navigation to reach its goal. If there are no more areas to explore, the robot moves back to the initial point of exploration. Figure 8 shows three steps of exploration based on a simulation.



**Fig. 8.** Simulation of exploration

Due to the procedure described above, in situation 1 of figure 8 the goal marked by a circle in upper left corner is chosen by the robot. The robot moves towards this point using local navigation techniques. In situation 2 the robot reaches the goal. Now a new evaluation is done and the result is in the lower left corner. The distance between goal and robot is above a certain threshold value and a path is computed by the A\*-algorithm. Due to the fact, that global navigation uses the information of the whole map, the generation of waypoints differs from the generation of goals during the local navigation. The resulting path is marked by a bold line in deep grey.

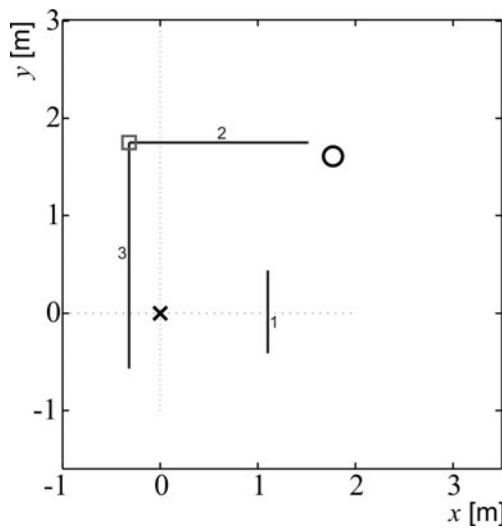


Fig. 9. Initial condition of exploration

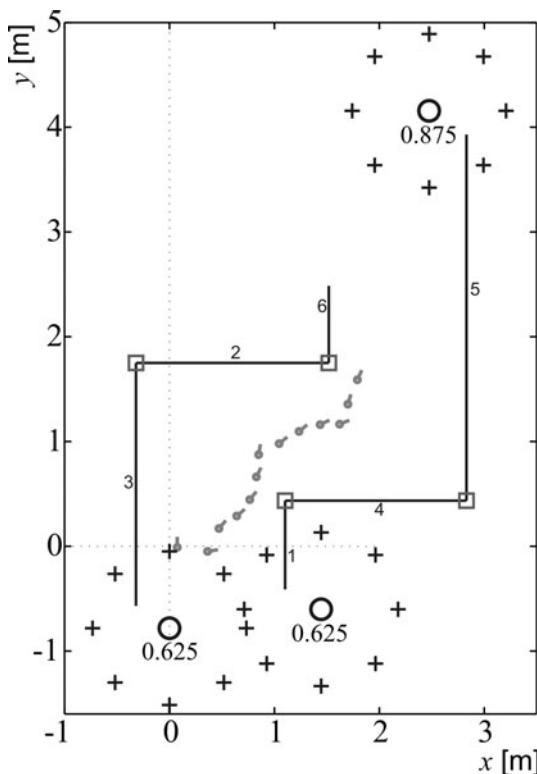
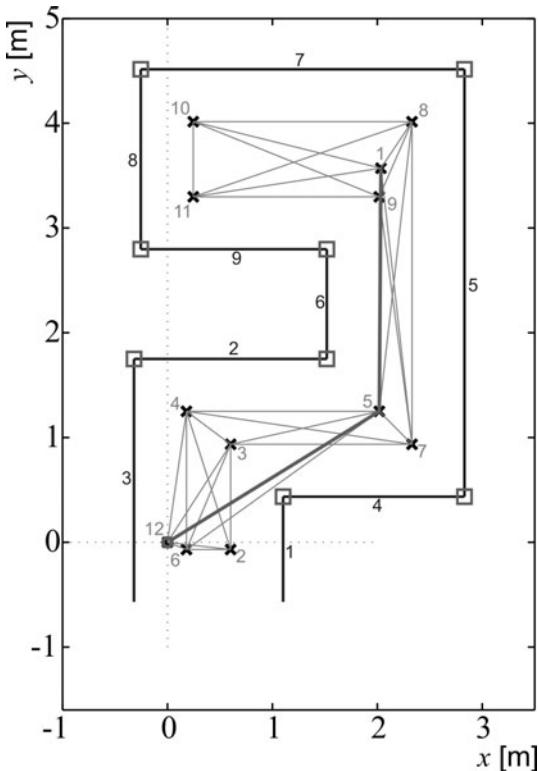


Fig. 10. Feature exploration during the mission



**Fig. 11.** Computation of an optimal path during the mission

Situation 3 shows the point, where the whole area is known. Now the robot gets back to the initial point by local navigation.

By the combination of global and local navigation dead ends are avoided and an optimal path for even distant goals will be found. With the assumption of a constant velocity the result is a time optimal exploration algorithm for an indoor environment.

## 6 Experimental Results

In this section the experimental results of a test environment in our laboratory are shown. During the experiments the robot acts fully autonomous. The robot starts at the origin of figure 9. Both coordinate planes are congruent at the initial point. Figure 9 shows the resulting map of the first scan.

The robot gets his first goal marked with a circle in the upper right corner by feature exploration. Choosing the goal, the robot starts to move towards it using local navigation. During the robots movement the map is growing because the SLAM-algorithm is online during the whole mission. If the goal is reached by the robot, a new feature exploration starts. This is shown in figure 10.

Three possible goals are evaluated by the robot. The most attractive goal is in the upper right corner of figure 10 and the robot starts moving towards it. If the robot reaches this goal, the upper region of the map is closed. A new feature exploration is done. At this time the most attractive goal is too far away for the local navigation. An optimal path is computed by the A\*-algorithm. The result is marked by a bold line in deep grey in figure 11.

Reaching the computed point, the robots mission comes to an end, because there is a wall from the left to the right in the lower region of the test area.

This small example shows the effective combination of global and local navigation for a reasonable time optimal exploration of an unknown indoor environment.

## 7 Conclusion

In line with this work the level of autonomy of the mobile robot was to be expanded in order to provide the ability to independently explore an unknown indoor environment. Therefore the main tasks, localisation and mapping, collision avoidance and path planning have been realized in several Matlab® functions.

For localisation and mapping a line feature based orthogonal SLAM algorithm was used. For the local navigation a simplified potential field method named AFA (**A**rtificial **F**orce **A**pproach) in combination with a rule based navigation for obstacle avoidance was presented. In order to avoid the disadvantages of local navigation, like local minima and dead ends a global navigation technique based on a visibility graph and the A\*-algorithm has been realized.

An efficient exploration is ensured by a feature based goal generation and evaluation. In that way goals which have a high value for the exploration of the unknown area can be found.

In our application we are dealing with a constant velocity of the robot. The A\*-algorithm always leads to the shortest path between two points. Thus, the combination of these local and global navigation methods leads to a reasonable and time efficient exploration of unknown indoor environments.

In experimental tests the robot proved his ability to localize, find and reach even distant goal points via optimal paths and to follow his track without collisions.

## References

1. Rao, S., Iyengar, S., Jorgensen, C., Weisbin, C.: Concurrent algorithms for autonomous robot navigation in an unexplored terrain. In: IEEE International Conference on Industrial Technology (ICIT), pp. 1137–1144 (1986)
2. Fujimura, K.: Motion planning in dynamic environments. Springer, Tokyo (1992)
3. Borenstein, J., Koren, Y.: The vector field histogram - fast obstacle avoidance for mobile robots. IEEE Robotics & Automation Magazine (June 1991)
4. Khatib, O.: Real-Time Obstacle Avoidance For Manipulators And Mobile Robots. Stanford University, CA 94305 USA (March 1985)
5. Fox, D., Burghard, W., Thrun, S.: The Dynamic Window Approach to Collision Avoidance. IEEE Robotics & Automation Magazine (March 1997)

6. Hart, P.E., Nilson, N.J., Raphael, B.: A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions of Systems Science and Cybernetics* (July 1968)
7. Lingelbach, F.: Pathplanning using Probabilistic Cell Decomposition. Centre for Autonomous Systems / Royal Institute of Techology, Stockholm, Sweden (April 2004)
8. Nilson, N.J.: A Mobile Automation: An Application of Artificial Intelligence Techniques. Stanford Research Institute, Menlo Park, CA USA (January 1969)
9. Alpen, M., Willrodt, C., Frick, K., Horn, J.: On-board SLAM for indoor UAV using a laser range finder. SPIE Defense, Security, and Sensing, Orlando, USA (2010)
10. Ibrahim, M.Y., Fernandes, A.: Study on Mobile Robot Navigation Techniques. In: IEEE International Conference on Industrial Technology (ICIT), pp. 230–236 (2004)
11. Newman, P., Bosse, M., Leonard, J.: Autonomous Feature-Based Exploration. Massachusetts Institute of Technology, Cambridge, USA (September 2003)

# Overhead High-Voltage Transmission Line Deicing Robot System and Experiment Study\*

Bai Yucheng, Wu Gongping, Xiao hua, Yan yu, Li Yingsong, and Fu Xingwei

School of Power And Mechanical Engineering,  
WuHan University, Wuhan, Hubei, P.R. China  
baiyc123@163.com

**Abstract.** Based on the further study of high-voltage transmission line deicing environment, this paper proposed a wheel milling deicing device and a compound deicing robot that can have double wheel milling and compressing wheel rolling simultaneously; implemented deicing experiment and obstacle negotiation simulation analysis on the transmission and robot mechanic system; proposed an obstacle negotiation navigation control method based on electromagnetic sensors, and finally implemented deicing experiment and obstacle negotiation experiment on deicing robot prototypes. The experiment demonstrates the feasibility of electromagnetic navigation technology in leading robot in obstacle negotiation, and also shows that it's possible to use compound deicing method of milling and compressing wheel rolling.

**Keywords:** Deicing robot, wheel milling, obstacle negotiation, electromagnetic navigation, transmission line.

## 1 Introduction

Severe icing on high-voltage transmission lines will lead to the tilting and collapse of towers and the fracture and flashover of transmission lines. It will result in the tripping of transmission lines and interrupts power supply, which has a huge negative effect on industrial and agricultural production as well as people's daily life. When icing damage happens, it is always in bad weather, blocked transportation and broken communication. It's hard to repair immediately, so it would be need a long time until the power supply system comes back to normal [1]. In early 2008, there was a large area of ice damage in Southern China, which resulted in a long period of power failure and caused great loss to China's national economy [2]. High-voltage transmission deicing is an effective way to reduce icing damage.

Currently there are many deicing methods adopted within and outside of China, such as thermal deicing, mechanical deicing, electric pulse, laser deicing, ultrasonic deicing and microwave deicing [3-8], etc. Thermal deicing mainly uses methods of increasing load current, alternating current short circuit and direct current short circuit

---

\* This work is partially supported by Chinese national science and technology supporting project (No. 2008BAF36B01).

[9]. Thermal deicing is highly efficient, but the deicing power is huge, so on-line deicing is not possible. Mechanical deicing is the most common and effective method currently. It allows on-line deicing but it mainly functions manually and has low efficiency and certain risk. Electric pulse, laser deicing, ultrasonic deicing and microwave deicing and other methods currently are still under research and not put into practice yet.

High-voltage transmission line deicing robot removes the icing on transmission line automatically, reduces the effect of icing damage on transmission lines and has fairly good application prospect and practical value. Currently both domestic and abroad have very few researches on deicing robot. Existing studies include the HQ LineRover remote control robot designed by Serge Montambault and others from the Hydropower Research Institute in Quebec, Canada [10]. He used sharp ice skates to impact and compress the icing on high-voltage transmission lines, which could remove the icing on the lines between towers but not capable of obstacle negotiation. In China, Sanxia University did research on the structure of deicing robot and proposed a rotary tool and rolling steel wheel combined pneumatic and telescopic deicing device [1, 11]; He Xiaohong and others from Lanzhou Technology University studies the obstacle negation control strategy of transmission line deicing robot [12, 13]; Hou Wenqi, Wang Jian, Ma Hongcu and others from National University of Defense and Technology implemented stimulation analysis on the body dynamics of deicing robots [14]; Wang Chao and others from Beijing University of Posts and Telecommunications proposed a compressing wheel pushing deicing robot [15]. Currently, most domestic studies of deicing robots are only theoretical. The deicing robot in China currently is mainly mechanical.

Based on the further study of high-voltage transmission line deicing environment, this paper proposed a wheel milling deicing equipment and a compound deicing robot that can have double wheel milling and compressing wheel rolling simultaneously. With the robot's climbing and obstacle negotiating ability, the wheel milling deicing equipment is designed on the executive arm of this deicing robot so that it would performance obstacle negotiation and deicing.

## 2 Deicing Environment and Deicing Devices

Overhead high-voltage transmission line is the main medium of electricity power transmission, with high-altitude suspended structure as its main feature: high current flows through the lines, surrounded by high-voltage electric field and magnetic field; furthermore, the lines have certain tension and flexibility. High-voltage transmission lines are exposed in the wild mostly. In winter, under the influence of low temperature, humidity and wind, dense glaze is easily formed around the lines. Glaze doesn't come off easily, and in snowing and raining weather, it forms icing and the icing would increase gradually [8]. Fig. 1 is the typical condition of icing on lines. According to on-site measurement and analysis on relevant documents, icing on lines has the following prominent feature [2]: 1) icing is semi-transparent, dense with no pore, hard, smooth, humid, oval shaped, its long and short sleeves have little different, easy to attach, typical glaze. 2) Icing has large density, usually no smaller than  $0.80 \text{ g/cm}^3$

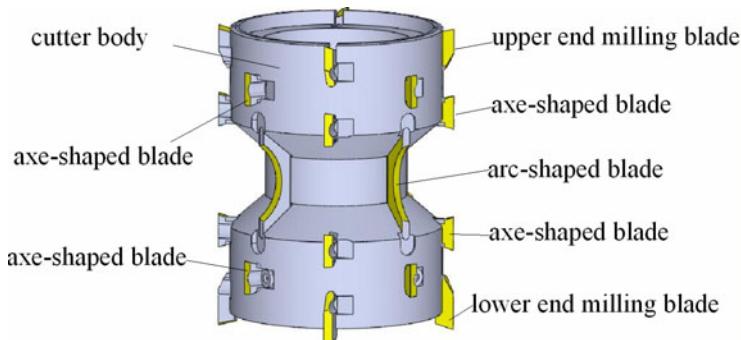
and reach as high as 0.92 g/cm<sup>3</sup>. 3) The thickness of icing is no smaller than 30mm, severe icing can be as thick as over 60mm. The icing on lines increase the burden of line towers, which results in the security risks of lines.



**Fig. 1.** Condition of Icing on Ground Wires

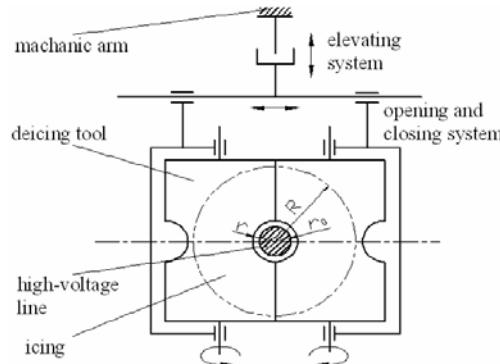
This paper proposed a wheel milling structured deicing device, the main bodies of which are two parallel copying milling ice tools. A sing deicing tool's external structure is shown in Fig. 2. The deicing tool is composed of cutter body and blades. The cutter body takes the shape of a rotary dumbbell, with several grooves evenly distributed for the installation of blades. The blades are of four different kinds: arc-shaped, axe-shaped, upper end milling blade and lower end milling blade. According to the position shown in the picture, place the blades on the cutter body. Arc-shaped blade is mainly used in copying milling to remove the icing near the high voltage lines; lower end milling blade is mainly used in obstacle negotiation to remove the icing on the other side of the obstacles, to do early preparations for the mechanic arm near high voltage lines; upper end milling blade is mainly to remove the icing near the upper end of milling blade; axe-shaped blade is used to remove the rest of the icing that the other three blades cannot remove.

Normally, two milling blades rapidly rotate in opposite direction and remove the icing around conduction wires through the scratching and colliding during the rapid movement. Normally, the high voltage line is at the center of the arc-shaped blade. Using milling blades that combine different blades can perform copying milling of the icing on high-voltage high voltage lines and remove the icing around the high-voltage lines efficiently.



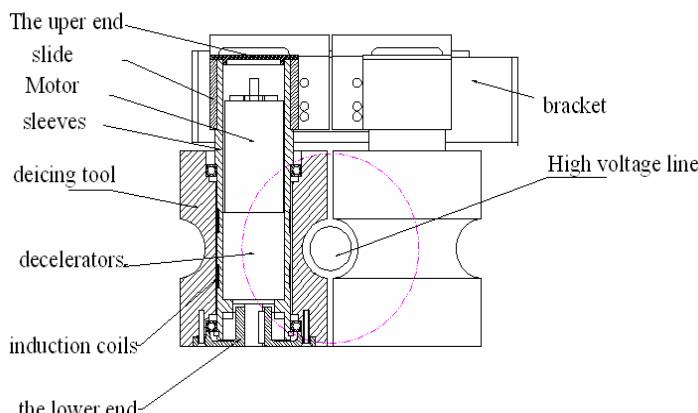
**Fig. 2.** Structure of Deicing Blade

Fig.3 shows a complete deicing device, which includes elevating system, opening and closing system, and deicing blade rotation system. Deicing elevating system is used to adjust the blade's relative position to high-voltage lines in the up and down direction. Deicing opening and closing system is used to adjust the cutting depth of deicing blades. The two blades rotate in opposite directions and perform high-speed copying milling.



**Fig. 3.** The mechanical structure of deicing device

The mechanical structure of deicing tools is shown in Fig.4. The deicing motors and decelerators are installed within the sleeves, which are connected with the stator of the motor decelerators; the sleeves are connected with the guide strip slides and nuts of the opening and closing system, and rolling bearings are installed between sleeves and deicing cutter body; the output shaft of motor decelerator is connected with the lower end, which is also connected with the cutter body. When the two deicing tools motors rotate respectively clockwise (CW) and counterclockwise (CCW),

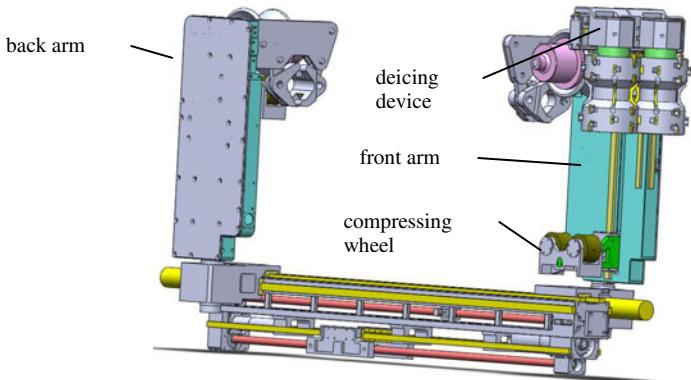


**Fig. 4.** Assemblage of Deicing Device

they drive the two blades to rotate CW and CCW through the lower end. This inlay structure satisfies the requirement of completely sealed motor during the deicing process. The opening and closing system is composed of CW and CCW screw nut pair, rolling guide strip slide pair and DC servo motor, which is that motor drives CW and CCW screw, and screw drives nuts and blades to move on rolling guide strips. On the sleeves, with the groove as axis center, two sets of induction coils are symmetrically winded in order to examine the magnetic field of high-voltage lines around the deicing tools and to determine the high-voltage lines' relative position to deicing tools, in case the blades scratch the high-voltage lines.

### 3 Structure of Deicing Robot

Deicing device can remove the icing on high voltage line, but not over the obstacles on the line. This paper proposed a deicing robot that can perform obstacle negotiation, the structure of which is shown in Fig.5.



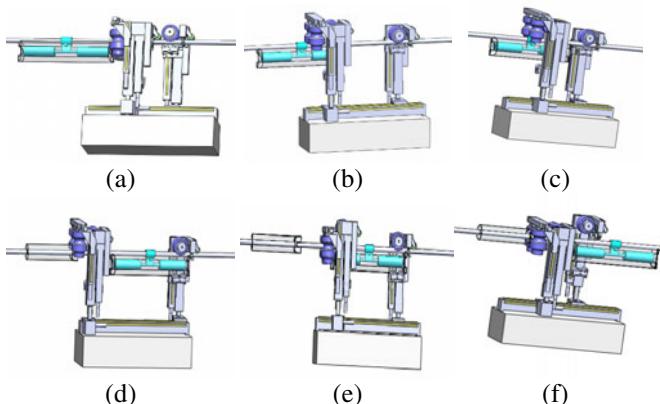
**Fig. 5.** The deicing robot

Deicing robot's structure is double-arm asymmetrical suspension, with one mechanical arm (front arm) installed with deicing device; every arm has an expansion joint and a rotation joint; on the body there is mobile guide, which allows for the closing and opening of two arms under the motor's drive; the freedom formed by these joints can allow the robot to move over the obstacles on the straight line (such as vibration damper and suspension clamp etc.). In order to climb with two arms and a single arm can be clamped when the obstacle negotiation, a pair of claws is installed at the end of each arm. In order to prevent slipping during running, there is a pair of compressing wheels with elastic buffer under the driving wheel of each arm. During deicing, the compressing wheel moves up and together with the road wheel compresses the high voltage line in case of slipping; Besides, a cooperative deicing working mode of back arm's claws clenching, front arm's road wheel and double arms' opening and closing system running at the same time, in order to increase thrust during deicing process; In order to improve the balance of the gravity center during the process of obstacle negotiation, another gravity center adjustment system is increase.

In the robot's deicing process, firstly the deicing device removes chunk of ice, and then driving wheel-compressing wheel roll over the remove the left smaller ice, which is a compound way of deicing.

## 4 The Solid Modeling and Simulation of Deicing Robot

The vibration damper and suspension clamp are the obstacles of robot's deicing process, and the robot has to change its own posture to cross over these obstacles. The purpose of solid modeling and simulation experiment is to implement simulation analysis on the designed theoretical model and verify the feasibility of the design of mechanic structure and obstacle negotiation. This paper used solidworks software to simulate deicing robot's structure and environment model. The simulated line model was LGJ400/50 high-voltage high voltage line, and vibration damper and suspension clamp were also simulated. A solid simulation modeling was implemented for deicing robot's crossing over vibration damper and suspension clamp, but because of the limited space of papers, this paper will only introduce the simulation process of robot crossing over vibration damper, which is shown in Fig.6.



**Fig. 6.** Simulation process of crossing over vibration damper

Before crossing over the obstacle, the robot closed its deicing tools, compressed its compressing wheels and removed the ice with a normal speed. It stopped in front of the vibration damper and was getting ready to cross over. The obstacle negotiation process can be broken down to the following steps (corresponding to the a-f in Fig.6):

- a)front arm's claw clenches, front compressing wheel compresses, which makes front arm fixed on the high voltage line, two arms close as much as they can;
- b)back arm's claw clenches, back compressing wheel compresses, front arm's claw and compressing wheel loosen up, blades open, front arm stretch to maximum length, move up on the deicing elevating system to the maximum, two arms open and the front arm's road wheel touches the icing on vibration damper;

c) back arm's compressing wheel fills up pressure, back arm's claw loosens, back arm's claw moves forward to the back of vibration damper, back arm's claw clenches;

d) both arms stretch, front arm's blades close, blades gradually descend and remove the icing at the lower end, both arms close and remove the icing at the lower end of front arm's road wheel;

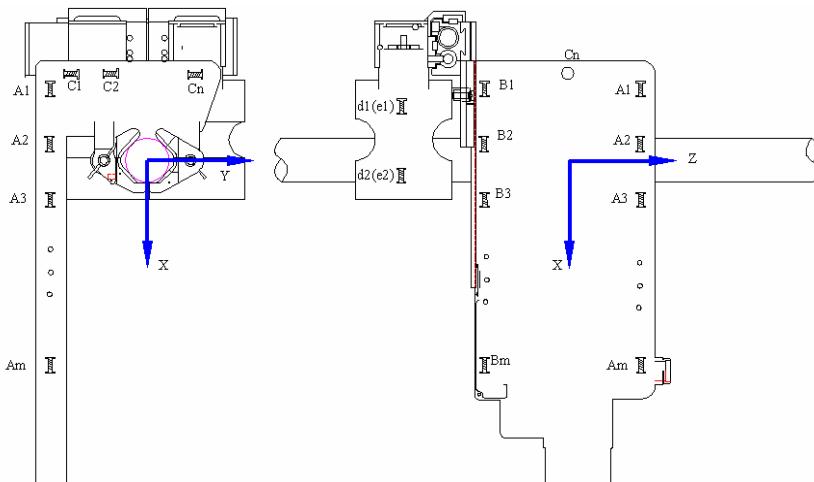
e) Deicing tools open, front arm shortens until road wheel touched the high voltage line, front arm's compressing wheel compresses, claw clenches, deicing tools close;

f) Back arm's claw loosens, compressing wheel loosens, back arm stretches; front arm's claw loosens, both arms close, road wheel moves forward, crosses over the vibration damper, simulation ends.

Through simulated obstacle negotiation experiment, it's shown that the designed structure parameters of the deicing robot are reasonable. Prototypes that follow the parameters design of the simulated structure can cross over the vibration damper.

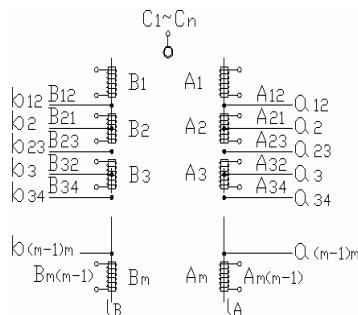
## 5 Electromagnetic Navigation Technology

When deicing robot is crossing over the obstacles, electromagnetic navigation technology is used to guide the arm to grasp or escape high-voltage lines without collision. The actually way is to put an array of electromagnetic sensors on its two mechanic arms, senses the magnetic field around the high-voltage lines through sensors, then the robot's control unit estimates the high-voltage line's relative position to the mechanic arms according to the collected signals of electromagnetic sensors, then constantly changes its posture to grasp or escape the high voltage line. The distribution of electromagnetic sensors on the front arm is shown in Fig. 7, with three groups



**Fig. 7.** Posture detection sensor model

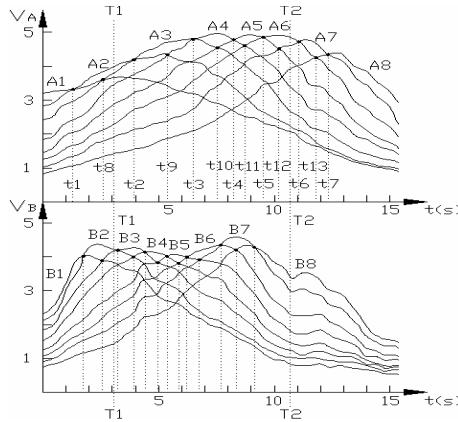
of sensors, which is group A, group B and group C. They are used to detect the high-voltage line's relative position to the mechanic arm and the other two groups of sensors on the blades group d and group e are used to detect the blades' relative position to the high voltage line. Group A, B and C sensors' centerlines are the same with their axes, which are  $l_A$ ,  $l_B$ ,  $l_C$ . Group A and B are distributed symmetrically, and their centerlines are parallel, and all the structure parameters of sensors are completely identical. Group A and B are used to detect if the plane of mechanic arm is parallel or vertical to the axis of high voltage line (Z direction). Group A or B sensors are used to detect the mechanic arm's relative x-direction position to the high voltage line, and Group C sensors are adopted to detect the mechanic arm's y-direction position to the high voltage line. The mechanic arm's relative posture to the high voltage line can be reflected to the working interval of sensors of all groups (Fig.8). Take Group A for example: Group A has m sensors which can determine  $a_{12}$ ,  $a_2$ ,  $a_{23}$ ,  $a_3$ , ...,  $a_{(m-1)m}$ , which is  $2m-3$  in total. The axes of sensors then are  $A_{12}$ ,  $A_{21}$ ,  $A_{23}$ ,  $A_{32}$ , ...,  $A_{m(m-1)}$ , which are separated into  $2m-2$  intervals.



**Fig. 8.** Division of intervals of posture detection sensors

The deicing robots in this paper use a posture detection sensor array of a total of 20 electromagnetic sensors with  $m=8$  and  $n=4$ . When the deicing robot's back arm fixes to the high-voltage wire, and the front arm lifts up, the front arm's Group A and B electromagnetic sensors' actual signals are shown in Fig.9.

Based on the intersection feature points of Group A sensor signals, corresponding time points  $t_1 \sim t_{13}$  can be labeled in Fig.9. ① at time  $t_1$ , the output curves of  $A_1$  and  $A_2$  cross, which means that the value of these two sensors are the same and the distance from the high voltage line to  $A_1$  and  $A_2$  is the same and is the closest, so

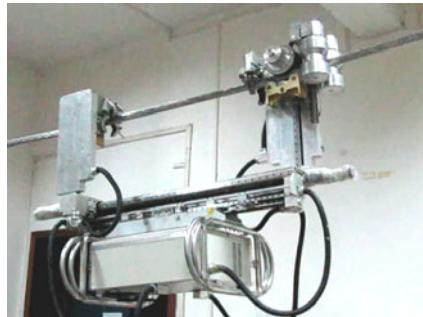


**Fig. 9.** Detected signals of electromagnetic sensors

at this time, the projection of high voltage line on  $l_A$  corresponds to the point  $a_{12}$  in Fig.8. By the same principle, it can be found that time point  $t_2$ 、 $t_3$ 、 $t_4$ 、 $t_5$ 、 $t_6$ 、 $t_7$  correspond with point  $a_{23}$ 、 $a_{34}$ 、 $a_{45}$ 、 $a_{56}$ 、 $a_{67}$ 、 $a_{78}$  in Fig.8 respectively, which is recorded as  $a_{ij}$ ,  $i = 1, 2, \dots, 7$ ,  $j = 2, 3, \dots, 8$ ,  $i \neq j$ . ② at time point  $t_8$ , the output value of  $A_2$  is bigger than that of  $A_1$  and  $A_3$  but the value of  $A_1$  and  $A_3$  is the same, meaning that the distance from high voltage line to  $A_2$  is the closest and is the same with the distance to  $A_1$  and  $A_3$ , so at this time the projection of high voltage line on  $l_A$  corresponds with point  $a_2$  in Fig.8. By the same principle, it can be found that time point  $t_9$ 、 $t_{10}$ 、 $t_{11}$ 、 $t_{12}$ 、 $t_{13}$  correspond with point  $a_3$ 、 $a_4$ 、 $a_5$ 、 $a_6$ 、 $a_7$  in Fig.8 respectively, which is recorded as  $a_k$ ,  $k = 2, 3, \dots, 7$ . ③ between time interval  $0 \sim t_1$ ,  $A_1$  has the biggest output value, is closest to the high voltage line, so the projection of high voltage line on  $l_A$  is in the  $A_{12}$  interval in Fig.8; between time interval  $t_1 \sim t_8$ ,  $A_2$  has the biggest output value and is closest to the high voltage line, and  $A_1$  has the second biggest output value and is second closest to the high voltage line, meaning that the projection of high voltage line on  $l_A$  is in the interval  $A_{21}$  in Fig.8; by the principle, it can be concluded that time interval  $t_8 \sim t_2$  corresponds with  $A_{23}$ , and  $t_2 \sim t_9$  corresponds with  $A_{32}$ 、...，and  $t_{13} \sim t_7$  corresponds with  $A_{87}$ ，which is recorded as  $A_{pq}$ ,  $p = 2, 3, \dots, 8$ ,  $q = 1, 2, \dots, 7$ ,  $p \neq q$ . By the same principle, the relative posture relationship between Group B,C sensors, deicing tool sensor and the high voltage line can be obtained.

## 6 Deicing Robot Prototype and Experiment

Deicing robot prototype is shown in Fig.10, which mainly includes body, front arm, back arm, deicing device, control chamber and other main parts. The prototype weighs 45 kg, the outside measurement is 900mm×500mm×700mm, deicing speed is 3m/min, and it can continuously work for more than 4 hours.



**Fig. 10.** Deicing robot prototype

### 6.1 Deicing Experiment

In order to verify the deicing performance of deicing robot, the simulated line deicing experiment was made in the experiment room. The straight steel pipe was adapted to simulated high-voltage line with a diameter of 35 mm and a length of 2.4m. Before the experiment, the steel pipe was put into a PVC pipe with 110mm in diameter and 1.2m in length. The PVC pipe was filled with water and then sealed. After sealed, it was put into an ice store for 48 hours. After freezing, the PVC pipe was removed, and there was a cylindrical icicle with a diameter of 110 mm and a density of  $0.91\text{g}/\text{cm}^3$  formed around the steel pipe. The iced steel pipe was put on the supporting shelf, and then the deicing robot was hanged on the un-iced section of the steel pipe. Then, the preparation for deicing experiment was completed.



**Fig. 11.** Walking deicing



**Fig. 12.** Pushing deicing

There are two parts of deicing experiment: one is walking deicing, and the other is pushing deicing.

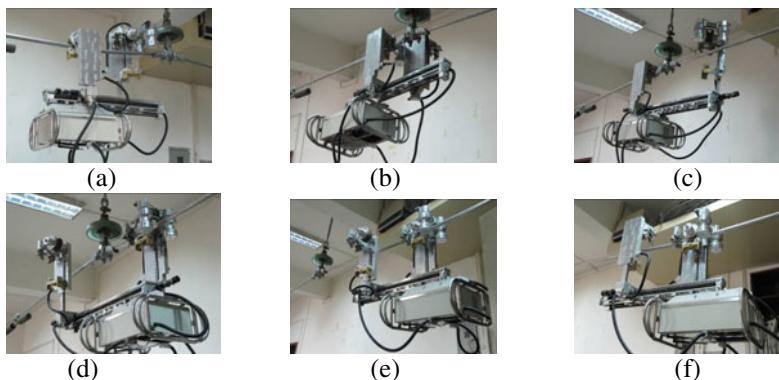
Walking deicing experiment is shown in Fig.11. Before the experiment, the height of the deicing tools was adjusted to a suitable position and compressed the compressing wheels of the front and back arms. The blades rotated with a speed of 450rpm and closed to its limit. The robot walking motor was switched on and it walked with a speed of 3m/min. When walking, the deicing tools gradually cut off the icing on the steel pipe, the crumbs splashed everywhere, and the robot's front arm was intensely shaking and huge chunks of ice cubes broke off and fell.

Pushing deicing experiment is shown in Fig.12. Before the experiment, the height of the blades was adjusted to a suitable position, and front arm and back arm closed to its minimum. The front arm's blade approached the bottom of icing, the compressing wheels of front arm and back arm compressed, back arm clenched to the steel pipe, the deicing blades rotated with a speed of 450rpm and closed to its limit, the front arm's road wheel motor and front arm back arm separation motor were switched on at the same time and pushed the front arm to move forward at 4m/min. When moving forward, the blades gradually cut off the icing on the steel pipe, the crumbs splashed everywhere, and the robot's front arm was intensely shaking and huge chunks of ice cubes broke off and fell. This is a highly efficient way of deicing.

Deicing experiment has shown that that it's feasible to use the compound deicing method of milling and compressing wheel rolling. It has also demonstrated the robot can remove icing by walking deicing and pushing deicing.

## 6.2 Operation Experiment on Non-icing Experimental Lines

In the experiment room, the high-voltage lines and matched fittings were installed on the shelf and compressed with certain tension. Limited by the experiment conditions, the robot has to run on non-icing high-voltage lines. This experiment was set to examine the obstacle negotiation of the suspension clamp on the direct high voltage line automatically. Before the experiment, the deicing robot was hanged on the straight line section of the high-voltage line suspension clamp. The high-voltage cable was put in 100A alternating current. According to planned movement control strategy, the deicing robot automatically began crossing over the obstacles, which is shown in Fig.13.



**Fig. 13.** Obstacle negotiation experiment of deicing robot prototype

In the experiment, the obstacle negotiation was totally automatic. There was no scratching of the high-voltage line and the mechanic arm's grasp of and escape from the high voltage line was very accurate and reliable. The experiment fully showed the feasibility of electromagnetic navigation technology leading robot in obstacle negotiation. Besides, the prototype's obstacle negotiation experiment also verified the accuracy of the deicing robot's structure model, which means that deicing robots can perform the crossing over of suspension clamps on direct high voltage line.

## 7 Conclusions

Based on the further study of high-voltage transmission line deicing environment, this paper proposed a wheel milling deicing equipment and a compound deicing robot that can have double wheel milling and compressing wheel rolling simultaneously; implemented deicing experiment and obstacle negotiation simulation analysis on the transmission and robot mechanic system; proposed an obstacle negotiation navigation control method based on electromagnetic sensors, and finally implemented deicing experiment and obstacle negotiation experiment on deicing robot prototypes. The experiment demonstrates the feasibility of electromagnetic navigation technology in leading robot in obstacle negotiation, and also shows that it's possible to use compound deicing method of milling and compressing wheel rolling. The study of this paper expands the methods of deicing robot and is worth of promoting in application.

## References

1. Yang, Y., Hongliang, G., Suiming, M., et al.: Structural Design of Overhead Line Deicing Robot. *Electric Power Construction* 30(3), 93–96 (2009)
2. Jingbo, Y., Zheng, L., Junke, H., et al.: Analysis of tower destroy about power transmission line in 2008. In: The Second (2009) National Overhead Power Transmission Line Exchange Meeting Symposium (2009)
3. Xiaoshu, S., qiu, N.: Discussion on Methods of De-icing for Overhead Transmission Lines. *High Voltage Engineering* 32(4), 25–27 (2006)

4. Ni, X., Zhao, J., Yang, M., et al.: Transmission line de-icing based on high-frequency resonance skin current. *Electric Power Automation Equipment* 29(2), 6–9 (2009)
5. Egbert, R.I., Schrag, R.L., Bernhart, W.D., et al.: Investigation of power line de-icing by electro-impulse methods. *IEEE Transactions on Power Delivery* 4(3), 1855–1861 (1989)
6. Huang, Q., Chen, Y., Zhang, C.: Design of an adaptive on-load de-icing scheme for overhead power transmission line. In: 2009 IEEE Power and Energy Society General Meeting (2009)
7. Gu, S., Chen, J., Cai, W., et al.: Experimental analysis and engineering designing of laser de-icing technology for transmission lines. *Gaodianya Jishu/High Voltage Engineering* 35(9), 2243–2249 (2009)
8. Landry, M., Beauchemin, R., Venne, A.: De-icing EHV overhead transmission lines using electromagnetic forces generated by moderate short-circuit currents. In: Proceedings of the IEEE International Conference on Transmission and Distribution Construction and Live Line Maintenance, pp. 94–100 (2000)
9. Yi, H., Ting, L.: The Features of Large Area Ice Damage of Power Network and The Prevent Measures, In: The Second, National Overhead Power Transmission Line Exchange Meeting Symposium (2009)
10. Montambault, S., Pouliot, N.: The HQ LineROVer: Contributing to innovation in transmission line maintenance. In: Proceedings of the IEEE International Conference on Transmission and Distribution Construction and Live Line Maintenance, ESMO, pp. 33–41 (2003)
11. Yi, Z., Wei, S., Hongliang, G., et al.: Mechanical Design of a Robot for Deicing Transmission Lines. *J. of China Three Gorges Univ.(Natural Sciences)* 30(6), 69–72 (2008)
12. Xiaopeng, L.: The Research and Design of the de-icing Robot of Transmission Line's Ice crossing the barrier methods. Lan Zhou university of technology (2006)
13. Xiaohong, H., Liang, G., Yuanyuan, M., et al.: Image processing technology applied in the vision system of the high - pressure deicing robot. *Industrial Instrumentation & Automation* (1), 106–108 (2010)
14. Wenqi, H., Jian, W., Hongxu, M., et al.: Design of Mechanical Configuration and Dynamics Simulation of Deicing Robot for Power Transmission Lines. *Machinery & Electronics* (10), 52–55 (2009)
15. Chao, W., ShiMin, W., Qizheng, L.: Design of Robot for Deicing High Voltage Transmission Lines. *Mechanical Engineering & Automationa* (1), 148–149 (2010)

# Projection on Convex Set and Its Application in Testing Force Closure Properties of Robotic Grasping

Li Miao, Yang Wenyu, and Zhang Xiaoping

State Key Lab of Digital Manufacturing Equipment and Technology,  
Huazhong University of Science and Technology, Wuhan 430074 P.R. China  
limiao712@yahoo.cn, mewyang@mail.hust.edu.cn,  
shoppinggre@hotmail.com

**Abstract.** Force closure is one of the most fundamental properties for a multi-contact robotic system to immobilize an object. Based on the projection of a point on a convex set, this paper presents an algorithm to compute the minimum distance from a point to a convex set in n-dimensional space. Further study of this projection algorithm, we found it can be used to test for closure properties of robotic grasping. Comparing with previous study, this method provides a clear physical interpretation and can be implemented very easily in practice. The superior efficiency and effectiveness of this algorithm are demonstrated by numerical examples.

**Keywords:** Multifingered grasp, Projection on convex set, numerical test, force closure.

## 1 Introduction

Force (form) closure is one of the most fundamental issues in multifingered grasping, dexterous manipulation and fixture design. It characterizes the capability of a grasp to completely restrain an object in the presence of external wrench or disturbances. For almost two decades, various tests and numerical algorithms for force closure and form closure have been investigated [1]-[20]. Reuleaux [1] firstly proposed the concept of force closure in the field of machine design. Salisbury [2] creatively introduced force closure into robotic grasping and proved that a grasping system is force closure only when the convex hull of the primitive contact wrench contains the neighborhood of the origin. As for a fixture workpiece system, Asada [3] proposed the relevant concepts of accessibility and detachability for optimal fixture design. Mishra *et al.* [4] and Markenscoff *et al.* [5] Lakshminarayana [6] derived the smallest number of frictionless contact points needed to construct a form (force) closure grasp. Bicchi [7] addressed the force closure problem as an ordinary differential equation and provided a framework for assessing the force closure properties based on Lyapunov's stability theory. Based on the duality between convex hull and convex polytopes, Liu [8] [9] formulated the force closure testing problem as a ray-shooting problem solved by the Simplex method. More recently, Buss *et al.* observed that the friction constraints can be represented by a semi-positive matrix [10]. Han *et al.* [11] further investigated the properties of the matrix and casted the force closure problem as a convex optimization problem involving linear matrix inequalities (LMIs) by checking the strictly internal

grasp force. As proved in [12], a grasp has force closure if and only if the grasp mapping  $G$  has full row rank and there exists an admissible strictly internal grasp force. Zhu *et al.* [13], [14], [15] used the Q distance and pseudo distance function to qualitative test and quantitative analysis of the force closure property. Xiong [19] constructed a J-function to detect force closure and stability for a fixture system, a detailed concepts and mathematical interpretation related to fixture system were also investigated. Based on the previous work of Liu [8], Yu *et al.* [17], [18], [19] used the distance algorithm for computing the intersection of the boundary of a compact convex with a ray emanating from an interior point of the set, which can be used in testing for force closure in robotic grasping. In [20], Yu further presented a finite convergence algorithm to compute the distance from a point to a convex cone and a comparison with many previous algorithms was also given, which demonstrated the efficiency and effectiveness of this algorithm by application to force closure test.

Based on a systematical study of previous works, especially [19], [20], this paper focuses on the problem of grasp analysis. We found connection between the force closure test and the orthogonal projection from a point to a convex set in n-dimensional Euclidean space. Projection on a convex set, especially on polyhedral cones is a problem of high impact on scientific community, which is still an open problem convex optimization [21]. For a force closure grasp, our algorithm needs finite iterations to obtain the result. For a non-force closure grasp, when the external wrench lies in the polar cone of the contacts wrench cone, our algorithm needs no iteration. Only when the external lies between the contacts wrench cone and its polar cone, the force closure test was formulated as a nonlinear optimization problem.

The rest of this paper is organized as follows. Section 2 presents a prerequisite about projection on a general convex set and a convex cone. The convex sets in multi-contact robotic system are studied in section 3 and the force closure test is formulated as a projection problem form a point to convex set. Simulation examples are given in Section 4. Section 5 concludes the paper with final remarks.

## 2 Projection Algorithm

Based on a systematical study of previous works of force closure test, we found many previous numerical algorithms [8], [9], [16], [17], [20], in fact, can be interpretation by projection on a convex set. In this section, we will discuss the algorithm to calculate the projection from a point to a convex set and cone, respectively.

### 2.1 Projection on Convex Set

Projection from a point in n-dimensional Euclidean space to a convex set, including subspace, affine set, convex cone, is defined as to find the minimum-distance between the point and the convex set and the corresponding point in the convex set that achieves the minimum distance. Convexity ensures the unique of the minimum distance [22].

Let  $P_C x$  denote projection of  $x$  on a closed convex set  $C \subseteq \mathbb{R}^n$  and  $P_C$  is the projection operator, obviously:

$$C = \{P_C x \mid x \in \mathbb{R}^n\} \quad (1)$$

The projection operator  $P_C$  is to find a point  $P_C x$  belongs to convex set  $C$  that is closest to point  $x$  in Euclidean norm, that is,

$$\|x - P_C x\|_2 = \inf_{y \in C} \|x - y\|_2 \quad (2)$$

Based on the dual interpretation of projection as a solution to a maximization problem [23]: Minimum distance from a point  $x \in \mathbb{R}^n$  to a convex set  $C \subseteq \mathbb{R}^n$  can be found by maximizing distance from  $x$  to hyperplane  $\partial H$  over the set of all hyperplanes separating  $x$  from  $C$ . Existence of a separating hyperplane presumes point  $x$  lies on the boundary or exterior to set  $C$ . Obviously, the optimal separating hyperplane should be the supporting hyperplane of set  $C$ , Fig. 1. Supposed  $x_0$  is a point on the boundary of set  $C$ , if  $a \neq 0$  satisfies  $a^T y \leq a^T x_0$  for all  $y \in C$ , then the hyperplane  $\partial H = \{y \in \mathbb{R}^n \mid a^T y = a^T x_0\}$  is called a supporting hyperplane to  $C$  at the point  $x_0$  and  $a \neq 0$  represents the normal direction of the hyperplane. The supporting function, i.e.,  $S_C : \mathbb{R}^n \rightarrow \mathbb{R}$  is defined as:

$$S_C(a) \triangleq \sup_{z \in C} a^T z \quad (3)$$

Then the supporting hyperplane with  $a$  as its normal direction can be represented by

$$\partial H(a) = \{y \in \mathbb{R}^n \mid a^T y = S_C(a)\} \quad (4)$$

The orthogonal projection of a point  $x$  on hyperplane  $\partial H$  is [22]:

$$P_{\partial H} x = x - a(a^T a)^{-1}(a^T x - S_C(a)) \quad (5)$$

The positive halfspace is denoted by, Fig. 2.

$$H_+ = \{y \in \mathbb{R}^n \mid a^T y \geq S_C(a)\} \quad (6)$$

Supposing  $x$  lies on the exterior to set  $C$ , as the optimal hyperplane will separate  $x$  from  $C$  and  $C \subset H_-$ , we have  $x \in H_+$ , then

$$\begin{aligned} \|x - P_C x\| &= \sup_{\partial H \ni x \in H_+} \|x - P_{\partial H} x\| \\ &= \sup_{a \ni x \in H_+} \|a(a^T a)^{-1}(a^T x - S_C(a))\| \\ &= \sup_{a \ni x \in H_+} \frac{1}{\|a\|} \|a^T x - S_C(a)\| \end{aligned} \quad (7)$$

Because the supporting function  $S_C(a)$  is a linear function of  $a$ , then the problem (7) can be formulated as a convex problem:

$$\begin{aligned} \|x - P_C x\| &= \max_a \text{imize } a^T x - S_C(a) \\ &\text{s.t. :} \|a\| \leq 1 \end{aligned} \quad (8)$$

The unique minimum-distance projection in convex  $C$  is therefore

$$P_C x = x - a^*(a^{*T} x - S_C(a^*)) \quad (9)$$

Where optimally  $\|a^*\| = 1$ .

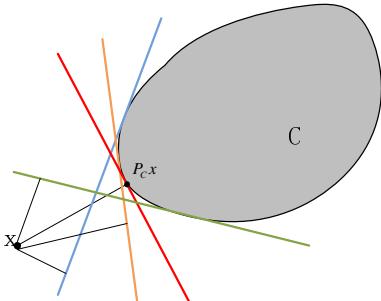


Fig. 1. Dual interpretation of projection [23]

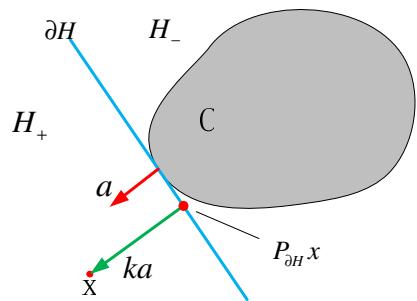


Fig. 2. Supporting hyperplane

However, the traditional projection only concerns the case that the point  $x$  lies on the exterior to set  $C$ . Here, for our application to force closure test in robotic grasping, we need to justify whether the point lie on the exterior to set  $C$ , though in this case, the projection is trivial and obvious. The following theorem can be used to check whether the point belongs to a convex set.

*Theorem 1:* Given a closed convex set  $C \subseteq \mathbb{R}^n$ , then a point  $x$  lies in  $C$  if and only if  $\max_a (a^T x - S_C(a)) \leq 0$ .

*Proof.* Necessity. If point  $x$  lies in  $C$ , from the definition of supporting hyperplane, we have  $x^T a \leq S_C(a)$ , obviously,  $\max_a (a^T x - S_C(a)) \leq 0$ .

Sufficiency. If there exist a point  $x$  that lying on the exterior to convex set  $C$  and  $\max_a (a^T x - S_C(a)) \leq 0$ , based on the separating hyperplane theorem, there exist a separating hyperplane  $\partial H_0 = \{y \in \mathbb{R}^n \mid a_0^T y = S_C(a_0)\}$  such that  $a_0^T x - S_C(a_0) > 0$ , which contradicts  $\max_a (a^T x - S_C(a)) \leq 0$ .  $\square$

*Projection Algorithm on Convex Set:* Given a convex set  $C \subseteq \mathbb{R}^n$  and a point  $x$ , we can obtain the unique minimum-distance projection by the solving following optimization problem:

*Step 1:* Set initial point  $a_0 \in \mathbb{R}^n$  and  $\varepsilon$  a small positive number as the termination tolerance.

*Step 2:* Solving the constrained nonlinear optimization problem:  $\max_a a^T x - S_C(a)$

$$\text{s.t. } \|a\| \leq 1$$

*Step 3:* If  $Pd = \max_a a^T x - S_C(a) < 0$ , that means the point  $x$  lies in  $C$ , then let

$Pd = \|x - P_C x\| = 0$  and  $P_C x = x$ , where  $Pd, P_C x$  is unique minimum distance and the corresponding projection in convex  $C$ , respectively; else, go to step 4.

*Step 4:* let  $P_C x = x - a^*(a^{*T} x - S_C(a^*))$  and  $Pd = \|a^{*T} x - S_C(a^*)\|$  and the algorithm stops.

Although there are some simple and efficient algorithms (such as Dykstra's algorithm [23]) to compute the minimum-distance projection, here, for the purpose of

emphasizing on robotic grasping analysis, we just use a general nonlinear optimization algorithm in step 2.

## 2.2 Projection on Cone

As a special case that the convex set  $C \subseteq \mathbb{R}^n$  is a cone, we can further simplify the projection algorithm when the point locates in the polar cone of  $C$ , Fig. 3. The polar cone is defined as:

$$C^\circ = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \leq 0, \forall x \in C\} \quad (10)$$

*Theorem 2:* Given a cone  $C \subseteq \mathbb{R}^n$ , then a point  $x$  lies in  $C^\circ$ ,  $x \in C^\circ$ , then the closest point to  $x$  in the cone is the origin, i.e.  $P_C x = 0$ ,  $Pd = \|x\|$ .

*Proof:* When  $x \in C^\circ$ , we have  $S_C(x) = \max_{z \in C} x^T z \leq 0$ , that means the supporting hyperplane  $\partial H(x) = \{y \in \mathbb{R}^n \mid x^T y = 0\}$  separates the cone  $C$  and the point  $x$ , then obviously that  $S_C(a^*) = 0$ . From (9), we have  $P_C x = x - a^* a^{*T} x = 0$  and  $Pd = \|x\|$ .  $\square$

Based on this property of cone, we can classify the location of the point with respect to the cone into three classes and therefore, give the simplified projection algorithm on cone.

*Projection Algorithm on Cone:* Given a cone  $C \subseteq \mathbb{R}^n$  and a point  $x$ , we can obtain the unique minimum-distance projection by the solving following optimization problem:

*Step 1:* Set initial point  $a_0 \in \mathbb{R}^n$  and  $\varepsilon$  a small positive number as the termination tolerance.

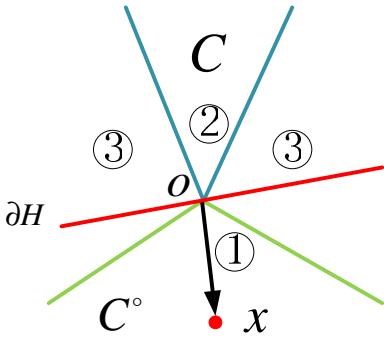
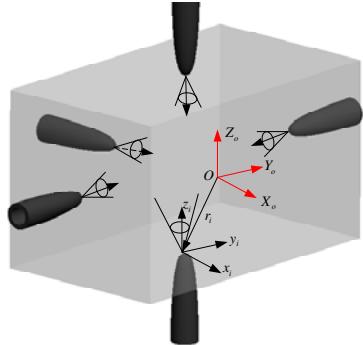
*Step 2:* If  $S_C(x) < 0$ , point  $x$  lies in area ①, let  $P_C x = 0$ ,  $Pd = \|x\|$  and the algorithm stops; else, go to step 3.

*Step 3:* Solving the constrained nonlinear optimization problem:  $\begin{aligned} \max_a & a^T x - S_C(a) \\ \text{s.t.} & \|a\| \leq 1 \end{aligned}$ .

*Step 4:* If  $Pd = \max_a a^T x - S_C(a) < 0$ , that means the point  $x$  lies in  $C$ , ②, then let  $Pd = \|x - P_C x\| = 0$  and  $P_C x = x$ ; else, go to step 4.

*Step 5:* Point  $x$  lies in area ③, let  $P_C x = x - a^*(a^{*T} x - S_C(a^*))$ ,  $Pd = \|a^{*T} x - S_C(a^*)\|$ , and the algorithm stops.

We should note here only in some cases that we need to solve the nonlinear optimization completely. When  $x$  lies in ①, we need no iteration to obtain the result; when  $x$  lies in ②, we can set the optimization termination tolerance  $\varepsilon$  larger to make terminate the iterations quickly, as the final result has nothing to do with  $\varepsilon$  in this circumstance; only when  $x$  lies in ③, we need to solve the optimization problem completely, however, this process can be finished very quickly using a general nonlinear optimization algorithm as there is only a norm constrain [22]. For a further

**Fig. 3.** Cone and polar cone**Fig. 4.** Multifingered grasp and contact cone

study, we would like to find some heuristics algorithm to improve the efficiency of our algorithm.

### 3 Contact Model and Force Closure Test

In this section, we will build a connection between the projection on convex set and the force closure test in multicontact robotic system, including multifingered grasp, fixture system, and biped robots.

#### 3.1 Frictional Constraints as Cone

Here we take the 3-D multifingered grasp as an example, Fig. 4. For a general multi-contact robotic system, the contact can be modeled as frictionless point contact (FPC), Point Contact with Friction (PCwF), soft finger contact with linear model (SFCI) and soft finger contact with elliptic model (SFCE). Let  $z_i$  denote the unit inward normal at contact  $i$  ( $i = 1, 2, \dots, m$ ), and  $x_i$  and  $y_i$  two unit tangent vectors. The contact force can be expressed in the local coordinate frame by  $f_i = [f_{i1}, f_{i2}, f_{i3}, f_{i4}]^T$ , where  $f_{i1}, f_{i2}$  and  $f_{i3}$  are the force components along  $z_i$ ,  $x_i$  and  $y_i$ , respectively, and  $f_{i4}$  is the torque about  $z_i$ . As is well known, to avoid separation and slide at contact, the contact force must comply one of the following contact constraints:

$$\text{FPC : } FC_i = \{f_i \mid f_{i1} \geq 0, f_{i2} = f_{i3} = f_{i4} = 0\} \quad (11)$$

$$\text{PCwF : } FC_i = \{f_i \mid f_{i1} \geq 0, \sqrt{f_{i2}^2 + f_{i3}^2} \leq \mu_i f_{i1}, f_{i4} = 0\} \quad (12)$$

$$\text{SFCI : } FC_i = \{f_i \mid f_{i1} \geq 0, \frac{\sqrt{f_{i2}^2 + f_{i3}^2}}{\mu_i} + \frac{|f_{i4}|}{\mu_{si}} \leq f_{i1}\} \quad (13)$$

$$\text{SFCE : } FC_i = \{f_i \mid f_{i1} \geq 0, \sqrt{\frac{f_{i2}^2 + f_{i3}^2}{\mu_i^2} + \frac{f_{i4}^2}{\mu_{si}^2}} \leq f_{i1}\} \quad (14)$$

Where  $\mu_i$  is the coefficient of friction and  $\mu_{si}$   $\mu'_{si}$  are the coefficient of torque for SFCI and SFCE, respectively.  $FC_i$  denotes the feasible set of the contact force, where is well known as the friction cone.

For a robotic grasping system, static force equilibrium of all forces exerted on the object implies that:

$$Gf = \sum_{i=1}^m G_i f_i = -w^{ext} \quad (15)$$

Where  $G = [G_1, G_2, \dots, G_m] \in \mathbb{R}^{6 \times 4m}$  is the grasp matrix mapping the force from contact space to object wrench space and

$$G_i = \begin{bmatrix} z_i & x_i & y_i & 0 \\ r_i \times z_i & r_i \times x_i & r_i \times y_i & z_i \end{bmatrix} \quad (16)$$

$f = [f_1^T, f_2^T, \dots, f_m^T]^T \in \mathbb{R}^{4m \times 1}$  is the contact force and  $w^{ext}$  is the external force exerted on the object.

As is well known in convex analysis, any closed convex set contain no lines can be expressed as the convex hull of its extreme points and extreme rays [23]. Then  $FC_i$  can be rewritten as

$$FC_i = cone(C_i) \quad (17)$$

Where  $C_i$  is the generator of the corresponding friction cone, which can be represented by

$$FPC : C_i = \{f_i \mid f_{i1} = 1, f_{i2} = f_{i3} = f_{i4} = 0\} \quad (18)$$

$$PCwF : C_i = \{f_i \mid f_{i1} = 1, \sqrt{f_{i2}^2 + f_{i3}^2} = \mu_i, f_{i4} = 0\} \quad (19)$$

$$SFCI : C_i = \{f_i \mid f_{i1} = 1, \frac{\sqrt{f_{i2}^2 + f_{i3}^2}}{\mu_i} + \frac{|f_{i4}|}{\mu_{si}} = 1\} \quad (20)$$

$$SFCE : C_i = \{f_i \mid f_{i1} = 1, \sqrt{\frac{f_{i2}^2 + f_{i3}^2}{\mu_i^2} + \frac{f_{i4}^2}{\mu'_{si}^2}} = 1\} \quad (21)$$

Then, we can obtain

$$G_i(FC_i) = G_i(cone(C_i)) = cone(G_i(C_i)) = cone(W_i) \quad (22)$$

Where  $W_i = G_i(C_i)$  is known as the primitive contact wrench [20] and  $G_i(FC_i)$  is called the feasible wrench cone of contact  $i$ .

Hence, (15) can be rewritten as

$$G(FC) = \sum_{i=1}^m cone(W_i) = cone(\sum_{i=1}^m W_i) = cone(W) \quad (23)$$

Where  $W = \sum_{i=1}^m W_i$  denotes the Minkowski sums of  $W_i$ .

Up to now, we have formulated the grasping mapping as a conic combination of the sum of the primitive wrench. In the following section, we will provide a numerical criterion to test the force closure property of a 3-D grasp.

### 3.2 Force Closure Test

A 3-D grasp has the force closure property means that it can resist any wrench in  $\mathbb{R}^6$ . Mathematically, a grasp is said to be force closure if and only if  $G$  is surjective and there exist a contact force that  $f_N \in \ker(G) \cap \text{int}(FC)$  [12], or equally,  $G(FC) = \mathbb{R}^6$  or  $\mathbb{R}^3$  for 3-D grasp and planar grasp, respectively. According to the conic expression of the feasible wrench set (23), the force closure can be rewritten as the following proposition [20].

*Proposition 1:* A 3-D robotic grasp is said to be force closure if and only if it satisfies the following two conditions:

- (1)  $\text{rank}(G) = 6$  ;
- (2) There exists a point  $w_e$  such that  $w_e \in ri\{\text{cone}(W)\}$  and  $-w_e \in \text{cone}(W)$ , where  $ri\{\bullet\}$  denotes the relative interior of a set.

*Proof:* This proposition is a generalized of proposition 1 in [20] and can easily be derived from force closure definition in [2].

The condition  $\text{rank}(G) = 6$  is very easy to verify, in order to test the force closure property, our main effort will focus on checking the second condition. What is more,

a point  $w_e \in ri\{\text{cone}(W)\}$  is also easy to find, such as  $w_e = \sum_{i=1}^m G_i f_i$ , where

$f_i = [1 \ 0 \ 0 \ 0]^T \in ri(FC_i)$ . The only difficult lies in that there was no simple and efficient method to check  $-w_e \in \text{cone}(W)$ . We found that the orthogonal projection of a point on convex set provide a geometric method to check  $-w_e \in \text{cone}(W)$ , which are summarized in the following proposition.

*Proposition 2:* The following statements are true:

- (1)  $x \in \text{cone}(W) \Leftrightarrow P_{\text{cone}(W)}(x) = x$  ;
- (2)  $x \in \text{cone}(W) \Leftrightarrow Pd = 0$ , where  $Pd$  is the unique distance between  $x$  and  $\text{cone}(W)$  ;
- (3)  $x \in \{\text{cone}(w)\}^* \Leftrightarrow P_{\text{cone}(w)}(x) = 0$ , where  $\{\text{cone}(w)\}^*$  denotes the polar cone of  $\text{cone}(w)$  ;
- (4)  $x \in \{\text{cone}(w)\}^* \Leftrightarrow Pd = \|x\|$ .

As discussed in section 2, the supporting function  $S_C(a) \triangleq \sup_{z \in C} a^T z$  plays an important role in implementing the projection algorithm. Therefore, we would like to indicate how to calculate the supporting function based on the following two properties when applying in a robotic grasping system.

*Theorem 3:*

$$(1) S_{A_1 + A_2 + \dots + A_m}(a) = \max_{i=1,2,\dots,m} S_{A_i}(a), \text{ where } A_1, A_2 \dots A_m \text{ are convex set.}$$

$$(2) S_{G(A)}(a) = S_A(G^T a), \text{ where } G \text{ is a real matrix.}$$

*Proof.* (1) According to definition of supporting function, we have the following derivation:

$$S_{A_1+A_2+\dots+A_m}(a) = \sup_{z \in A_1+A_2+\dots+A_m} a^T z = \sup_{z \in A_1 \cup A_2 \cup \dots \cup A_m} a^T z = \max_{i=1,2,\dots,m} S_{A_i}(a).$$

The second equality comes from the fact that an affine function defined in a convex set can achieve its extreme value only in the extreme point of the convex set.

(2) Similarly,  $S_{G(A)}(a) = \sup_{z \in G(A)} a^T z = \sup_{z=GK, K \in A} a^T GK = S_A(G^T a)$  where  $K$  is generator of  $A$ .  $\square$

From (22) and (23), we have:

$$S_{G(FC)}(a) = \max_i S_{cone(W_i)}(a) = \max_i S_{C_i}(G_i^T a) \quad (24)$$

Hence, for a given grasp configuration  $G$ , we only need to calculate  $S_{C_i}(G_i^T a)$  for each contact, denote  $b_i = G_i^T a = [b_{i1} \ b_{i2} \ b_{i3} \ b_{i4}]^T$ , we have:

$$\text{FPC: } S_{c_i}(b_i) = b_{i1} \quad (25)$$

$$\text{PCwF: } S_{c_i}(b_i) = b_{i1} + \mu_i \sqrt{b_{i1}^2 + b_{i2}^2} \quad (26)$$

$$\text{SFCl: } S_{c_i}(b_i) = b_{i1} + \max\{\mu_i \sqrt{b_{i1}^2 + b_{i2}^2}, \mu_{si} |b_{i4}| \} \quad (27)$$

$$\text{SFCE: } S_{c_i}(b_i) = b_{i1} + \sqrt{\mu_i^2 (b_{i1}^2 + b_{i2}^2) + \mu_{si}^2 b_{i4}^2} \quad (28)$$

Then, the force closure test can be done by the following algorithm:

*Force Closure Test:* Given a grasp,  $G = [G_1, G_2, \dots, G_m]$ ,  $C = [C_1, C_2, \dots, C_m]$ , we can check the force closure property of the grasp immediately.

Step 1: Compute  $\text{rank}(G)$ , if  $\text{rank}(G) = 6$ , go to step 2; else, non force closure and the algorithm stops;

Step 2: Compute  $w_e = \sum_{i=1}^m G_i f_i$  where  $f_i = [1 \ 0 \ 0 \ 0]^T$  and  $S_{G(FC)}(-w_e)$  from (24)-(28);

Step 3: Compute  $P_{G(FC)}(-w_e)$  and  $Pd$  by using the Projection Algorithm on Cone;

Step 4: If  $Pd = 0$ , force closure; else, non force closure and algorithm stops.

From the above algorithm, we note that  $S_{G(FC)}(-w_e)$  has an analytical expression so that step 1 and step 2 cost little time. The time will mostly be used in the Projection Algorithm on Cone. What is more, we found that  $Pd$  can be used as a quantitative grasp metric indicating how far a grasp departing from a force closure grasp. In the next section, we will provide some numerical examples to demonstrate the efficiency of our method.

## 4 Numerical Example

We provide some numerical examples to verify the efficiency of our algorithm. The algorithm was tested using MATLAB on laptop with Pentium-M 1.50GHz and 512 RAM.

*Example 1:* the first numerical example comes from the example used by [20]. There are two types of four-fingered grasps on a gold ingot. Their contact position and unit inward normal are as follows:



Fig. 5. Gold ingot gripped by four (a) PCwF (b) SFC

$$\text{Grasp (a): } r_1 = [0, -2, 3\sqrt{3}/4]^T, z_1 = [0, \sqrt{3}/2, 1/2]^T$$

$$r_2 = [3/4, 0, \sqrt{3}/2]^T, z_2 = [-\sqrt{3}/2, 0, 1/2]^T$$

$$r_3 = [0, 2, 3\sqrt{3}/4]^T, z_3 = [0, -\sqrt{3}/2, 1/2]^T$$

$$r_4 = [-3/4, 0, \sqrt{3}/2]^T, z_4 = [\sqrt{3}/2, 0, 1/2]^T$$

$$\text{Grasp (b): } r_1 = [0, -3/2, \sqrt{3}/2]^T, z_1 = [0, \sqrt{3}/2, 1/2]^T$$

$$r_2 = [0, 3/2, \sqrt{3}/2]^T, z_2 = [0, -\sqrt{3}/2, 1/2]^T$$

$$r_3 = [3\sqrt{3}/8, 0, 3\sqrt{3}/4]^T, z_3 = [-3/5, 0, -4/5]^T$$

$$r_4 = [-3\sqrt{3}/8, 0, 3\sqrt{3}/4]^T, z_4 = [3/5, 0, -4/5]^T$$

From the projection algorithm in section 3, we obtain:

$$\text{Grasp (a): } \text{rank}(G) = 6 \text{ and } Pd = 2, P_{G(FC)}(-w_e) = [0, 0, 0, 0, 0, 0]^T;$$

$$\text{Grasp (b): } \text{rank}(G) = 6 \text{ and } Pd = 0, P_{G(FC)}(-w_e) = [0, 0, 0.6, 0, 0, 0]^T;$$

Then grasp (a) is non force closure while grasp (b) is force closure. The average CPU time for grasp (a), (b) is 0.210ms, 82.76ms respectively.

*Example 2:* Consider a 3-D grasp on a ceramic bowl with four PCwF fingers, Fig. 6. The positions of the four fingers are fixed while the normal direction of the fourth fingers can change in the Y-Z plane. The angle between the normal of the fourth finger and the positive  $Y_4$  direction is denoted by  $\alpha$ . Then the position and inward normal of the four fingers are as follows:

$$r_1 = [5, 0, 5\sqrt{3}/3]^T, z_1 = [-1/2, 0, \sqrt{3}/2]^T$$

$$r_2 = [-5, 0, 5\sqrt{3}/3]^T, z_2 = [1/2, 0, \sqrt{3}/2]^T$$

$$r_3 = [0, 8, 8]^T, z_3 = [0, \sqrt{2}/2, -\sqrt{2}/2]^T$$

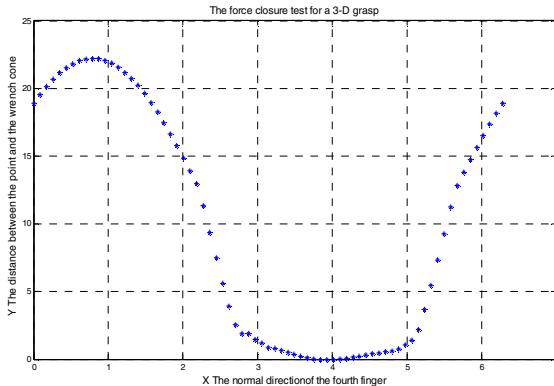
$$r_4 = [0, -8, 8]^T, z_4 = [0, \cos(\alpha), \sin(\alpha)]^T$$

As  $\alpha$  varies over the range  $[0, 2\pi]$ , the force closure property of the grasp are tested used our proposed numerical algorithm, which are shown in Fig. 7. It is observed that

the grasp achieves force closure property if  $\alpha$  lies in a small range centered at 4 rad. In this example, we test 73 different grasps and the average CPU time is 166.4ms. We note here that we did not refer to some simple and efficient algorithms (such as the famous Dykstra's algorithm [23]) to compute the minimum-distance projection. We just used the function **fmincon** in MATLAB to obtain the minimum-distance projection. Comparing with other methods [8], [14], [16], [20], our method make a connection between the projection on convex set and force closure test, which can be easily further developed with application to some multicontact robotic system.



**Fig. 6.** Ceramic bowl grasped by four PCwF fingers



**Fig. 7.** The force closure property varies with the fourth finger normal

## 5 Conclusion

This paper presents a numerical algorithm to the force closure test based on projection on convex set. We found that a minus projection distance is obtaining when the point lies in convex set, this property can be used as a numerical criterion for qualitative test of force closure. The proposed numerical test can be calculated by solving a single norm constrained convex optimization problem, which is very easy to implement in practice. We also note that some specific projection algorithm (alternative projection algorithm, Dykstra's algorithm) can be imbedded in our numerical test to improve its computation efficiency, which is one of our future research topics.

**Acknowledgments.** This work was supported by National Basic Research Program of China (973 Program), No. 2006CB705404.

## References

1. Reuleaux, F.: *The Kinematics of Machinery*. Macmillan, New York (1876)
2. Salisbury, J.K., Roth, B.: Kinematic and force analysis of articulated hands. *ASME J. Mech., Transmissions, Automat., Des.* 105(1), 35–41 (1983)
3. Asada, H., By, A.B.: Kinematic analysis of workpart fixturing for flexible assembly with automatically reconfigurable fixture. *J. Robot* 1(2), 86–94 (1985)
4. Mishra, B., Schwartz, J.T., Sharir, M.: On the existence and synthesis of multifinger positive grips. *Algorithmica* 2, 541–558 (1987)
5. Markenscoff, X., Ni, L., Papadimitriou, C.H.: The geometry of grasping. *Int. J. Robot. Res.* 9, 61–74 (1990)
6. Lakshminarayana, K.: In Mechanics of Form Closure, ASME paper 78-DET-32 (1978)
7. Bicchi, A.: On the closure properties of robotic grasping. *Int. J. Robot. Res.* 14(4), 319–344 (1995)
8. Liu, Y.H.: Qualitative test and force optimization of 3D fictional force closure grasps using linear programming. *IEEE Trans. Robot. Autom.* 15(1), 163–173 (1999)
9. Liu, Y.H., Lam, M.-L., Ding, D.: A complete and efficient algorithm for searching 3-D form-closure grasps in the discrete domain. *IEEE Trans. Robot.* 20(5), 805–816 (2004)
10. Buss, M., Hashimoto, H., Moore, J.: Dextrous hand grasping force optimization. *IEEE Trans. Robot. Automat.* 12, 406–418 (1996)
11. Han, L., Trinkle, J.C., Li, Z.X.: Grasp analysis as linear matrix inequality problems. *IEEE Trans. Robot. Automat.* 16, 663–674 (2000)
12. Murray, R.M., Li, Z.X., Sastry, S.S.: *A Mathematical Introduction to Robotic Manipulation*. CRC Press, Boca Raton (1994)
13. Zhu, X.Y., Wang, J.: Synthesis of force-closure grasps on 3-D objects based on the Q distance. *IEEE Trans. Robot. Automat.* 19(4), 669–679 (2003)
14. Zhu, X.-Y., Ding, H., Wang, Y.: A numerical test for the closure properties of 3D grasps. *IEEE Trans. Robot. Automat.* 20(3), 543–549 (2004)
15. Zhu, X.Y., Ding, H., Tso, S.K.: A pseudodistance function and its applications. *IEEE Trans. Robot. Automat.* 20(2), 344–352 (2004)
16. Xiong, Y., Xiong, X.: Algebraic structure and geometric interpretation of rigid complex fixture systems. *IEEE Trans. Robot. Autom.* 4(2), 252–264 (2007)
17. Zheng, Y., Qian, W.-H.: Simplification of the ray-shooting based algorithm for 3-D force-closure test. *IEEE Transactions on Robotics* 21(3), 470–473 (2005)
18. Zheng, Y., Qian, W.-H.: An enhanced ray-shooting approach to force-closure problems. *J. of Manufacturing Science and Engineering* 128(4), 960–968 (2006)
19. Zheng, Y., Chew, C.-M.: A numerical solution to the ray-shooting problem and its applications in robotic grasping. In: Proceedings of IEEE International Conference on Robotics and Automation, Kobe, Japan, pp. 2080–2085 (May 2009)
20. Zheng, Y., Chew, C.-M.: Distance between a point and a convex cone in n-dimensional space: computation and application. *IEEE Trans. Robot. Autom.* 25(6), 1397–1412 (2009)
21. See the popularity of the Wikimization page Projection on Polyhedral Cone, <http://www.convexoptimization.com/wikimization/index.php/Special:PopularPages>
22. Boyd, S., Vandenberghe, L.: *Convex Optimization*. Cambridge University Press, Cambridge (2004)
23. Dattorro, J.: *Convex Optimization & Euclidean Distance Geometry*. Meβoo (2005)

# Buffering Indices for Robot Manipulators Based on the Energy Distribution

Hao Wang<sup>1,2,\*</sup>, Zhongqin Lin<sup>1,2</sup>, Kai Zhao<sup>1</sup>, and Genliang Chen<sup>1</sup>

<sup>1</sup> City Key Laboratory of Automobile Engineering of Shanghai, Shanghai Jiao Tong University,  
Shanghai, 200240, P.R. China  
Tel.: +8621 3420 6786; Fax.: +8621 3420 4542  
[wanghao@sjtu.edu.cn](mailto:wanghao@sjtu.edu.cn)

<sup>2</sup> State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University,  
Shanghai, 200240, P.R. China  
[{wanghao,zqlin,kzhao,leungchan}@sjtu.edu.cn](mailto:{wanghao,zqlin,kzhao,leungchan}@sjtu.edu.cn)

**Abstract.** This paper proposes the buffering indices for evaluating the capacity of robot manipulators to comply with unpredictable and heavy external loads. According to the roles of mechanical components, a robot is divided into two parts: the buffering structure and the protected structure. The proposed indices are based on the energy distribution between these parts when they are in contact with the environment. Two indices are proposed: the Static Buffering Index (SBI) which reveals the ratio of the energy distributed in the buffering structure to the whole system under a specified configuration and the Kineto-static Buffering Index (KBI) reflects energy distribution along a predefined trajectory. The general stiffness mapping of robots with both flexible actuators and linkages is developed for the calculation of these indices. In the end, a heavy forging manipulator is studied as an example, and some numerical results are provided to demonstrate the invariance of the indices.

**Keywords:** Buffering indices, robot manipulator, energy distribution, stiffness mapping.

## 1 Introduction

Robots have been traditionally used as positioning devices without much regard to external forces. Compliance endows robots with the ability to perform the task in constrained work space, such as the surface machining [1,2], industrial assembly [3,4], etc. In these robotic tasks, the fast and precise manipulation is of main concern. So the buffering capability plays an important role on the modification of the robot trajectory based on the contact forces.

However, not all the robotic tasks put the positioning accuracy in the first place. For heavy manipulators working under extreme environment, the ability to buffer external loads acted on the end-effector is in the most important place. Furthermore, the buffering for impact during the working process will assure both the continuous

---

\* Corresponding author.

operation and safety for robots. Moreover, the biped robots make use of buffering to improve the landing performance by reducing the contact energy cost. In these cases, the buffering of contact forces is more important than the positioning accuracy of the end-effector for the following three reasons. First of all, the position of the end-effector doesn't need accurate control, which means the position error in a limited range is acceptable. Secondly, the contact force has exceeded load driving ability. So actuators are not able to make compensation to the position error of the end-effector. Last but not least, the external loads transmitted to the mechanical bodies have exceeded their intensities, which will cause serious damage to the equipment. So the positioning accuracy has to make concessions to the buffering ability according to the requirements of these robotic tasks.

For the purpose of positioning accuracy, the active compliance is widely used. But it will suffer from the slow response if we consider the short process of the buffering for contact forces. Since the passive compliance often provides faster and more reliable responses for dynamic contact, it is a better choice here. The passive compliance makes use of the natural compliance inherent in robotic structures to provide the compliance effect required, such as flexible elements [5,6], dissipative elements, topology-varying mechanism [7], etc. The most famous one is the RCC device invented by Whitney [8]. Then developed RCC mechanisms are proposed to adapt different working conditions [9,10]. Some mechanisms also have compliance characteristics and part of them can be used to realize the buffering for a specific robotic task [11-14].

For the realization of buffering, the compliance matrix is the key factor for both the design of passive compliance and the control of active compliance. The initial researches are limited in the investigation of the structure of compliance matrices via several methods, such as screw theory and Lie algebra [15-17]. Then some researchers propose the procedure to synthesize realizable compliance or stiffness matrix via simple springs connected in serial or parallel, such as Ciblak and Lipkin [18], Roberts [19], and Huang and Schimmels [20-22].

However, the compliance matrix can't directly reflect the ability for robots to buffer external loads because it only reveals the static relation between external loads and the deformation of the end-effector for robot manipulators. Furthermore, there're several disadvantages if we use matrix to establish an index, such as the relativity to the coordinate system, indefinitely physical meaning, etc. So the need to establish proper buffering indices has become quite obvious as the goal for the design, control and optimization of robots to comply with heavy external loads. This paper presents the buffering indices based on the energy distribution to evaluate the buffering ability for manipulators to comply with excessive external loads during working process.

## 2 Buffering Indices for Robot Manipulators

The buffering of robot manipulators is realized by the flexible elements such as links, joints, spring and dissipative elements as damping, as well as the topology-varying mechanisms to buffer excessive external loads. The elastic links, joints and spring can store and release the energy while the damping dissipates the energy to prevent the important structure of manipulators from being damaged. Additionally, manipulators

with topology-varying mechanism can change the energy flow to realize the buffering effect. Then according to the modular concept, a manipulator can be divided into two kinds of parts: one part with flexible and dissipative elements, and topology-varying mechanism as the buffering structure; the other as the protected structure.

Two buffering indices are proposed based on the ratio of energy distributed in the buffering structure to the whole system. The Static Buffering Index (SBI) is used to evaluate the buffering ability of manipulators at a certain configuration whereas the Kineto-static Buffering Index (KBI) reflects the ability to buffer external loads along a predefined trajectory. Since the buffering structure is used for absorbing, dissipating and changing the energy caused by external loads to protect the manipulators, the higher percent of energy distributed in this structure means better buffering ability.

## 2.1 Static Buffering Index (SBI)

In static condition, the total work for a manipulator with N bodies can be expressed as:

$$\Delta W = \sum_{i=1}^N (\mathbf{F}_i \cdot \Delta \mathbf{r}_{ic} + \mathbf{M}_i \cdot \Delta \boldsymbol{\varphi}_{ic}) = 0 \quad (1)$$

where  $\mathbf{F}_i$  and  $\mathbf{M}_i$  are the applied force and moment exerted on each body of the manipulator,  $\mathbf{r}_{ic}$  and  $\boldsymbol{\varphi}_{ic}$  are the linear and angular displacement vectors of the mass center for each body of the manipulator.

Suppose all the external loads are acted on the end-effector and can be simplified as a principal vector and moment at a reference point D. Then the work of external loads will be balanced by energy flow inside the manipulator  $\Delta E'$  as:

$$\Delta W = \Delta \mathbf{D}^T \hat{\mathbf{F}}_e + \Delta E' = 0 \quad (2)$$

where  $\hat{\mathbf{F}}_e = [\mathbf{F}_e, \mathbf{M}_e]^T$  represents the generalized applied forces on the bodies. And  $\Delta \mathbf{D}$  is the displacement of point D from an equilibrium position.

Since the robot manipulator consists of the buffering structure and the protected structure, the energy flow inside the manipulator can be divided into two parts:

$$\Delta E' = \Delta E_p + \Delta E_b \quad (3)$$

where  $\Delta E_p$  is the energy of the protected structure and  $\Delta E_b$  is the energy of the buffering structure.

The Static Buffering Index  $\eta$  is defined as the ratio between the energy of the buffering structure and the whole energy flow:

$$\eta = \frac{\Delta E_b}{\Delta E'} \quad (4)$$

It is obvious that the value of the SBI is between zero and one. If it approaches one, the manipulator has better buffering property because the energy mainly concentrated in the buffering structure, which means the remaining structure is well protected.

Now we're going to explain the physical meaning of the SBI from the viewpoint of energy and work. The energy flow inside the manipulator can be rewritten by:

$$\Delta E' = \Delta W_{nc} + \Delta W_c \quad (5)$$

where  $\Delta W_c$  is the work of conservative force and  $\Delta W_{nc}$  is the remaining part. The work of conservative force has the relationship with the potential energy  $V$  under the static condition by:

$$\Delta W_c = -\Delta V \quad (6)$$

The work of remaining part will cause the energy dissipation  $E$ , so we have to introduce the dissipative function  $\psi_q$  to express it:

$$\frac{dE}{dt} = -2\psi_q \quad (7)$$

So the term  $\Delta W_{nc}$  can be expressed as:

$$\Delta W_{nc} = -\Delta E \quad (8)$$

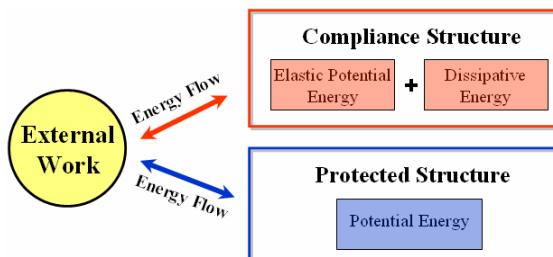
Then, Eq. (2) can be rewritten by:

$$\Delta W = \Delta \mathbf{D}^T \hat{\mathbf{F}}_e + \Delta E' = \delta \mathbf{D}^T \hat{\mathbf{F}}_e - \Delta V - \Delta E = 0 \quad (9)$$

The work caused by the deformation of flexible elements can be expressed in the form of elastic potential energy  $V_b$  while dissipative energy  $\psi_q$  reflects the work of damping. Suppose the buffering structure has  $N_1$  degrees of freedom (DOF), then the energy of the buffering structure  $\Delta E_b$  and the protected structure  $\Delta E_p$  can be rewritten in the form of generalized coordinate  $q_j$  as:

$$\Delta E_b = -\sum_{j=1}^{N_1} \left[ \frac{\partial V_b}{\partial q_j} + \frac{\partial \psi_q}{\partial \dot{q}_j} \right] \Delta q_j, \quad \Delta E_p = -\sum_{j=1}^{N_2} \frac{\partial V_p}{\partial q_j} \Delta q_j \quad (10)$$

So under the static condition, the energy flow will be stored by elastic potential energy and/or dissipated by dissipative energy in the buffering structure to protect the important structure. This energy distribution can be illustrated in Fig. 1.



**Fig. 1.** Energy distribution of robot manipulators under static condition

## 2.2 Kineto-static Buffering Index (KBI)

When a robot manipulator is in motion, the effect of inertia force and moment has to be considered seriously. The general equation of dynamics for the robot manipulator with  $N$  bodies can be expressed as:

$$\Delta W = \sum_{i=1}^N ((\mathbf{F}_i - m_i \ddot{\mathbf{r}}_{ic}) \cdot \Delta \mathbf{r}_{ic} + (\mathbf{M}_i - \dot{\mathbf{H}}_{ic}) \cdot \Delta \boldsymbol{\varphi}_{ic}) = 0 \quad (11)$$

where  $\mathbf{H}_{ic}$  is the angular momentum of the mass center for each body of the manipulator. Similar to the SBI, the instantaneous work of the external load can be balanced by energy flow inside the manipulator. So  $\Delta E'$  and the instantaneous buffering index  $\eta$  have the same expression as that in Eq. (3) and Eq. (4).

Suppose that a manipulator will follow a predefined trajectory to finish a specified task. Then at each single discrete point, we have an instantaneous buffering index. So the Kineto-static Buffering Index  $\eta_k$  is defined as the instantaneous buffering index at each discrete point of the trajectory integrated over the trajectory and divided by the integration of the trajectory  $T_a$  itself:

$$\eta_k = \frac{\int_{T_a} \eta dT_a}{\int_{T_a} dT_a} \quad (12)$$

KBI has the value between zero and one because the index  $\eta$  is between zero and one. This index offers a relatively comprehensive evaluation for buffering property of manipulators during the operation process along a predefined trajectory. Similar index can also be defined to evaluate the buffering effect of a manipulator over the whole working space  $W$  by the Global Buffering Index (GBI)  $\eta_G$ :

$$\eta_G = \frac{\int_W \eta dW}{\int_W dW} \quad (13)$$

Because of the introduction of inertia force and moment, kinetic energy appears. The kinetic energy  $T$  can be expressed as:

$$T = \sum_{i=1}^N \frac{1}{2} [m_i v_i^2 + \boldsymbol{\omega}_i \cdot (\mathbf{J}_i \cdot \boldsymbol{\omega}_i)] \quad (14)$$

where  $\boldsymbol{\omega}_i$  and  $v_i$  are the angular and linear velocity of mass center for each body.

Suppose the manipulator has  $f$  DOF and if we rewrite the Eq. (11) in the form of generalized coordinates with kinetic energy  $T$ , then we can get:

$$\sum_{j=1}^f [Q_j - \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_j} \right) + \frac{\partial T}{\partial q_j}] \Delta q_j = 0 \quad (15)$$

where  $Q_j$  is the generalized force and it has the relation with the potential energy  $V$ :

$$Q_i = -\frac{\partial V}{\partial q_j} \quad (j = 1, 2, \dots, f) \quad (16)$$

Because of the damping, we also introduce the dissipative energy here, then the total energy of the system utilized to balance the external work is:

$$\Delta \mathbf{D}^T \hat{\mathbf{F}}_e = -\delta E' = \sum_{j=1}^f \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} + \frac{\partial \psi_q}{\partial \dot{q}_j} \right] \Delta q_j \quad (17)$$

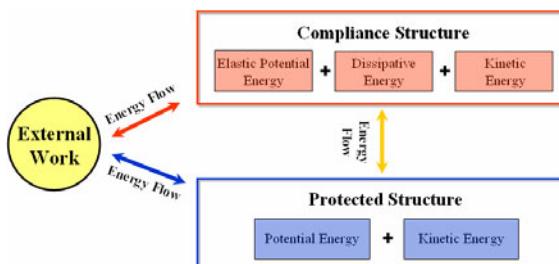
where  $L$  is the Lagrange function, satisfying:

$$L = T - V \quad (18)$$

Suppose the buffering structure has  $N_1$  DOF and the protected structure has  $N_2$ . Then we can express the energy of these two structures in the form of generalized coordinates as:

$$\begin{aligned} \Delta E_b &= -\sum_{j=1}^{N_1} \left[ \frac{d}{dt} \left( \frac{\partial L_q}{\partial \dot{q}_j} \right) - \frac{\partial L_q}{\partial q_j} + \frac{\partial \psi_q}{\partial \dot{q}_j} \right] \Delta q_j \\ \Delta E_p &= -\sum_{j=1}^{N_2} \left[ \frac{d}{dt} \left( \frac{\partial L_p}{\partial \dot{q}_j} \right) - \frac{\partial L_p}{\partial q_j} \right] \Delta q_j \end{aligned} \quad (19)$$

From Eq. (19), we find that the KBI has the different physical meaning compared with the SBI. Because of the introduction of inertia force and moment, the kinetic energy appears in both buffering structure and protected structure. Moreover, since the robot manipulator is in motion, the value of instantaneous buffering index will change during the operation process. That's because the flexible elements will store and release elastic potential energy during a cycle of operation and the dissipative and kinetic energy are closely related to kinematic parameters of the manipulator. This energy distribution is illustrated in Fig. 2.



**Fig. 2.** Energy distribution of the robot manipulator under kineto-static condition

### 3 Calculation Method for Buffering Indices

With the definition of buffering indices, we'll present the detailed calculation method in this section. The key step is the calculation of energy based on stiffness mapping.

#### 3.1 Calculation for the Energy of Protected Structure

In order to calculate the energy of the protected structure, the buffering structure can be removed from the robot manipulator temporarily. Then, according to the Eq. (2), external load will be balanced only by the energy of the protected structure as:

$$\Delta \mathbf{D}_p^T \hat{\mathbf{F}}_e + \Delta E_p = 0 \quad (20)$$

Then the dynamic equation of the protected structure at an instantaneous time is:

$$\Delta \dot{\mathbf{D}}_p^T \hat{\mathbf{F}}_e + \Delta \dot{\mathbf{q}}^T \hat{\mathbf{F}}_d + \Delta \dot{\mathbf{p}}^T \hat{\mathbf{G}} + \Delta \dot{\mathbf{p}}^T \hat{\mathbf{F}}^* = 0 \quad (21)$$

where  $\hat{\mathbf{F}}_d$ ,  $\hat{\mathbf{F}}_e$  and  $\hat{\mathbf{F}}^*$  are the generalized driving force of actuators, the generalized external force and the generalized inertia force respectively, all of which are the combination of force and torque.  $\hat{\mathbf{G}}$  is the gravity force vector of each body

The holonomic constraint equation of the system is:

$$\Phi(\mathbf{u}, \mathbf{q}) = 0 \quad (22)$$

where  $\mathbf{u}$  is the dependent coordinate.

According to the projection between different coordinate systems, we can get:

$$\hat{\mathbf{F}}_e = -\mathbf{J}_D^T \hat{\mathbf{F}}_d - \mathbf{J}_P^T \hat{\mathbf{G}} - \mathbf{J}_P^T \hat{\mathbf{F}}^* \quad (23)$$

Then we can get the stiffness mapping of the forces and torques on the end-effector to the corresponding linear and angular Cartesian displacement of end-effector by:

$$\mathbf{K}_{pe} = \frac{\partial \hat{\mathbf{F}}_e}{\partial \mathbf{D}_p} = -\mathbf{J}_D^T \mathbf{K}_d \mathbf{J}_D - \frac{\partial \mathbf{J}_D^T}{\partial \mathbf{D}_p} \hat{\mathbf{F}}_d - \frac{\partial \mathbf{J}_P^T}{\partial \mathbf{D}_p} \hat{\mathbf{G}} - \frac{\partial \mathbf{J}_P^T}{\partial \mathbf{D}_p} \hat{\mathbf{F}}^* \quad (24)$$

where  $\mathbf{K}_d$  is the stiffness matrix of actuators and is expressed as:

$$\mathbf{K}_d = \frac{d\hat{\mathbf{F}}_d}{d\mathbf{q}} = \text{diag}(K_{d_1}, K_{d_2}, \dots, K_{d_n}) \quad (25)$$

If we want to calculate the energy of the protected structure for the SBI, then the term of generalized inertia force  $\hat{\mathbf{F}}^*$  is zero. From the expression of stiffness matrix, we can find that the stiffness of the rigid manipulator is determined by the configuration of robot manipulators and the stiffness of actuators. Then, the energy of the protected structure will be calculated by the stiffness matrix  $\mathbf{K}_{pe}$  as:

$$\Delta E_p = -\Delta \mathbf{D}_p^T \mathbf{K}_{pe} \Delta \mathbf{D}_p \quad (26)$$

### 3.2 Calculation for the Energy of Robot manipulators

The calculation of the energy of the whole system is similar to the protected structure. So based on the stiffness matrix  $\mathbf{K}_e$ , the energy flow inside the manipulator is:

$$\Delta E' = -\Delta \mathbf{D}^T \mathbf{K}_e \Delta \mathbf{D} \quad (27)$$

However, for the introduction of flexible and dissipative elements, the stiffness mapping is more complicated than that for rigid body system. Suppose  $\Delta W_1$  and  $\Delta W_2$  are the work of spring and damping respectively, which have the relationship with elastic potential energy  $V_s$  and dissipative function  $\psi_q$ . So we can introduce these two items into the dynamic equation for the manipulator as:

$$\Delta \dot{\mathbf{D}}^T \hat{\mathbf{F}}_e + \Delta \dot{\mathbf{q}}^T \hat{\mathbf{F}}_d + \Delta \dot{\mathbf{p}}^T \hat{\mathbf{G}} + \Delta \dot{\mathbf{p}}^T \hat{\mathbf{F}}^* + \Delta \dot{\mathbf{q}}^T \frac{\partial W_1}{\partial \mathbf{q}} + \Delta \dot{\mathbf{q}}^T \frac{\partial W_2}{\partial \mathbf{q}} = 0 \quad (28)$$

Similar to the rigid manipulator, the stiffness matrix of robot manipulators with spring and damping can be obtained by introducing the Jacobian matrix as:

$$\mathbf{K}_e = \frac{\partial \hat{\mathbf{F}}_e}{\partial \mathbf{D}} = -\mathbf{J}_D^T \mathbf{K}_d \mathbf{J}_D - \frac{\partial \mathbf{J}_D^T}{\partial \mathbf{D}} \hat{\mathbf{F}}_d - \frac{\partial \mathbf{J}_P^T}{\partial \mathbf{D}} \hat{\mathbf{G}} - \frac{\partial \mathbf{J}_P^T}{\partial \mathbf{D}} \hat{\mathbf{F}}^* + \left( \frac{\partial \mathbf{J}_D^T}{\partial \mathbf{D}} \frac{\partial V_s}{\partial \mathbf{q}} + \frac{\partial \mathbf{J}_D^T}{\partial \mathbf{D}} \frac{\partial \psi_q}{\partial \mathbf{q}} \right) \quad (29)$$

Since the flexible joints can be simplified as rigid joints with nonlinear springs, the stiffness mapping is the same as that we discussed above. However, we have to develop the stiffness mapping of robot manipulators with flexible links by the finite element method under heavy loads because the equivalent model of simple springs to present the axial, rotational and flexural stiffness of flexible links isn't accurate enough when the deformation of these links is relatively large.

Suppose there is a flexible link named  $B_i$  with  $l_i$  units, then it will have totally  $l_i + 1$  nodes. Then, the dynamic equation of the flexible link under static condition is:

$$\Delta \dot{\mathbf{p}}_i^T \hat{\mathbf{F}}_{Ci}^a + \Delta \dot{\mathbf{a}}_i^T \mathbf{f}_i^O - \Delta \dot{\mathbf{a}}_i^T \mathbf{K}_i \mathbf{a}_i = 0 \quad (30)$$

where  $\dot{\mathbf{p}}_i = [\dot{\mathbf{r}}_i, \boldsymbol{\omega}_i]^T \in \mathbb{R}^{6 \times 1}$ , and  $\dot{\mathbf{r}}_i$  and  $\boldsymbol{\omega}_i$  are the velocity of body-fixed coordinate,  $\mathbf{f}_i^O$  is the generalized applied force,  $\mathbf{F}_i^O$  is the generalized deformation force,  $\mathbf{C}_i$  and  $\mathbf{K}_i$  are the mode damping matrix and mode stiffness matrix of link  $B_i$ , and  $\hat{\mathbf{F}}_{Ci}^a$  is the principal vector and moment at the original point of the body-fixed coordinate system. If there are totally  $n$  flexible links in a manipulator, then:

$$\sum_{i=1}^n (\Delta \dot{\mathbf{p}}_i^T \hat{\mathbf{F}}_{Ci}^a + \Delta \dot{\mathbf{a}}_i^T \mathbf{f}_i^O - \Delta \dot{\mathbf{a}}_i^T \mathbf{K}_i \mathbf{a}_i) = 0 \quad (31)$$

All applied forces of the flexible link are gravity force, driving force and external force, so:

$$\mathbf{f}_i^O = \sum_{k=1}^{l_i+1} \Phi_i^{kT} \mathbf{F}_i^k = \sum_{k=1}^{l_i+1} (\Phi_i^{kT} m^k) \mathbf{g} + \Phi_{di}^{kT} \mathbf{F}_i^d + \Phi_e^{kT} \mathbf{F}_i^e = \gamma_i \mathbf{A}_i \mathbf{g} + \Phi_{di}^{kT} \mathbf{F}_i^d + \Phi_e^{kT} \mathbf{F}_i^e \quad (32)$$

where  $\Phi_{di}^k$  and  $\Phi_e^k$  are the mode matrixes of the nodes acted by driving force and external force respectively. With the expression of  $\mathbf{f}_i^O$  and  $\dot{\mathbf{a}}_i^T$ , we will get the expression of last two terms in Eq. (31) by:

$$\sum_{i=1}^n \Delta \dot{\mathbf{a}}_i^T \mathbf{K}_i \mathbf{a}_i = \Delta \dot{\mathbf{p}}^T \mathbf{K}_a, \quad \sum_{i=1}^n \Delta \dot{\mathbf{a}}_i^T \mathbf{f}_i^O = \Delta \dot{\mathbf{p}}^T \tilde{\mathbf{F}}_a = \Delta \dot{\mathbf{p}}^T (\gamma \hat{\mathbf{g}} + \tilde{\mathbf{F}}_d + \tilde{\mathbf{F}}_e) \quad (33)$$

Finally, the power for robot manipulators with flexible links is:

$$\Delta \dot{\mathbf{q}}^T \hat{\mathbf{F}}_d + \Delta \dot{\mathbf{D}}^T \hat{\mathbf{F}}_e + \Delta \dot{\mathbf{p}}^T \hat{\mathbf{G}} + \Delta \dot{\mathbf{p}}^T \hat{\mathbf{F}}^* + \Delta \dot{\mathbf{p}}^T \tilde{\mathbf{F}}_a - \Delta \dot{\mathbf{p}}^T \mathbf{K}_a = 0 \quad (34)$$

By introducing the Jacobian matrixes  $\mathbf{J}_p$  and  $\mathbf{J}_D$ , we will get the expression of generalized external force:

$$\hat{\mathbf{F}}_e = -\mathbf{J}_D^T \hat{\mathbf{F}}_d - \mathbf{J}_p^T \hat{\mathbf{G}} - \mathbf{J}_p^T \hat{\mathbf{F}}^* - \mathbf{J}_p^T \tilde{\mathbf{F}}_a + \mathbf{J}_p^T \mathbf{K}_a \quad (34)$$

So the stiffness matrix  $\mathbf{K}_e$  can be obtained by:

$$\mathbf{K}_e = \frac{\partial \hat{\mathbf{F}}_e}{\partial \mathbf{D}} = -\mathbf{J}_D^T \mathbf{K}_d \mathbf{J}_D - \frac{\partial \mathbf{J}_D^T}{\partial \mathbf{D}} \hat{\mathbf{F}}_d - \frac{\partial \mathbf{J}_p^T}{\partial \mathbf{D}} \hat{\mathbf{G}} - \frac{\partial \mathbf{J}_p^T}{\partial \mathbf{D}} \hat{\mathbf{F}}^* - \frac{\partial \mathbf{J}_p^T}{\partial \mathbf{D}} \tilde{\mathbf{F}}_a + \frac{\partial \mathbf{J}_p^T}{\partial \mathbf{D}} \mathbf{K}_a \quad (36)$$

By comparing this equation with the stiffness expression of rigid bodies in Eq. (27), it is easy to find that there're two extra terms in this expression: the coupling stiffness and the elastic stiffness of flexible links.

### 3.3 Calculation for the Energy of the Buffering Structure

With the Eq. (26) and Eq. (27), it is easy to calculate the energy of the buffering structure  $E_b$  at an instantaneous time by:

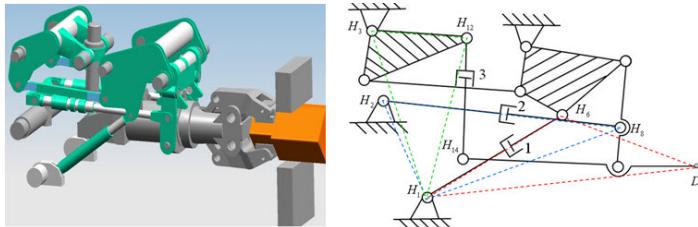
$$\Delta E_b = \Delta E - \Delta E_p = -\Delta \mathbf{D}^T \mathbf{K}_e \Delta \mathbf{D} + \Delta \mathbf{D}_p^T \mathbf{K}_{pe} \Delta \mathbf{D}_p \quad (37)$$

With the energy of both buffering structure and the whole robot manipulator, we can calculate the SBI and the KBI via the ratio of these two items.

## 4 Case Study

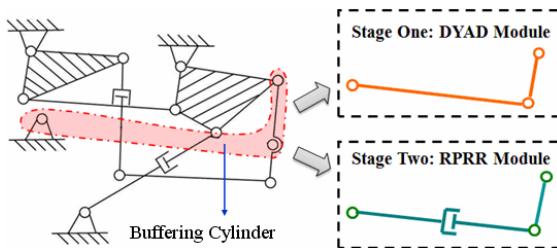
In this section, the buffering effects of a heavy forging manipulator with time-varying topology buffering structure are evaluated. And the simulation results prove that the proposed indices are reasonable and invariant.

Forging manipulator is widely used in industry to hold the workpiece during the forging process. It can undertake the huge anti-forces caused by the transfiguration of the workpiece. Here, a 3DOF mechanism consisting of 15 joints and 11 bodies is established in Fig. 3. Thereinto, 3 prismatic joints are actively actuated by hydraulic cylinders which are labeled as 1, 2 and 3 respectively. The end-effector is labeled as the point D.



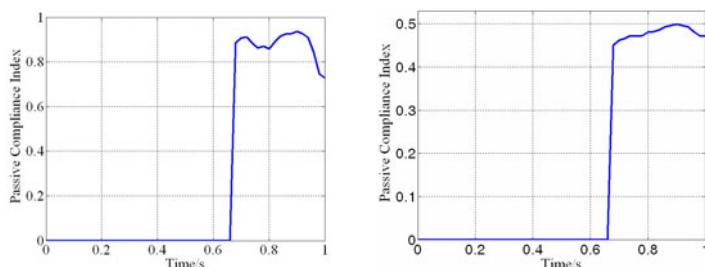
**Fig. 3.** Geometric model and the 3DOF mechanism of a forging manipulator

Forging manipulator experiences two stages with different topology during the working process, as shown in Fig.4. The “lock-unlock” characteristic of hydraulic cylinder 2, controlled by a flooding valve, is responsible for the topology change. At stage one, it is locked so the DOF of the system is  $f=2$ . Then the increasing anti-forces will cause the increase of oil pressure and at last unlock this hydraulic cylinder to offload forces. Then at stage two the topology is changed and the DOF is  $f=3$ .



**Fig. 4.** Topology change process of a forging manipulator

Then, according to the discussion in the above sections, the buffering indices can be obtained. And Fig. 5 illustrates the results according to two cases: horizontal external load only and the same horizontal and vertical external loads.



(a) Passive buffering index (horizontal) (b) Passive buffering index (vertical and horizontal)

**Fig. 5.** Buffering index along predefined trajectory under two different cases

The KBI  $\eta_k$  can be calculated based on Eq. (13) as:

$$\begin{aligned}\eta_{k1} &= \int_{T_a} \eta_1 dT_a / \int_{T_a} dT_a = 0.0444 / 0.005 = 0.888 \\ \eta_{k2} &= \int_{T_a} \eta_2 dT / \int_{T_a} dT_a = 0.0024 / 0.005 = 0.48\end{aligned}\quad (38)$$

where  $\eta_{k1}$  is the KBI with horizontal external load only and  $\eta_{k2}$  is the KBI with both horizontal and vertical external loads. So the index reveals that buffering cylinder will absorb most of the horizontal energy caused by the external load to protect the remaining structure.

## 5 Conclusions

In this paper, two buffering indices have been established to evaluate the ability of robot manipulators to comply with large and unpredictable external loads from environment: Static Buffering Index (SBI) and Kineto-Static Buffering Index (KBI). Different from the compliance matrix, the proposed indices reveal the energy distribution directly between the buffering structure and the protected structure inside a robot manipulator. One industrial example proves that the manipulators with higher value of buffering indices will have better buffering ability. However, since the KBI is based on the kineto-static analysis, the dynamic stiffness has to be introduced if we want to evaluate the dynamic buffering behavior for robot manipulators. Furthermore, the design and optimization of buffering structure guided by these indices haven't been well developed yet. This paper offers the fundamental work only and the further research of design and optimization will be presented in future.

## Acknowledgments

This research is jointly supported by the National Natural Science Foundation of China (Grant No. 50605042) and the National Basic Research Program of China (973 Program) (Grant No. 2006CB705400).

## References

1. Asada, H., Goldfine, N.: Optimal Compliance Design for Grinding Robot Tool Holders. In: Proceedings of the 1985 IEEE International Conference on Robotics and Automation, San Louis, MO (1985)
2. Schimmels, J.M.: Multidirectional Compliance and Constraint for Improved Robotic Deburring: Part 1 Improved Positioning, Part 2 Improved Bracing. *Robotics and Computer Integrated Manufacturing* 17, 277–286 (2001)
3. Leu, M.C., Jia, Y.L.: Mating of Rigid Parts by a Manipulator with Its Own Compliance. *ASME Journal of Engineering for Industry* 117, 240–247 (1995)
4. Yun, S.K.: Compliant Manipulation for Peg-in-Hole: is Passive Compliance a Key to Learn Contact Motion? In: Proceedings of the 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA (2008)

5. Khatait, J.P., Mukherjee, S., Seth, B.: Compliant Design for Flapping Mechanism: A Minimum Torque Approach. *Mechanism and Machine Theory* 41, 3–16 (2006)
6. Tantanawat, T., Kota, S.: Design of Compliant Mechanisms for Minimizing Input Power in Dynamic Applications. *ASME Journal of Mechanical Design* 129, 1064–1075 (2007)
7. Zhao, K., Wang, H., Chen, G.L.: Dynamics Simulation of Topology-Varying Mechanisms by Replacing Local Modeling Units. In: Proceedings of ASME Design Engineering Technical Conference, Brooklyn, New York (2008)
8. Whitney, D.E.: Quasi-static Assembly of Compliantly Supported Parts. *ASME Journal of Dynamic System, Measurement and Control* 104(1), 65–77 (1982)
9. Asada, H., Kakumoto, Y.: The Dynamic Analysis and Design of a High-speed Insertion Hand Using the Generalized Centroid and Virtual Mass. *ASME Journal of Dynamic System, Measurement and Control* 112, 646–652 (1990)
10. Schimmels, J.M., Huang, S.: A Passive Mechanism that Improves Robotic Positioning Through Compliance and Constraint. *Robotics and Computer-Integrated Manufacturing* 12, 65–71 (1996)
11. Hudgens, J., Cox, D., Tesar, D.: Classification structure and compliance modeling for serial manipulators. In: Proceedings of the 2000 IEEE International Conference on Robotics and Automation, San Francisco, CA (2000)
12. Kim, W., Li, B.J., Cho, W.: RCC Characteristics of Planar/Spherical Three Degree-of-Freedom Parallel Mechanisms with Joint Compliances. *ASME Journal of Mechanical Design* 122, 10–16 (2000)
13. Ang, M.H., Wang, W., Robert, N.K., Low, T.S.: Passive Compliance from Robot Limb and Its Usefulness in Robotic Automation. *Journal of Intelligent and Robotic Systems* 20, 1–21 (1997)
14. Chakarov, D.: Study of the Passive Compliance of Parallel Manipulators. *Mechanism and Machine Theory* 34, 373–389 (1999)
15. Loncaric, J.: Geometrical Analysis of Compliant Mechanisms in Robotics. Ph.D. Dissertation, Harvard University (1985)
16. Patterson, T., Lipkin, H.: Structure of Manipulator Compliance Matrices. In: 21st ASME Mechanisms Conference, Chicago, IL (1990)
17. Patterson, T., Lipkin, H.: A Classification of Robot Compliance. *ASME Journal of Mechanical Design* 115, 581–584 (1993)
18. Ciblak, N., Lipkin, H.: Synthesis of Cartesian Stiffness for Robotic Applications. Proceedings of the 1999 IEEE International Conference on Robotics and Automation, Detroit, MI (1999)
19. Roberts, R.G.: Minimal Realization of a Spatial Stiffness Matrix with Simple Springs Connected in Parallel. *IEEE Transactions on Robotics and Automation* 15, 953–958 (1999)
20. Huang, S., Schimmels, J.M.: The bounds and realization of spatial stiffnesses achieved with simple springs connected in parallel. *IEEE Transactions on Robotics and Automation* 14, 466–475 (1998)
21. Huang, S., Schimmels, J.M.: The bounds and realization of spatial compliances achieved with simple serial elastic mechanisms. *IEEE Transactions on Robotics and Automation* 16, 99–103 (1998)
22. Huang, S., Schimmels, J.M.: The Duality in Spatial Stiffness and Compliance as Realized in Parallel and Serial Elastic Mechanisms. *ASME Journal of Dynamic System, Measurement and Control* 124, 76–84 (2002)

# Robust Robotic Grasping Force Optimization with Uncertainty

Yang Wenyu, Li Miao, and Zhang Xiaoping

State Key Lab of Digital Manufacturing Equipment and Technology,

The School of Mechanical Science and Engineering

Huazhong University of Science and Technology, Wuhan 430074 P.R. China

limiao712@yahoo.cn, mewyang@mail.hust.edu.cn,

shoppinggre@hotmail.com

**Abstract.** The contact location and the coefficient of friction are two of the main uncertainties that rise in the robotic grasping force analysis problem. In many robotic tasks, we seek robust solutions to problem, where the problem data suffer from known statistical uncertainty, that is, a solution minimizes the objective function for the worst-case while satisfying the constraints for every possible case. Our objective is to reveal the effect of the uncertainty on the optimal solution. First we propose an uncertain point contact model (UPCM), which can describe the contact location uncertainty. Then together with the robotic grasping constraints, we formulate the constrained grasping force optimization problem as an unconstrained least-squares problem based on exterior penalty method. Finally, a numerical example is given to illustrate the method.

**Keywords:** Robust Optimization, Robotic Grasping Force, Exterior Penalty Method.

## 1 Introduction

The robotic community has a long history to deal with the uncertainty problems in the precise robotic tasks. The uncertainty mainly due to the discrepancy between the actual world and the used model: the robot's motions, the sensors, and all of the relevant physical parameters, such as the coefficient of friction. To deal with uncertainty, the robot can either use an effective and efficient control scheme based on sensory feedback or a robust model that is not sensitive to perturbation in the world. Under both cases, a robot must include uncertainty in its model of objects and physical processes [1], [2].

As for a robotic grasping task, one of the key issues is to determine the contact force between the manipulated object and the fingers. Without considering the uncertainties in the grasping model, in the previous research, the robotic grasping force analysis problem was usually transformed into an optimization problem. Kerr and Roth [3] used the linear programming (LP) technique in multifingered hands kinematic planning by using linear constraints to express friction constraints. Nakamura [4] proposed a nonlinear programming method to obtain the contact forces and many

scholars further developed the nonlinear programming method to improve its computation performance. Boyd [5] used the Second-Order Cone (SOC) to describe the friction cone constraint and recast the problem as a SOC problem. Based on the important observation that a friction constraint can be expressed as a semi-positive matrix [6], Han [7] further formulated the grasping forces analysis problem as a linear matrix inequality problems (LMI), which can be easily solved using the barrier method in convex optimization [8]. However, all of these works are based on the assumption that the contact location and the coefficient of friction are precisely known. Some researchers have taken the grasping uncertainties into consideration from other aspects, such as the impact of uncertainty on the force-closure analysis [9], the contact representation [10] and the motion planning [11]. Bone and Du [12] introduced a new metric for measuring the sensitivity of a grasp to positioning errors.

As far as our knowledge, considering the grasping uncertainties, few researches have been done in the grasping force analysis problem.

In this paper, we proposed an uncertain point contact model, which takes the imprecise contact location into consideration. In the optimization process, we further examine the influence of the coefficient of friction on the grasping forces, which is also one of our further research issues. The remainder of this paper is organized as follows. In Section 2, we formulate the grasping forces analysis problem in mathematics. In Section 3, we present the uncertain point contact model. In Section 4, we transform constrained grasping force optimization problem as an unconstrained least-squares problem based on external point method. In Section 5, a numerical example is given to illustrate this method. Finally, we make some discussions and point out the direction for the future work.

## 2 Problem Formulation

Consider a  $p$ -fingered robot hand grasping an object using an enveloping grasp, with a total of  $n$  contact points and  $k$  contact points on the palm. Assume that the contact model is point contact with friction and in the next section we will give the uncertain point contact model. Denote the grasp matrix at contact point  $i$  as  $\mathbf{G}_i \in R^{6 \times 3}, i = 1 \dots n$  and the whole grasp matrix can be written as:

$\mathbf{G} = [\mathbf{G}_1, \dots, \mathbf{G}_k, \mathbf{G}_{k+1}, \dots, \mathbf{G}_n] \in R^{6 \times 3n}$ , in which the former  $k$  columns are the grasp matrixes of the contact point on the palm. Static force equilibrium of all forces exerted on the object implies that:

$$\mathbf{G}\mathbf{t} = -\mathbf{w}_o \quad (1)$$

Where  $\mathbf{w}_o \in R^6$  is the external force exerting on the object and  $\mathbf{t} = [t_{1n}, t_{1t1}, t_{1t2}, \dots, t_{in}, t_{it1}, t_{it2}, \dots, t_{nn}, t_{nt1}, t_{nt2}]^T \in R^{3n}$  denotes the contact force at all the contact points and  $t_{in}, t_{it1}, t_{it2}$ , are the normal and tangential contact forces at contact point  $i$  respectively.

Next we will build the kinematic relationship between the contact forces and the joint torques of the robot hand. Denote  $\tau_{ij}$  the joint torque of the  $j^{th}$  joint of the finger

$i$  and  $\mathbf{t}_{in_i} \in R^3$  the contact forces of the contact point  $n_i$  at the finger  $i$ . So the relationship of joint torques and contact forces at the finger  $i$  can be written as [13]:

$$\boldsymbol{\tau}_i = \begin{bmatrix} \mathbf{J}_{i1}^T, \dots, \mathbf{J}_{in_i}^T \end{bmatrix} \begin{bmatrix} \mathbf{t}_{i1} \\ \vdots \\ \mathbf{t}_{in_i} \end{bmatrix} \stackrel{\Delta}{=} \mathbf{J}_i^T \mathbf{t}_i; \quad (2)$$

Where  $\boldsymbol{\tau}_i = [\tau_{i1}, \dots, \tau_{iqi}]^T \in R^{qi}$ , that means there are  $qi$  joints in the finger  $i$  and assuming there are total  $m$  joint torques at all the fingers, obviously,  $\sum_{i=1}^p qi = m$ .

Assembling joint torques and contact forces at all the fingers, we have the following linear mapping from the contact forces to the joint torques:

$$\boldsymbol{\tau} = \mathbf{J}^T \mathbf{t} \quad (3)$$

The essence of a grasping forces analysis problem is to determine  $\mathbf{t}$  in the force and torque equilibrium equation (1), (3) under some other grasping constraints, such as the friction constraint. According to the Coulomb friction law, the friction constrain can be represented as:

$$u_i t_{in} \geq \sqrt{t_{i1}^2 + t_{i2}^2} \quad (4)$$

Where  $u_i$  is the static coefficient at the  $i$  th contact point. Inequality (4) presents the friction cone constraint at  $i$  th contact, which is denoted as:

$$fc_i = \{\mathbf{t}_i \in R^{3 \times 1} \mid g_i(t) = u_i t_{in} - \sqrt{t_{i1}^2 + t_{i2}^2} \geq 0\} \quad (5)$$

Then the whole friction cone constraint at all the contact points can be represented by the Cartesian product of  $fc_i, i = 1, \dots, n$  denote by:

$$FC = fc_1 \times \dots \times fc_n = \{(\mathbf{t}_1, \dots, \mathbf{t}_n) \mid \mathbf{t}_i \in fc_i, i = 1, \dots, n\} \quad (6)$$

As discussion above, if we choose a proper objective function according to the task, then a grasping force analysis problem can be casted into an optimization problem:

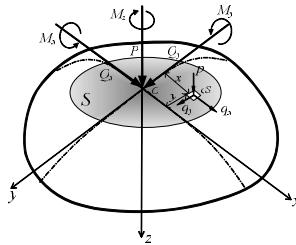
$$\begin{aligned} & \text{minimize : } F_{task} \\ & \text{subject to } \mathbf{G}\mathbf{t} = -\mathbf{w}_o \\ & \quad \boldsymbol{\tau} = \mathbf{J}^T \mathbf{t} \\ & \quad \mathbf{t} \in FC \end{aligned} \quad (7)$$

Here we find that the contact location, orientation and the coefficient of friction are related to the grasping force optimization problem, but there is not such work to clearly describe their impact on the grasping forces. So here we try to make some effort to reveal their relationship and we will focus on the uncertainty of location and coefficient of friction.

### 3 Uncertain Point Contact Model

In the robotic grasping force analysis problem, the types of contact usually are classified as: frictionless point contact (FPC), point contact with friction (PCWF), soft finger contact (SFC) [14]. Actually, all of these contact models are based on the basic contact mechanics principle [15]. The equivalent force transmitted from one contact surface to another through a point of contact can be resolved into a normal force  $P$ , a tangential force  $Q$  and the moment  $M$ , which can be express as follows and shown in Fig. 1.

$$\begin{aligned} P &= \int_S p dS & Q_x &= \int_S q_x dS & Q_y &= \int_S q_y dS \\ M_x &= \int_S py dS & M_y &= -\int_S pxdS & M_z &= \int_S (q_y x - q_x y) dS \end{aligned} \quad (8)$$



**Fig. 1.** Force and moment acting on a contact area [15]

In the point contact model, we usually not consider the spin moment  $M_z$ , arising from friction within the contact area, and we can also always find a point, on which we have:

$$M_x = 0 \quad M_y = 0 \quad (9)$$

In fact, this special point not always locates on the center of the contact area, to some degree, it is related to the distribution of contact force. Here, we use the uncertain point contact model (UPCM) to illustrate this characteristic of the equivalent force point. The location of each contact area between the fingers and object can be expressed as:

$$\mathbf{a}_i \in \mathcal{E}_i = \{\bar{\mathbf{a}}_i + \mathbf{P}_i \mathbf{u} \mid \|\mathbf{u}\|_2 \leq 1\} \quad (10)$$

Where  $\mathbf{P}_i \in R^{3 \times 3}$  is the uncertain matrix and  $\bar{\mathbf{a}}_i$  is the theoretic equivalent contact point.

- (1) When  $\text{rank}(\mathbf{P}_i) = 3$ , the uncertain equivalent contact point locates in an ellipsoid;
- (2) When  $\text{rank}(\mathbf{P}_i) = 2$ , the uncertain equivalent contact point locates in an ellipse;

(3) When  $\text{rank}(\mathbf{P}_i) = 1$ , the uncertain equivalent contact point locates on a line;

(4) When  $\mathbf{P}_i = 0$ , the equivalent contact point is determinate;

If the uncertain matrix is known, then we can use the robust optimization to solve (7), which is represented as:

$$\begin{aligned} \minimize & : F_{\text{task}} \\ \text{subject to} & \quad \mathbf{G}(a)\mathbf{t} = -\mathbf{w}_o \\ & \quad \mathbf{\tau} = \mathbf{J}^T \mathbf{t} \\ & \quad \mathbf{t} \in FC \\ & \quad \mathbf{a} = \{(\mathbf{a}_1, \dots, \mathbf{a}_n) \mid \mathbf{a}_i \in \varepsilon_i, i = 1, \dots, n\} \end{aligned} \quad (11)$$

Here,  $\mathbf{a} = \{(\mathbf{a}_1, \dots, \mathbf{a}_n) \mid \mathbf{a}_i \in \varepsilon_i, i = 1, \dots, n\}$  represents the real location of the contact points, which can be easily expressed by the uncertain point contact model (UPCM). We found that when  $\mathbf{a}$  is given; the optimization problem (11) is equal to the primal problem (7). When there are some uncertainties in  $\mathbf{a}$ , that is, at least one of the uncertain matrix satisfies:  $\mathbf{P}_i \neq 0$ . The problem (11) becomes a robust optimization problem; we require that the constraints be satisfied for all of the possible value of the parameters  $\mathbf{a} = \{(\mathbf{a}_1, \dots, \mathbf{a}_n) \mid \mathbf{a}_i \in \varepsilon_i, i = 1, \dots, n\}$ . However, in practice, we concentrates on the case where the perturbation or uncertainty is stochastic with know statistics. Then our main objective of this paper is to quantify the effect of the stochastic uncertainty of problem data on the optimal solution.

## 4 Robust Grasping Force Computation Model

Using the exterior penalty method, when the problem data is determinate, the constrained nonlinear programming problem (11) can be transformed into an unconstrained nonlinear programming problem (12), where  $r$  is a positive penalty parameter and  $h_i(t)$  represents all of the equality constraints in (11) and  $g_j(t)$  describes all of the inequality constraints in (11). It is has been proven that as  $r$  is large enough, the optimization problem (12) is equal to (11) [16].

$$\min : \{ F_{\text{task}}^T F_{\text{task}} + r[\sum_i (h_i(t))^2 + \sum_j (g_j(t))^2] \} \quad (12)$$

In fact, the problem (12) can be further formulated as a least-squares problem, which can be rewritten as:

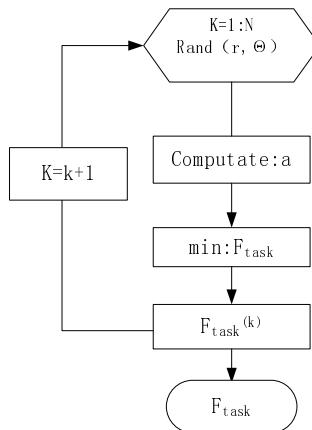
$$\begin{aligned} \min : & \phi = \psi(\mathbf{t})^T \psi(\mathbf{t}); \\ \psi(\mathbf{x}) = & \begin{bmatrix} F_{\text{task}} \\ \sqrt{r} g(\mathbf{t}) \\ \sqrt{r} h(\mathbf{t}) \end{bmatrix} \end{aligned} \quad (13)$$

Consequently, we can obtain the optimal solution of (13) easily using some common unconstrained optimization algorithm, such as Newton method or L-M method.

Here we assume that the contact area is an ellipse, which can be expresses as:

$$\varepsilon_i = \{\bar{\mathbf{a}}_i + \begin{bmatrix} m & 0 \\ 0 & n \end{bmatrix} \begin{bmatrix} r \cos \theta \\ r \sin \theta \end{bmatrix} \mid \theta \in [0, 2\pi], \mid r \leq 1\} \quad (14)$$

Where  $m$  and  $n$  stand for the two axis of the ellipse. We use the Monte Carlo method to express the uncertainty of the contact location and  $r, \theta$  are the random variables, which stands for the distance and direction of the real location from the theoretic location, respectively. The flow chart of the optimization process is given as follows:

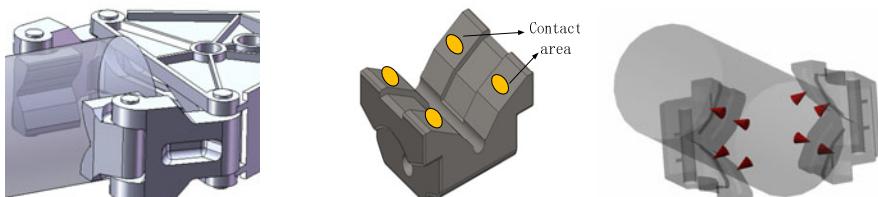


**Fig. 2.** The flow chart of the robust optimization process

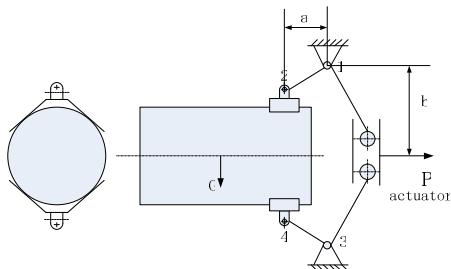
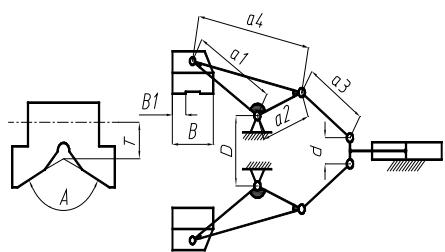
## 5 Numerical Example

In this section, a numerical example is presented to illustrate robust grasping force optimization method. From the derivation of the computation model, it can be found that this method can be applicable for other contact forces indeterminate system, such as workpiece-fixture system.

We use the method discuss above to optimize the minimal actuator force for a heavy-duty manipulation in forge industry, Fig. 3, Fig. 4.



**Fig. 3.** The structure of the gripper and the contact area

**Fig. 4.** The structure of the heavy-duty manipulation**Fig. 5.** The structure parameters for the manipulation

We can easily derive the relationship between the moment on joint 1 and 3, Fig.4 and the actuator force.

$$P = \frac{M_1 + M_3}{a_2 \sin \theta_{23}} \cos \theta_3 \quad (15)$$

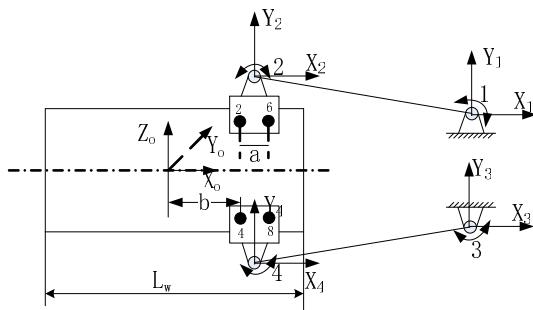
Where  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$  are the angle between  $a_1$  and horizontal line,  $a_2$  and horizontal line,  $a_3$  and horizontal line, respectively.

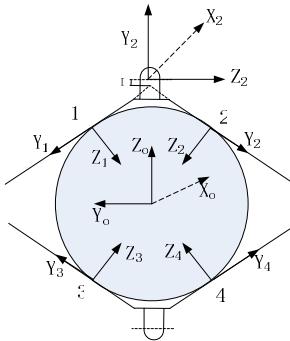
$$\theta_1 = \arcsin\left(\frac{C - D}{2a_1}\right) \quad (16)$$

$$\theta_2 = \pi - \arcsin\left(\frac{C - D}{2a_1}\right) - \arccos\left(\frac{a_1^2 + a_2^2 - a_4^2}{2a_1 a_2}\right) \quad (17)$$

$$\theta_3 = \arcsin\left(\frac{D + 2a_2 \sin \theta_2 - d}{2a_3}\right) \quad (18)$$

The coordinate of the contact and joint are shown as Fig. 6 and Fig. 7. There are eight contact points between the grippers and the workpiece, Fig. 7, each of which is point contact with friction (PCWF).

**Fig. 6.** The position of coordinates



**Fig. 7.** The position of coordinates on contact

Here we make an assumption that the real contact normal direction does not change. The real contact position and orientation can be represented as;

$$\mathbf{R}_{ci} = \overline{\mathbf{R}}_{ci} \quad (19)$$

$$\mathbf{P}_{ci} = \overline{\mathbf{P}}_{ci} + \overline{\mathbf{R}}_{ci} \mathbf{P}_i \mathbf{u} \quad (20)$$

Here,  $\mathbf{R}_{ci}$ ,  $\mathbf{P}_{ci}$  are the real contact orientation and position,  $\overline{\mathbf{R}}_{ci}$ ,  $\overline{\mathbf{P}}_{ci}$  are the theoretic contact orientation and position, which can be compute easily from Fig.5 and Fig. 6. Here we give the orientation and position of point 1 as an example:

$$\overline{\mathbf{R}}_{ci} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \quad \overline{\mathbf{P}}_{ci} = \begin{bmatrix} b \\ \frac{1}{2}R \\ \frac{\sqrt{3}}{2}R \end{bmatrix}$$

Because we use the PCWF contact model, the real grasp matrix can be express as:

$$\mathbf{G} = [\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_8] \in R^{6 \times 24} \quad (21)$$

Where  $\mathbf{G}_i = \begin{bmatrix} \mathbf{R}_{ci} & \mathbf{0} \\ [\mathbf{P}_{ci} \times] \mathbf{R}_{ci} & \mathbf{R}_{ci} \end{bmatrix} \mathbf{H}_i$  and the contact type matrix  $\mathbf{H}_i = \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix}_{6 \times 3}$  [13], which represents this kind of contact only transmits three contact forces.

In the same way, we can formulate the Jacobian matrix for the gripper.

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1^T & \mathbf{J}_2^T & \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & \mathbf{J}_5^T & \mathbf{J}_6^T & \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} \\ \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & \mathbf{J}_3^T & \mathbf{J}_4^T & \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 3} & \mathbf{J}_7^T & \mathbf{J}_8^T \end{bmatrix}^T \in R^{24 \times 4} \quad (22)$$

Because the joint 2 and 4 are passive joints, which can not provide any actuator forces, we have the passive joint constraint.

$$\mathbf{H}_p \mathbf{J}^T \mathbf{t} = \mathbf{0} \quad (23)$$

Where  $\mathbf{H}_p = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$  is a selection matrix.

Here we find that in order to determine the minimal actuator force, we are in fact only need to determine the minimal moment on the active joints 1 and 3. So we can choose our objective function as:

$$\min : \| \boldsymbol{\tau}_a = \mathbf{H}_a \mathbf{J}^T \mathbf{x} \| \quad (24)$$

Where  $\mathbf{H}_a = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$  is a selection matrix.

After choosing the objective function, we can use the method in section 3 and 4 to solve this optimization problem efficiently.

We use the 20T/50T.m manipulation as an example. The workpiece is made of ST41 with  $E_w = 116Gpa$ ,  $v_w = 0.43$  when the temperature is  $T = 950^{\circ}C$ . The grippers are made of 45# with  $E_f = 206Gpa$  and  $v_f = 0.31$  the structure parameters are given as follows:

**Table 1.** The structure parameters of the heavy-duty manipulation

Parameter:	Manipulation type 20T/50Tm		
A	1	130°	a4
B	0.57m	d	11.211m
B1	0.17m	Workpiece weight	20T
T	0.168m	Workpiece moment	50T.m
D	0.64m	Workpiece radius	0.6m
a1	0.943m	Actuator radius	400mm
a2	0.636m	Actuator pressure	20MPa
a3	0.7m		

When the coefficient of friction is  $u_i = 0.5, i = 1...8$ , the penalty parameter is,  $r = 10^{10}$ , and we don't take the contact position uncertainty into consideration. We obtain the minimal actuator force under a horizontal grasp and a vertical grasp, respectively.

$$F_{hor} = 1.0e+005 *$$

$$\begin{bmatrix} 1.5587 & 0.0000 & -1.7264 & -0.0000 & 0.1292 & 1.6878 & -0.0921 & -1.8185 \\ -2.5354 & -0.0000 & -2.3627 & 0.0000 & 0.1479 & -1.9262 & 0.1576 & -1.7438 \end{bmatrix} N$$

$$P_{hor} = 1.5781e+006 N$$

$$F_{ver} = 1.0e+005 *$$

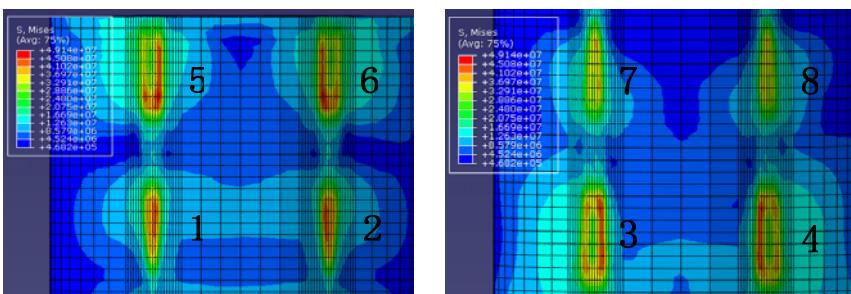
$$\begin{bmatrix} -0.3310 & 0.0003 & -1.7616 & 1.9097 & -1.8489 & 1.9180 & 0.1742 & 0.6011 \\ -0.1439 & -0.0000 & 0.3781 & 0.7695 & -0.4569 & -0.4240 & 0.1819 & -0.0325 \\ 0.7219 & 0.0005 & 3.6034 & 4.1178 & 3.8091 & 3.9286 & 0.5037 & 1.2040 \end{bmatrix} N$$

$$P_{ver} = 1.5863e+006 N$$

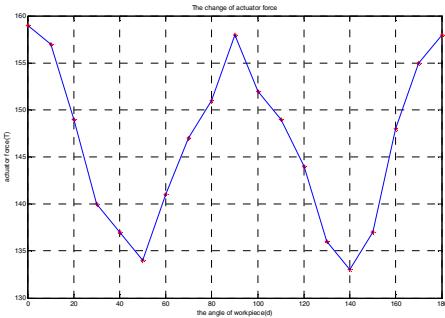
Where the  $i$ th column in  $F_{hor}$  and  $F_{ver}$  represents the three elements of the contact force at  $i$ th contact point, where the upper two rows are the tangential contact forces and the lowest row are the normal contact forces. We found that the contact forces in the contact points 3,4,5,6 are much bigger than others in vertical grasp and the contact forces in the contact points 1,3,6,8 are much bigger than others in horizontal grasp, which are conforming to the FEA simulation results, as shown in Fig.8. When the workpiece rotates through 180 degree, the minimal actuator force varies as Fig.9, which indicates that grasp configuration has a significant influence on the minimal actuator force. This issue is also one of our future research area that to plan a robotic grasping configuration with minimal actuator forces.

We also show the relationship between the coefficient of friction and the minimal actuator force, as shown in Fig. 10. Here we assume all the coefficients at all the contact points are the same. We find that when the coefficient of friction increases, the minimal actuator force will decrease. This also coincides with our intuition, that is, the more rough an object is, the more easily we can grasp it. In the further research, we will try to clarify relationship between the force-closure prosperity and uncertainty of coefficient of friction.

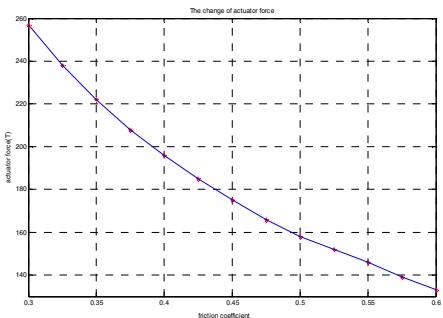
When the coefficient of friction is  $\mu_i = 0.5, i=1\dots8$ , the penalty parameter is,  $r=10^{10}$ , and the contact point uncertainty is expresses as  $r=\text{rand}(8,1)$   $\theta=2\pi \cdot \text{rand}(8,1)$   $m=0.2$   $n=0.1$ ; As shown in Fig. 11, we give the minimal actuator force when we run the optimization process ten times,  $N=10$ , that is, we check ten contact point in the uncertain contact area and obtain the minimal actuator for each possible contact point. When the contact area decreases, in another word, the uncertainty of the contact point position decrease, we also compute the minimal actuator force,  $m=0.1$   $n=0.05$ , Fig. 12.



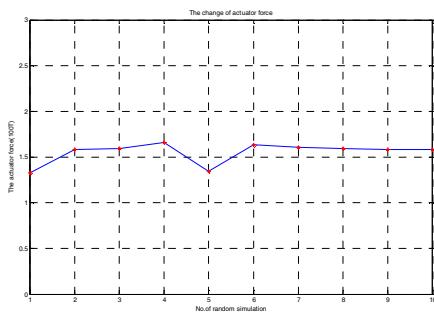
**Fig. 8.** The simulation of contact forces at contact points of vertical grasp



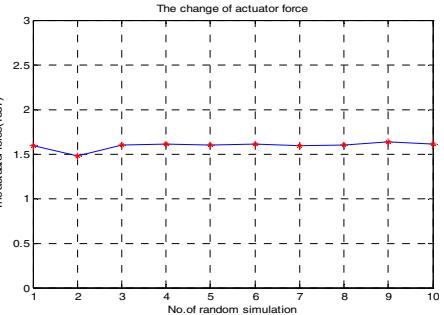
**Fig. 9.** The minimal actuator force varies when workpiece rotate 180 degree



**Fig. 10.** The relationship between the coefficient of friction and the minimal actuator force



**Fig. 12.** The minimal actuator force varies in 10 random optimization



**Fig. 13.** The minimal actuator force with small contact area

## 6 Conclusions and Future Work

This paper proposes an uncertainty point contact model to deal with grasping force analysis problem. We try to find a solution that satisfies the constraints under all possible cases. We formulate this problem as a robust optimization problem and solve it using exterior penalty method, which can be further transformed into an unconstrained least-squares problem. A numerical example is given to illustrate the method. However, our work is based on the assumption that the contact area is an ellipse and the uncertain contact point locates with it. In fact, when the contact area varies, contact point changes, as well as the contact normal direction. Furthermore, we don't give any analysis result about the impact of coefficient of friction on the contact forces. For a future work, we will clarify the relationship between the force-closure prosperity and uncertainty of coefficient of friction. Also, we will try to find a metric function that can quantify the impact of uncertainty on the objective optimal solution and an experiment will also be done to verify our method.

**Acknowledgments.** This work was supported by National Basic Research Program of China (973 Program), No. 2006CB705404.

## References

1. Goldberg, Mason, et al.: Geometric Uncertainty in Motion Planning: Summary Report and Bibliography, Catalina Island, CA (1992)
2. Erdmann, M.: Using Back Projections for Fine Motion Planning with Uncertainty. *The International Journal of Robotics Research* 5(1), 19–45 (1986)
3. Kerr, J., Roth, B.: Analysis of Multifingered Hands. *The International Journal of Robotics Research* 4, 3–17 (1986)
4. Nakamura, Y., Nagai, K., Yoshikawa, T.: Dynamics and Stability in Coordination of Multiple Robotic Mechanisms. *The International Journal of Robotics Research* 8, 44–61 (1989)
5. Lobo, M., Vandenberghe, L., Boyd, S., Lebret, H.: Applications of Second-Order Cone Programming. *Linear Algebra and its Applications* 284, 193–228 (1998)
6. Buss, M., Hashimoto, H., Moore, J.: Dextrous Hand Grasping Force Optimization. *IEEE Transactions on Robotics and Automation* 12(3), 406–418 (1996)
7. Li, H., Trinkle, J.C., Li, Z.: Grasp Analysis as Matrix Inequality Problems. *IEEE Transactions on Robotics and Automation* 16(6), 663–674 (2000)
8. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
9. Zheng, Y., Qian, W.-H.: Coping with the Grasping Uncertainties in Force-Closure Analysis. *The International Journal of Robotics Research* 24, 311–327 (2005)
10. Xiao, J., Zhang, L.: Towards Obtaining All Possible Contacts — Growing A Polyhedron by Its Location Uncertainty. *IEEE Transactions on Robotics and Automation* 12(4), 553–565 (1996)
11. Xiao, J., Volz, R.: A Replanning Approach towards Uncertainty Handling in Robot Assembly. Invited paper for the Third IEEE International Symposium on Intelligent Control, Arlington, VA, pp. 326–331 (1988)
12. Bone, G.M., Du, Y.: Multi-Metric Comparison of Optimal 2D Grasp Planning Algorithms. In: Proc. IEEE ICRA, pp. 3061–3066 (2001)
13. Murray, R.M., Li, Z., Sastry, S.S.: A Mathematical Introduction to Robotic Manipulation. CRC Press, Boca Raton (1994)
14. Cutkosky, R.: Robotic Grasping and Fine Manipulation. Kluwer, Boston (1985)
15. Johnson, K.L.: Contact Mechanics. Cambridge University Press, New York (1985)
16. Wismer, D.A., Chattergy, R.: Introduction to Nonlinear Optimization –A Problem Solving Approach, pp. 63–79. Elsevier North-Holland, Inc., New York (1978)

# Analysis of Forging Compliance Process and Design of the Forging Simulator

Pu Zhang, Zhenqiang Yao, and Zhengchun Du

The State Key Lab of Mechanical System and Vibration,  
Shanghai Jiao Tong University, Shanghai 200240  
[{zhe,zqyao,zcdt}@sjtu.edu.cn](mailto:{zhe,zqyao,zcdt}@sjtu.edu.cn)

**Abstract.** This paper presents a new simulator which has a good force capability, and the simulator can reproduce working environment of forging manipulator, by simulating force and movement. In this paper, firstly, forging compliance process of the forging manipulator DDS750 is analyzed, and there is a mechanism topological change during the process which includes initiative compliance and passive compliance. Secondly, a 2-DOF parallel planar mechanism is introduced as the simulator based on the above analysis, and the optimal work space of the mechanism is obtained based on the dexterity, force capability and stiffness performance index. Finally, an example of the simulator is introduced, which is a new approach for the testing and calibration of the heavy forging manipulator.

**Keywords:** Forging manipulator; Topological change; Parallel mechanism; Force capability; Compliance.

## 1 Introduction

Force capability and compliance capability is the core performance of heavy forging manipulator, especially the passive compliance capability in clamp's axial is critical for protecting the forging manipulator's body, therefore, the testing and calibration of heavy forging manipulator's force capacity and compliance capability is very important.

Heavy forging manipulator mainly depends on import at present in our country, correlative research has just begun, and the relevant content of testing and calibration for heavy forging manipulator has not reported in the world. There are many cases about testing and calibration for the other equipments, such as: Zengxiang Fu [1] calibrated the heavy press's force capability by forging specimens and found the best force capability through the specimens' stress-strain relationship; Hehua Zhu [2] used the soil box to simulate a variety of soil conditions, accordingly tested the shield's ground adaptability; Yonggang Lin [3] built a pitch-controlled wind turbine system to test the electro-hydraulic actuator's performance; Tongli Chang [4] built a HIL ground experiment simulator for testing the on-orbit docking mechanism. Literature [1] only calibrated the heavy press's force capability, but literatures [2-4] reproduced

the working environment by experimental methods to testing a variety of performances of the equipments, therefore they enlighten us a new approach for the testing and calibration of the heavy forging manipulator.

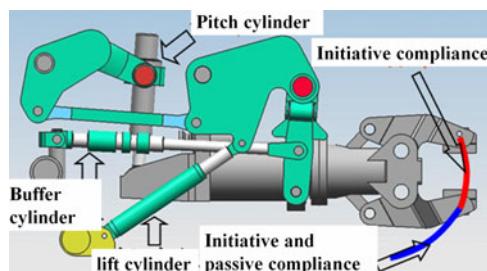
In this paper, forging manipulator DDS750 is the study object, and it's forging compliance process is analyzed, then a parallel mechanism is designed as the forging simulator based on the above analysis, to reproduce working environment by simulating force and movement of the forging blank, thus we can test and calibrate the equipment by mean of the simulator.

## 2 Analysis of Forging Compliance Process

Analysis of forging compliance process is the basis of designing heavy forging manipulator and forging simulator. Figure 1 shows forging manipulator's working environment. Figure 2 shows the mechanism of forging manipulator.



**Fig. 1.** The working environment of forging manipulator



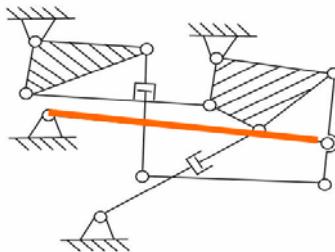
**Fig. 2.** The mechanism of forging manipulator

Through field research and analysis on the mechanism of forging manipulator, we find that there is a mechanism topological change during the forging process which includes two stages:

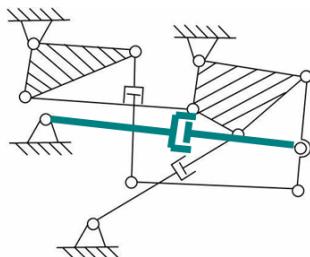
1. Initiative compliance stage. As shown in figure 3, the buffer cylinder is in a locked state, the lift cylinder's liftup motion realizes the initiative compliance of forging manipulator with 2 degrees of freedom.

2. Initiative and passive compliance stage. As shown in figure 4, the lift cylinder's liftpup motion realizes the initiative compliance, the buffer cylinder's state changes from lock into move, this changing leads to the passive compliance and mechanism topological change, therefore forging manipulator's degrees of freedom change to 3 this moment.

In both stages, the manipulator's clamp endures the deformation resistance in active compliance and passive compliance directions.



**Fig. 3.** The mechanism of initiative compliance



**Fig. 4.** The mechanism of initiative and passive compliance

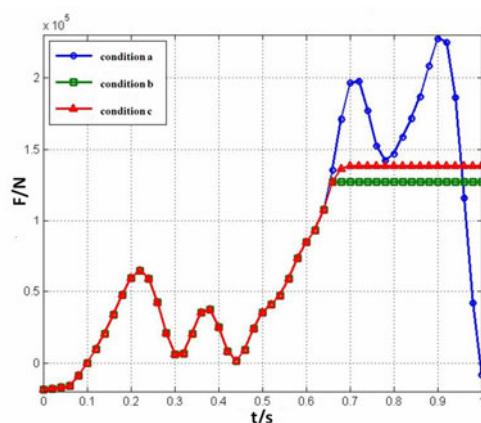
Based on the above analysis, forging manipulator's kinematics and dynamics simulation in the rounding process of a long shaft is done, to analyze the compliance's impact on key components of forging manipulator in three kinds of conditions, three kinds of conditions are: **a.** the buffer cylinder is locked; **b.** the lift cylinder opens normally without regard to hydraulic effects; **c.** the lift cylinder opens normally considering hydraulic effects.

As shown in figure 5 and 6, the forging process duration is 1 second; the pressure of buffer cylinder reaches the limit pressure of overflow valve at 0.65 s, therefore 0.65s is the opening time of buffer cylinder, it is initiative compliance stage before 0.65s, correspondingly it is initiative and passive compliance stage after 0.65s.

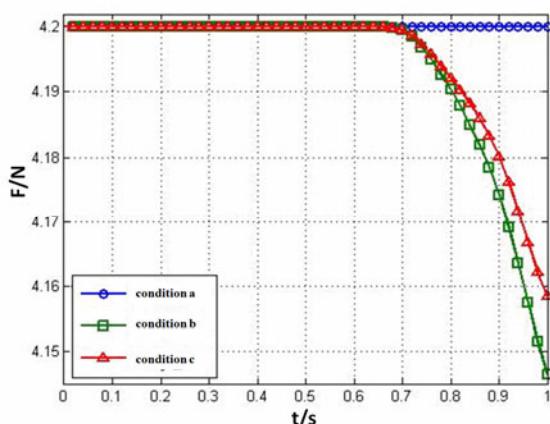
Comparing three kinds of conditions, before 0.65s the overflow valve of buffer cylinder is not open, the displacement of buffer cylinder is 0mm, the pressure of buffer cylinder is rising, the buffer cylinder's forces and displacements in three kinds of

conditions are identical; after 0.65s, in condition a, the force of the buffer cylinder is biggest, the buffer cylinder is locked all the time, cylinder displacement is zero with no any compliance , at about 0.7s and 0.9s two peak forces of buffer cylinder appear, the maximum force value is  $2 \times 10^5$ N; in the other two conditions, the buffer cylinder normally opens, the displacement is increasing gradually with passive compliance, two peak forces valued about 130kN and 140kN are at 0.65s and 0.7s; comparing with condition b, cylinder displacement in condition c is more shorter, peak forces appear later ,and those values are bigger.

The above analysis shows that passive compliance motion can reduce the buffer cylinder hinge point's constraint reaction, and reduce the driving force of the buffer cylinder to avoid the appearance of the peak force. Therefore, the passive compliance is especially important for protecting manipulator body.



**Fig. 5.** Curve of the buffer cylinder's driving force



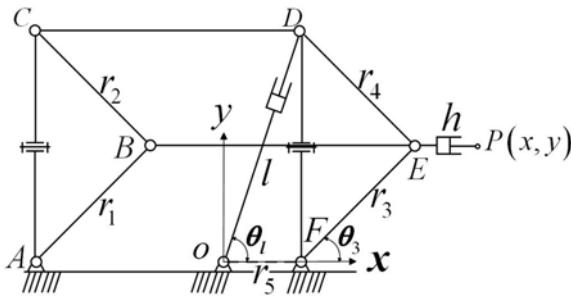
**Fig. 6.** Curve of the buffer cylinder's length

Analysis of forging compliance process shows that the forging manipulator's clamp endures the deformation resistance in active compliance and passive compliance directions, force capability and compliance capability of forging manipulator in passive compliance direction is especially important.

### 3 Design of the Forging Simulator

Parallel mechanisms have been used more and more extensive because of characteristics of high rigidity, strong capacity, high precision and compact structure, therefore combined with analysis of forging compliance process we design a kind of parallel mechanism as the forging simulator to reproduce working environment by simulating force and movement of the forging blank.

As shown in figure 7, the 2-DOF parallel mechanism combines sarrut parallel mechanism with parallelogram mechanism. The fixed platform of parallel mechanism is AF, the movable platform is BP, and the end effector is the point P. The driving cylinder OD and EP make the platform BP move with 2-DOF by changing the length  $l$ ,  $h$  of cylinders.



**Fig. 7.** 2-DOF parallel mechanism

#### 3.1 Kinematic Analysis

The coordinate system is fixed to the platform AF, as shown in figure 7,  $\overline{AB} = r_1$ ,  $\overline{BC} = r_2$ ,  $\overline{DE} = r_3$ ,  $\overline{EF} = r_4$ ,  $\overline{OF} = r_5$ ,  $\overline{OD} = l$ ,  $\overline{EP} = h$ ,  $r_1 = r_2 = r_3 = r_4 = r_5$ ,  $\angle FOD = \theta_1$ ,  $\angle XFE = \theta_3$ ,  $\theta_1, \theta_3 \in (0, \pi/2)$ ,  $P = \{x, y\}^T$ , the kinematics equation is:

$$\left\{ \begin{array}{l} x = r_5 + \frac{\sqrt{4r_3^2 + r_5^2 - l^2(t)}}{2} + h(t) \\ y = \frac{\sqrt{l^2(t) - r_5^2}}{2} \end{array} \right. \quad (1)$$

The velocity Jacobian matrix is obtained as equation (2) from derivation of equation (1):

$$J = \begin{Bmatrix} 1 & -\varphi(\theta_l) \tan \theta_3 \\ 0 & \varphi(\theta_l) \end{Bmatrix} \quad (2)$$

In the equation (2),  $\varphi(\theta_l) = \frac{1}{2 \sin \theta_l} = \frac{\sqrt{4y^2 + r_s^2}}{4y}$ ,  $\tan \theta_3 = \frac{y}{\sqrt{r_3^2 - y^2}}$ .

Considering the ideal plane, force Jacobian matrix is:

$$G = (J^{-1})^T = \begin{Bmatrix} 1 & 0 \\ \tan \theta_3 & 2 \sin \theta_l \end{Bmatrix} \quad (3)$$

### 3.2 Performance Analysis

The forging simulator is used to reproduce working environment by simulating force and movement of the forging blank, so it must has good performances including dexterity, force capability and stiffness performance, and an optimal work space of the mechanism should be obtained based on the performance analysis.

#### 3.2.1 Dexterity Analysis

The condition number of Jacobian matrix is dexterity performance index of parallel mechanism [5], [6], [7], the condition number of this mechanism is:

$$\kappa_J = \|J\| \|J^{-1}\| \quad (4)$$

In the equation (4),  $\|\bullet\|$  is arbitrary norm of matrix, usually taking Euclidean norm,  $\|J\| = \sqrt{\lambda_{\max}(J^T J)}$ , if  $K_J$  is smaller, dexterity and control accuracy will be better.

From equation (2), (4), we can know that the condition number of Jacobian matrix is decided by variable  $y$ . Figure 8 shows the curve of the Jacobian's condition number, minimum condition number is 1.17 when  $y = 0.09$  mm and condition number is between 1 and 2.5 when  $y = 0.05 - 0.42$  mm, where is the optimal condition number area.

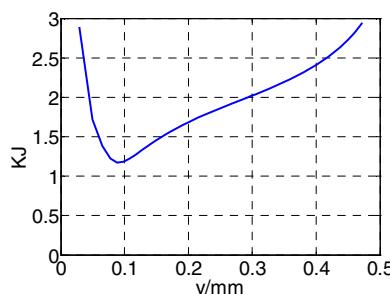
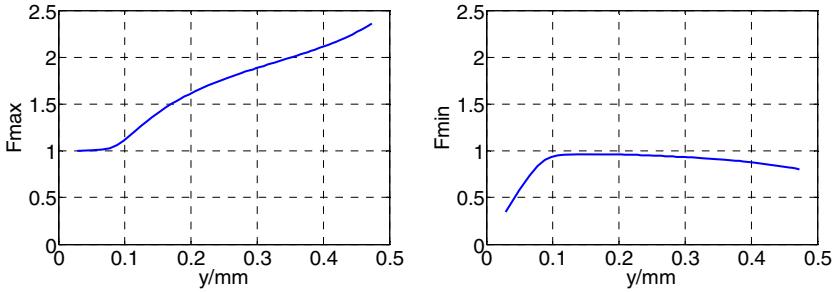


Fig. 8. Curve of the Jacobian's condition number



**Fig. 9.** Curve of the force capability index

### 3.2.2 Force Capability Analysis

Force capacity extremes are used to evaluate the force capacity of parallel mechanism in the work space, this parallel mechanism's maximum and minimum force capacity indexes [8] are:

$$F_{max} = \sqrt{\lambda_{F_{max}}(G^T G)} = \sigma_{F_{max}}(G) \quad (5)$$

$$F_{min} = \sqrt{\lambda_{F_{min}}(G^T G)} = \sigma_{F_{min}}(G) \quad (6)$$

In the equation (5), (6),  $G$  is the force Jacobian matrix,  $\lambda_{F_{max}}$ ,  $\lambda_{F_{min}}$  are the Maximum and minimum eigenvalues of matrix  $G^T G$ , if  $F_{max}$ ,  $F_{min}$  are bigger, force capacity of the parallel mechanism will be better.

From equation (3), (5), (6), we can know that the force capacity index is decided by variable  $y$ . Figure 9 shows the curve of the force capacity index, parallel mechanism has a good force capacity performance when  $0.1 < y < 0.4$  mm, force capacity performance will be better if variable  $y$  is bigger in this area.

### 3.2.3 Stiffness Analysis

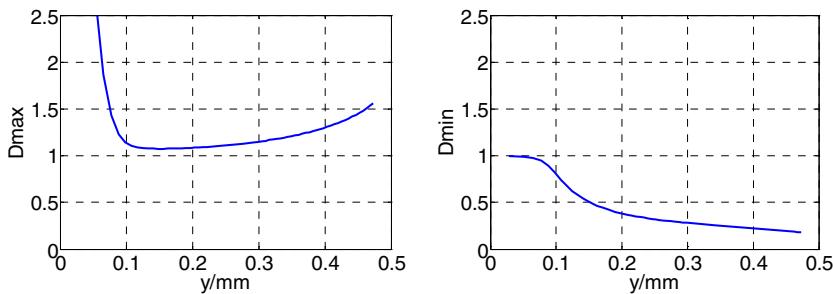
The end deformation extremes are used to evaluate the stiffness performance of parallel mechanism in the work space, this parallel mechanism's maximum and minimum stiffness performance indexes [8] are:

$$D_{max} = \sqrt{\lambda_{D_{max}}(C^T C)} = \sigma_{D_{max}}(C) \quad (7)$$

$$D_{min} = \sqrt{\lambda_{D_{min}}(C^T C)} = \sigma_{D_{min}}(C) \quad (8)$$

In the equation (7), (8),  $C$  is the flexibility matrix,  $C = JJ^T$ ,  $\lambda_{D_{max}}$ ,  $\lambda_{D_{min}}$  are the maximum and minimum eigenvalues of matrix  $C^T C$ , if  $D_{max}$ ,  $D_{min}$  are smaller, stiffness performance of the parallel mechanism will be better.

From equation (2), (7), (8), we can know that the stiffness performance index is decided by variable  $y$ . Figure 10 shows the curve of the stiffness performance index, parallel mechanism has a good stiffness performance when  $0.1 < y < 0.4$  mm, stiffness performance will be better if variable  $y$  is bigger in this area.



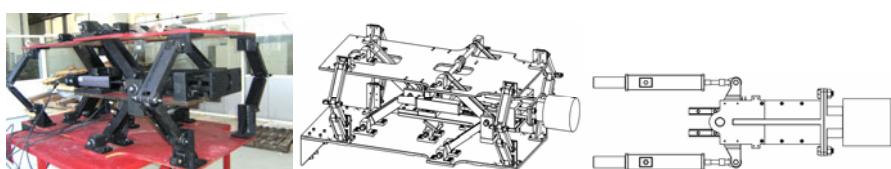
**Fig. 10.** Curve of the stiffness performance index

### 3.3 Optimal Work Space

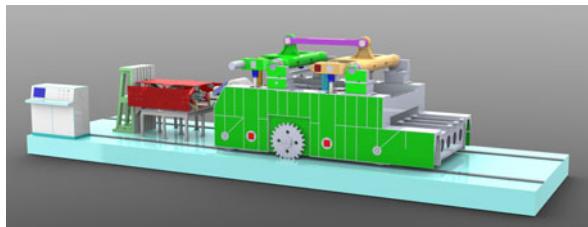
Based on the dexterity, force capability and stiffness performance index analysis, we can get the optimal work space:  $0.1 \leq y \leq 0.4$  mm.

## 4 Applications

As shown in Figure 11, 12<sup>10</sup> the above parallel mechanism can be designed as the forging manipulator. The forging manipulator drove by electric cylinders includes control system, power system, and mechanical system. The idea of reproduce working environment is: first of all we can get the blank's deformation rule in the forging process by CAE software (ABAQUS, DEFORM, etc.), and then the control system changes the blank's deformation rule into every cylinder's force, displacement control signals, so cylinders can drive the mechanical system with 2 DOF to simulate force and movement of the forging blank, finally force and motion is transmitted to the forging manipulator for testing and calibration.



**Fig. 11.** The forging simulator



**Fig. 12.** The forging simulator platform

## 5 Conclusion

In this paper, forging manipulator DDS750 is the study object, and it's forging compliance process is analyzed, the analysis shows that force capability and compliance capability of forging manipulator in passive compliance direction is especially important, then based on the above analysis a parallel mechanism with a good force capability performance is designed as the forging simulator to reproduce working environment, by simulating force and movement of the forging blank, and the optimal work space of the mechanism is obtained based on the dexterity, force capability and stiffness performance index. Finally, an example of the simulator is introduced, which is a new approach for the testing and calibration of the heavy forging manipulator.

## Acknowledgements

The research is supported by National Basic Research Program of China (973 Program) under research grant No.2006CB705400.

## References

1. Fu, Z., Wang, H., Su, S.: The importance and key techniques for load/energy calibration of forging equipment with large capacity. *Mechanical Science and Technology* (6), 890–891 (2001)
2. Zhu, H., Xu, Q., Liao, S., et al.: Experimental study on working parameters of EPB shield machine. *Chinese Journal of Geotechnical Engineering* (5), 553–557 (2006)
3. Lin, Y., Xu, L., Li, W., et al.: Semi-physical simulation test-bed for electro-hydraulic proportional pitch-controlled wind turbine system. *China Mechanical Engineering* (8), 667–670 (2005)
4. Chang, T., Cong, D., Ye, Z., et al.: Simulation on HIL ground experiment simulator for on-orbit docking. *Acta Aeronautica et Astronautica Sinica* (4), 975–979 (2007)
5. Salisbury, J.K., Craig, J.: Articulated hands: Kinematic and force control issues. *International Journal of Robotics Research* 1(1), 4–12 (1982)
6. Gosselin, C.M., Angeles, J.: A global performance index for the kinematic optimization of robotic manipulators. *Transaction of the ASME, Journal of Mechanical Design* 113, 220–226 (1991)
7. Xiong, Y., Ding, H., Liu, E.: *Robotics*. China Machine Press, Beijing (1993)
8. Liu, X.: The relationships between the performance criteria and link lengths of the parallel manipulators and their design theory. *Yanshan University, Qinhuangdao* (1999)

# Field Calibration Method of Camera for Measuring the Dimensions of Large Forging

Wei Liu, Jian Du, Bangguo Wang, Xinghua Jia, Shuangjun Liu, and Zhenyuan Jia

Key Laboratory for Precision and Non-traditional Machining Technology of the Ministry of Education, Dalian University of Technology, Dalian 116024, P.R. China  
lw2007@dlut.edu.cn

**Abstract.** Stereo vision technology has the potential for measuring the dimensions of hot large forging. However, under the high temperature operation circumstance in forging workshop, it is difficult to calibrate the stereo vision sensor by the traditional calibration method, such as planar patterns method and self-calibration method. In this paper, a field calibration method of binocular stereo vision sensor is presented by virtue of projected patterns from a projector, for improving the dimensional measurement precision of large forging. The intrinsic parameters of CCD camera can be obtained by capturing the projected patterns during orthogonal movements of camera, and recognizing the feature points of patterns, while the extrinsic parameters of CCD camera is measured by detecting the distance between the feature points of patterns with the electronic theodolite. Experiments on the field calibration of camera at forging workshop are conducted. And the experimental results show that the relative error of dimensional measurement reaches 0.363%, which indicates that the proposed calibration method is effective for calibrating the dimensional measurement system of large forging.

**Keywords:** Stereo vision; field calibration method; projected pattern; feature points.

## 1 Introduction

Stereo vision system, which is well-known for smart and easy operation, has already been employed for measuring the dimension of large forging in time for controlling precisely in the process of manufacturing. However, radiation and oxide coating of large forgings in high temperature result in that the profile of forging is unclear in shot image from stereo vision system, which makes it difficult to extract the feature points that are desirable in the dimensional calculation. To solve this problem, we have proposed the spectrum selective method to capture the clear image of large forging [1]. Then, in order to acquire further the dimensions of large forging by stereo vision system, the camera must be calibrated, and the calibration results affect markedly the measurement precision. Therefore, a field calibration method of camera, which is practicable under the high temperature operation circumstance in forging workshop, is necessary for dimensional measurement of large forging.

According to the traditional calibration methods based on the stereo vision technology, the camera is usually calibrated by using the 2D or 3D reference objects, including the planar pattern board or calibration block [2], which can usually obtain high calibration precision. Zhang et al [3] has established the calibration method using the planar pattern, which is easy for manufacturing. However, due to the high temperature operation circumstance, it is difficult to calibrate newly when the vision field of camera is changed, according to the dimensional measurement requirement of large forging. Another calibration method proposed by Sun et al [4], can operate under the circumstance of wide vision field by baseline ruler. However, the baseline ruler has to be placed in different positions, which is also difficult to operate in forging workshop.

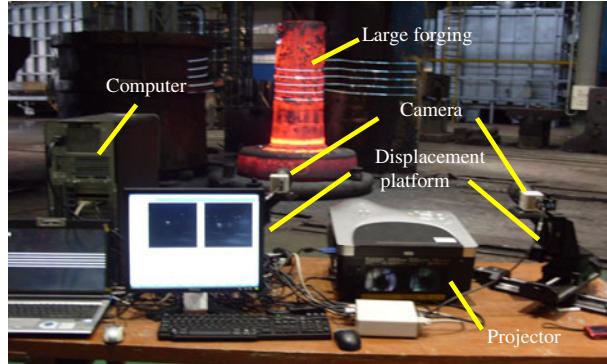
The self-calibration method of camera, which can calibrate the camera only by using the relationship among the intrinsic parameters of the cameras, is not accurate [5-7]. Zhang et al proposed a calibration theory based on the parallel vanishing points realized the calibration of the binocular stereo vision sensor [8]. However, the initial conditions such as the ideal parallel cannot be obtained in forging workshop.

In this paper, we present a field calibration method of CCD camera based on projected patterns technique. The binocular vision measurement system for the dimension of large forging is introduced in Section 2. Thereafter, the calibration method of intrinsic and extrinsic parameters of camera is elaborated in detail in Section 3. In Section 4, the projected patterns for detecting intrinsic and extrinsic parameters are designed. Then, the experiments on field calibration of camera and dimensional measurement are given in Section 5. Finally, the conclusion of this paper is given in Section 6.

## 2 Binocular Stereo Vision Measurement System for Dimension of Large Forging

The spectrum selective method is employed to capture the clear and high-quality image of large forging [1]. The light stripes are projected on the forging, and images of the forging with the light stripes can be captured by the left and right cameras with the spectrum selective method, respectively. Then, the feature points of the stripes can be extracted and matched with the calibration parameters to calculate the dimension of the forging.

The dimensional measurement system for large forging, as shown in Fig.1, consists of two digital CCD-cameras (REDLAKE, ES4020) with a low-pass filter (BD, 450SP) in front of each camera, a DLP projector (3M, PD80X), a computer and two electrically-controlled platforms. The distance between the two digital CCD-cameras is set by 1m, which can obtain low error [9]. The DLP projector is arranged between the two cameras. The cameras and projector are both connected to a computer with control wires. In addition, two electrically-controlled platforms are used to install cameras, and change their positions and angles for calibrating the intrinsic and extrinsic parameters of camera, which is introduced in detailed in Section 3 and 5.



**Fig. 1.** The dimensional measurement system for large forging based on binocular vision

### 3 Calibration Method of Camera Parameters

#### 3.1 Camera Parameters

In this paper, camera model involved is described by the pinhole model. As shown in Fig. 2, four coordinates are given as follows: world coordinate system is  $o_w x_w y_w z_w$ ; computer image coordinate system is  $o_f u v$ , with the scales of pixel, where origin  $o_f$  is at the top left corner of the CCD image plane; image coordinate system is  $O'XY$ , with the scales of millimeter, where origin  $O'$  is defined at the intersection point of the camera optical axis and the image plane, and  $X$  and  $Y$  axes are parallel with  $u$  and  $v$  axes respectively; camera coordinate system is  $Oxyz$ , where center point is the point  $O$  (optical axis center), moreover,  $x$  and  $y$  axes are parallel with  $X$  and  $Y$  axes respectively. The 3D coordinate of  $P_0$  in the established camera coordinate system can be written as  $(x, y, z)$ .

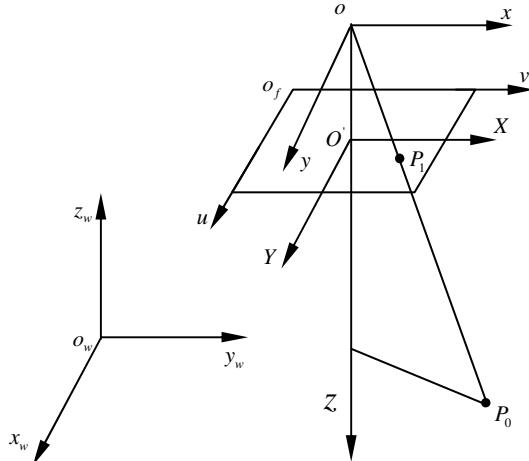
According to the linear pinhole model of camera, the perspective transformation of coordinate system from world to image system can be expressed as [9]:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & u_0 & 0 \\ 0 & f_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

where  $f_u = f / dx$ ,  $f_v = f / dy$ ;  $dx$  and  $dy$  represent the physical dimensions of every pixel on  $X$  and  $Y$  direction, respectively;  $f$  is the focus of camera;  $f_u, f_v, u_0, v_0$  are just the intrinsic parameters of camera to be calibrated; rotation matrix  $R = R_{3 \times 3}$  and translation vector  $T = [t_x \ t_y \ t_z]^T$  describe the position relationship between the camera coordinate system and world coordinate system, which are just the extrinsic

parameters of camera to be calibrated. Moreover, the intrinsic parameters matrix can be given by:

$$K = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}.$$



**Fig. 2.** Schematic of the linear camera model

According to the mapping relationship between the coordinate systems, the following formula can be given by:

$$\frac{f}{z} = \frac{(u - u_0)dx}{x} = \frac{(v - v_0)dy}{y} \quad (2)$$

where \$(u, v)\$ is the coordinate of spatial point \$(x, y, z)\$ in image coordinate system, and \$(u\_0, v\_0)\$ is the intersection point of optical axis and image coordinate system.

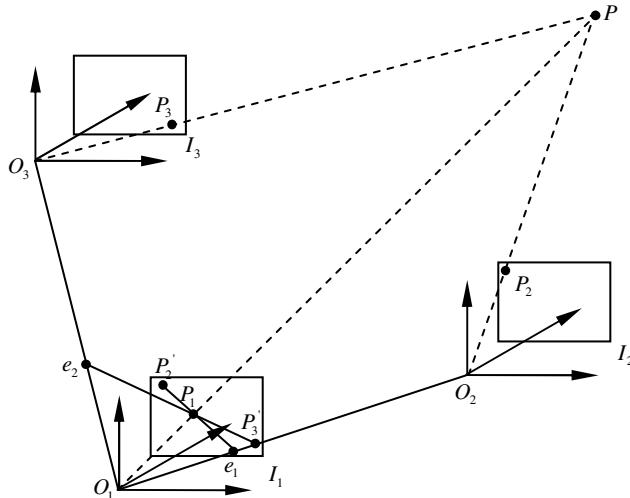
### 3.2 Calibration Method of Camera Intrinsic Parameters

In this paper, camera intrinsic parameter is calibrated by four group of orthogonal motion of camera with simultaneously changing the poses of camera, driven by the electrically-controlled platform.

Figure 3 illustrates the calibration principle of camera intrinsic parameters. \$I\_1\$ and \$I\_2\$ are the image before motion and the image after motion respectively, as shown in Fig. 3. The points \$P\_1\$ and \$P\_2\$ are the projects of the same space point \$P\$ in the images \$I\_1\$ and \$I\_2\$. Line \$\overline{P\_1P\_2}\$ is defined by the corresponding point line, where \$P\_2\$ is the

corresponding point of  $P_2$  in the image  $I_1$ , with the same coordinate as that in the image  $I_2$ . The corresponding point lines of the image before motion and the one after motion cross the each other at the pole, and the vector  $O_1e_1$  from the optical center to the pole  $e_1$  is parallel with the motion vector [10]. According to (2), we can obtain:

$$O_1e_1 = ((u - u_0)dx, (v - v_0)dy, f)^T = f((u - u_0)/f_u, (v - v_0)/f_v, 1)^T \quad (3)$$



**Fig. 3.** Calibration principle of camera intrinsic parameters

As shown in Fig. 3, when the camera is controlled to conduct one group orthogonal motion in one plane,  $O_1e_1$  and  $O_1e_2$  represent two orthogonal vector, moreover,  $O_1e_1 \cdot O_1e_2 = 0$ . Then, when three group motions of camera in the same as the one above is finished, according to (3), we can obtain:

$$(u_{11} - u_0)(u_{12} - u_0)dx^2 + (v_{11} - v_0)(v_{12} - v_0)dy^2 + f^2 = 0 \quad (4)$$

$$(u_{21} - u_0)(u_{22} - u_0)dx^2 + (v_{21} - v_0)(v_{22} - v_0)dy^2 + f^2 = 0 \quad (5)$$

$$(u_{31} - u_0)(u_{32} - u_0)dx^2 + (v_{31} - v_0)(v_{32} - v_0)dy^2 + f^2 = 0 \quad (6)$$

$$(u_{41} - u_0)(u_{42} - u_0)dx^2 + (v_{41} - v_0)(v_{42} - v_0)dy^2 + f^2 = 0 \quad (7)$$

where  $(u_{ij}, v_{ij})$  ( $i = 1, 2, 3, 4; j = 1, 2$ ) are the coordinates of the poles in each image coordinate system. Denote that  $a = u_0$ ,  $b = v_0 f_u^2 / f_v^2$ ,  $c = f_u^2 / f_v^2$ , then, (4), (5), (6), and (7) can be written as:

$$\begin{bmatrix} u_{11} + u_{12} - u_{21} - u_{22} & v_{11} + v_{12} - v_{21} - v_{22} & v_{21}v_{22} - v_{11}v_{12} \\ u_{21} + u_{22} - u_{31} - u_{32} & v_{21} + v_{22} - v_{31} - v_{32} & v_{31}v_{32} - v_{21}v_{22} \\ u_{31} + u_{32} - u_{41} - u_{42} & v_{31} + v_{32} - v_{41} - v_{42} & v_{41}v_{42} - v_{31}v_{32} \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} u_{11}u_{12} - u_{21}u_{22} \\ u_{21}u_{22} - u_{31}u_{32} \\ u_{31}u_{32} - u_{41}u_{42} \end{bmatrix} \quad (8)$$

The four groups of the orthogonal motions in the same plane can be classified into four modes: three tilting, three panning, two panning and one tilting, one panning and two tilting. The matrix equation of one panning and two tilting is proved to exist an unique stable solution [11]. Therefore, we select the orthogonal motion mode of one panning and two tilting to calibrate the intrinsic parameters of camera.

### 3.3 Calibration Method of Camera Extrinsic Parameters

The origin of the world coordinate system is set by the origin of the camera coordinate. Then, the extrinsic parameters can be acquired according to the position relationship of the left and right cameras, which can be solved by the essential matrix  $E$ . However, the position relationship deduced from  $E$  is not the actual relation, but the one with a proportional factor. Thus, we employ a projector to project pattern onto the pillars of the forging press, and detect the actual distance between the feature points of patterns by the electronic theodolite. Then, the extrinsic parameters of camera can be obtained by solving of the proportional factor between the distance deduced from the matrix  $E$ , and the actual distance.

The eight or more than eight feature points in the left and right images are extracted and matched by the *SIFT* algorithm, which can provide high matching precision. Then, the fundamental matrix  $F$  can be calculated, which reflects the interrelationship of these corresponding feature points in the left and right images. Suppose that  $x_l$  and  $x_r$  represent the matched feature point of left and right image respectively, then  $x_l^T F x_r = 0$ .

The essential matrix  $E$ , which is equal to the product of the intrinsic parameters  $K_l$  (left camera),  $K_r$  (right camera) and the fundamental matrix  $F$ , is the special form of the fundamental matrix  $F$  in the normalization image coordinate, and the relationship can be given by:

$$E = K_r^T (K_r^T [T]_x R K_l^{-1}) K_l = K_r^T F K_l \quad (9)$$

where  $[T]_x$  is the anti-symmetrical matrix of  $T$ .

The matrixes  $R$  and  $T$  can be solved by the eigenvalue decomposition of  $E$  [12],  $E = USV^T$ , where  $S = \text{diag}(k, k, 0)$ . Then the matrix  $R$  and  $T$  can be expressed as:

$$R \approx U Q V^T, [T]_x = V Z V^T \quad (10)$$

$$\text{where } Q = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, Z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

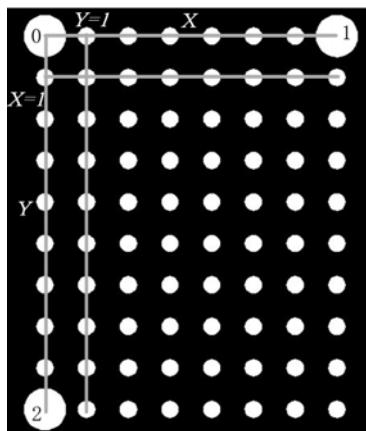
According to (10) and the relation of  $[T]_x T = 0$ , the translation vector between two cameras can be described by  $T = U(0, 0, 1)^T$ . The solution of (10) is obtained by the assumption of  $\|T\| = 1$ . Since the epipolar line constraint cannot define the proportional factor of 3D information, the actual translation vector  $T'$  can be given by  $T' = sT$ , while the actual rotation matrix is just the matrix  $R$ .

## 4 Projected Pattern Design

Projected pattern design is the key point of camera calibration. Thus, in order to calibrate the intrinsic and extrinsic parameters of camera efficiently in forging workshop, a kind of practical pattern projected by a projector is designed in this section.

### 4.1 Projected Pattern Design for Intrinsic Parameters

The projected pattern for calibrating the intrinsic parameters of camera is composed of circularity array, as shown in Fig. 4, where three larger-size circularities for locating the projected pattern are placed in the corner of the array, which is the location datum mark for the other circularities.



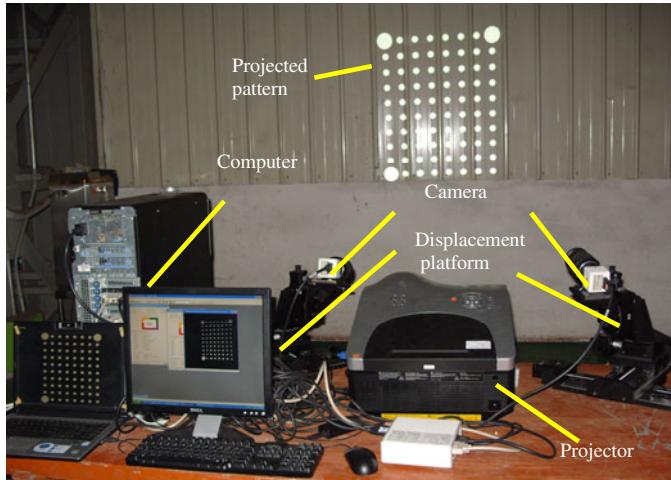
**Fig. 4.** Coordinates identification of the pattern for calibrating intrinsic parameters

The location process of the circularity array consists of the pretreatment, automatic threshold segmentation, noise reduction, image filtering, circularity identification, elliptic center extraction, and topological location.

After the identification of the all circularity, the centers of three larger-size circularities are recognized and located by the circular area. For each center, the summation of the distance to the other centers is calculated, and the minimum is defined by the origin  $O$ . The nearer center point to origin is defined by the 1<sup>st</sup> point, and the other by the 2<sup>nd</sup> point. Consequently, the  $OXY$  axes are determined. The X axis and the Y

axis are the lines from origin to the 1<sup>st</sup> point and the 2<sup>nd</sup> point respectively, as shown in Fig. 4. The circle centers of pattern are lined to fabricate the line  $Y=1, 2, \dots, n$  line, as well as the line  $X=1, 2, \dots, n$ . If a point is both on the  $X=m$  line and on  $Y=n$  line simultaneously, its coordinate is defined by  $(m, n)$ , which finishes the location of all the points on the pattern.

During the motion of the camera, different images of the pattern projected onto the wall of workshop are captured, as shown in Fig. 5. It is noticeable that the pattern would fully fill the screen to reduce the distortion error.



**Fig. 5.** Field picture of camera intrinsic parameters calibration

## 4.2 Projected Pattern Design for the Extrinsic Parameters

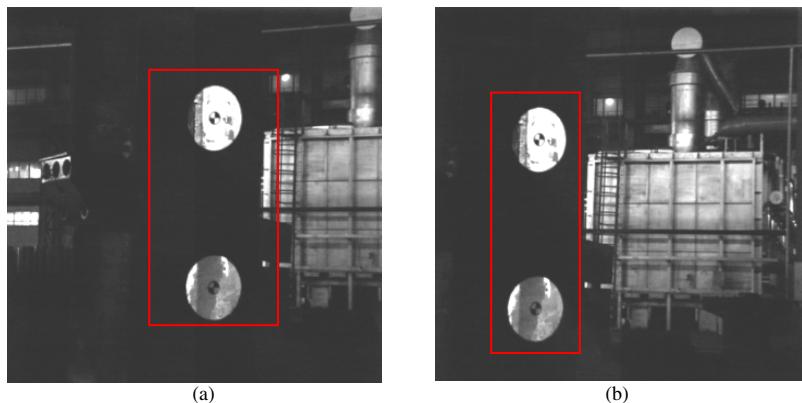
The projected pattern for calibrating extrinsic parameters of camera is composed of two circularities, which is suitable for projection and feature point extraction, as shown in Fig. 6. Moreover, the actual distance between the feature points of patterns projected onto the pillar of forging press can be detected by two electronic theodolites (Kelida, DJ-02C), for acquiring the proportional factor of translation vector.

The circle center coordinates of circularities are determined by the same method as the intrinsic parameters one. According to (9) and (10), the rotation matrix  $R$  and the translation vector  $T$  are obtained, and with combining  $K$  and  $K'$ , the distance  $L'$  between the two center feature points is reconstructed. Then, the proportional factor can be written as:

$$s = L / L' \quad (11)$$

where  $L$  is the actual distance measured by the theodolite.

In addition, in order to guarantee measurement precision of  $L$ , several experiments should be conducted for the mean value of the actual distance.



**Fig. 6.** Pattern for the extrinsic parameters calibration (a) left image (b) right image

## 5 Experiments and Analysis

The proposed method in this paper is verified by the experiments on field calibration of camera in forging workshop. The image size is selected by  $2048 \times 2048$  pixels. The two CCD cameras are fixed on two electrically-controlled platforms for calibrating the camera parameters.

**Table 1.** Motion parameters for the four orthogonal motions

	Camera pose	Location parameter (x,y)/mm		
First group	Tilting angle 15°/Panning angle 40°	0, 100	0, 0	100, 0
Second group	Tilting angle 0°/ Panning angle 40°	100, 0	0, 0	0, 100
Third group	Tilting angle 0°/ Panning angle 50°	0, 100	0, 0	100, 0
Fourth group	Tilting angle -15°/ Panning angle 50°	100, 0	0, 0	0, 100

In the process of the calibration of the intrinsic parameters, the orthogonal motions of one panning and two tilting are performed, as shown in Table 1. The projected patterns onto the wall of workshop are captured by two CCD cameras, and the calculated intrinsic parameters of two cameras are listed in Table 2. Moreover, according to the actual distance of the feature points of pattern, the extrinsic parameters of two CCD camera is also shown in Table 2.

In order to testify the calibration precision of the proposed method in this paper, a calibration experiment with the planar pattern is performed in the forging workshop. The actual length value of each rectangle in planar pattern is 60mm. The measured value of the length is 59.782mm, and the relative error of dimensional measurement is 0.363%, which indicates that the proposed method can calibrate well the intrinsic and extrinsic parameters of camera with high precision.

**Table 2.** Calibration result of camera parameters with the projected pattern

	Left camera	Right camera
Intrinsic parameter $K$	$\begin{bmatrix} 6036.56 & 0 & 1043.28 \\ 0 & 6036.56 & 1122.85 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 5979.11 & 0 & 1122.86 \\ 0 & 5979.11 & 1097.87 \\ 0 & 0 & 1 \end{bmatrix}$
Rotation matrix $R$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -0.9251 & -0.0353 & 0.3781 \\ 0.0663 & -0.9954 & 0.0692 \\ 0.3739 & 0.0891 & 0.9232 \end{bmatrix}$
Translation vector $T$	$[0 \ 0 \ 0]^T$	$[1072.4 \ 215.1570 \ 6062.3]^T$

## 6 Conclusions

In this paper, a field calibration method of binocular stereo vision system is presented, for measuring the dimension of hot large forging. The intrinsic parameters of CCD camera are calibrated by combining the projected pattern and orthogonal motion of camera, and the extrinsic by the projected pattern and the electronic theodolite. Experimental results in forging workshop indicate that the proposed method is practicable and suitable for calibrating the CCD camera of dimensional measurement system of large forging, under the high temperature operation circumstance. Further study is suggested focusing on improving further the calibration precision by the nonlinear optimization of camera parameters.

## Acknowledgements

This paper is supported by the National Basic Research Program of China 973 Project under Grant No. 2006CB705406.

## References

1. Jia, Z.Y., Wang, B.G., Liu, W., Sun, Y.W.: An Improved Image Acquiring Method for Machine Vision Measurement of Hot Formed Parts. *Journal of Materials Processing Technology* 210(2), 267–271 (2010)
2. Tsai, R.Y.: An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 364–374. IEEE Press, New York (1986)
3. Zhang, Z.Y.: A Flexible New Technique for Camera Calibration. *IEEE Trans. Pattern Analysis and Machine Intelligence*. 22(11), 1330–1334 (2000)
4. Sun, J.H., Wu, Z.Y., Liu, Q.Z., Zhang, G.J.: Field calibration of stereo vision sensor with large FOV. *Optics and Precision Engineering* 17(3), 633–640 (2009) (in Chinese)
5. Maybank, S.J., Faugeras, O.A.: A Theory of Self-Calibration of a Moving Camera. *International Journal of Computer Vision* 8(2), 123–151 (1992)

6. Li, Y., Hung, Y.S., Lee, S.: A Stratified Self-Calibration Method for Circular Motion in Spite. *Image and Vision Computing* 26, 731–739 (2008)
7. Qian, G., Chellappa, R.: Bayesian Self-Calibration of a Moving Camera. *Computer Vision and Image Understanding* 95(3), 287–316 (2004)
8. Zhang, J.X., Duan, F.J., Ye, S.H.: Two step Method for Binocular Sensor Self-Calibration. *Opto-Eletronic Engineering* 25(5), 37–41 (1998)
9. Zhang, G.J.: Machine Vision. Science Press, Beijing (2005) (in Chinese)
10. Ma, S.D.: A Self-Calibration Technique for Active Vision System. *IEEE Transactions on Robotics and Automation* 12(1), 114–120 (1996)
11. Yang, C.J., Wang, W., Hu, Z.Y.: An Active Vision Based Camera Intrinsic Parameters Self-Calibration Technique. *Chinese Journal of Computers* 21(5), 428–435 (1998) (in Chinese)
12. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, London (2000)

# Velocity Control of the Horizontal Buffer System for Heavy Load Forging Manipulator

Jiaoyi Hou, Hua Zhou, Jun Zou, and Xin Fu

State Key Laboratory of Fluid Power Transmission and Control,

Zhejiang University, 310027 Hangzhou, China

houjiaoyi@gmail.com, zhoub@cmee.zju.edu.cn, junzou@zju.edu.cn,  
fu@zju.edu.cn

**Abstract.** This paper introduces the working mode of quick-forging for heavy load forging manipulator. Simulation model of horizontal buffer system is built according to the hydraulic circuit. Based on different control algorithms, three controllers are designed and applied in the simulation model. The performances of the three controllers are studied based on the simulation, and results show that the variable argument proportional integral derivative (VAPID) works best. The other factors that influence the velocity control performance are also analyzed, some suggestions on the improvement of velocity control are provided based on the study.

**Keywords:** Forging manipulator; PID control; Various argument; Velocity control; Quick forging.

## 1 Introduction

As shown in Fig. 1, forging manipulators are heavy industrial robots, whose main task is to hold and manipulate the work piece during the forging process. Its capability of cooperating with the forging press makes the forging manipulator widely used in forging field. Due to this new forging mechanism, the efficiency of the free forging is increased and energy consumption is reduced. Quick-forging is a complicated process, during which the forging manipulator is working closely with the forging press. Advantages such as low energy cost, high production efficiency and quality make quick-forging an important working mode in free forging.

Up to now, there have been very few investigations publicly reported exclusively concentrating on the forging frequency of the manipulator. Simulations show that the forging frequency is related to properties of the raising system and can be increased by improving the raising system [1].

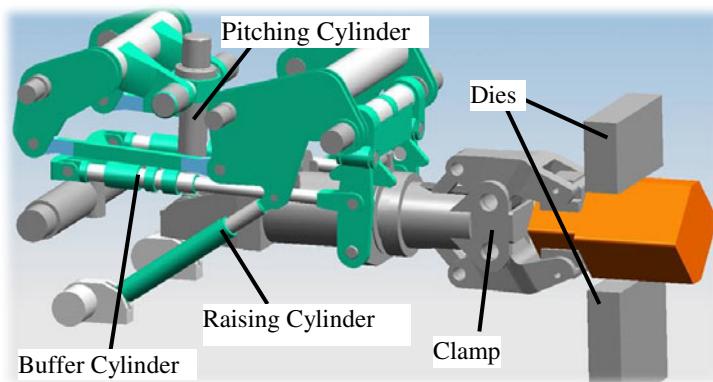
### 1.1 Mechanism of the Forging Manipulator

Special mechanism design of the forging manipulator ensures that the displacement of the clamp in horizontal direction is proportional to that of the horizontal buffer

cylinder, the displacement of the clamp in vertical direction is proportional to that of the raising cylinder and that the angle between the clamp and the horizontal line is proportional to the displacement of the pitching cylinder as shown in Fig. 2. [2] [3] [4].



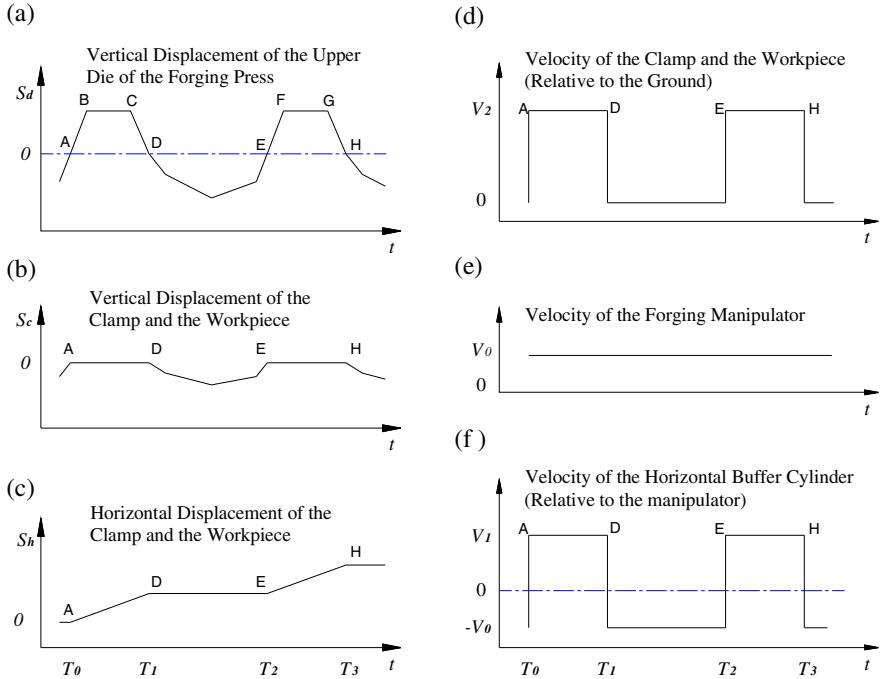
**Fig. 1.** One type of forging manipulator



**Fig. 2.** Simplified mechanism

## 1.2 Quick-Forging Mode

In order to maintain high-quality of the work piece's mechanical property and reduce the energy dissipation of the hot work piece, high forging frequency is needed. The work piece must be kept still in horizontal when the press dies touch it and feeds when the press dies depart from it. As shown in Fig. 3(c), the scope from point A to point D indicates the separation of press dies from the work piece and the feeding of the work piece in horizontal direction. The scope from point D to point E indicates the touch of the press anvils with the work piece while the work piece keeps still in horizontal direction.



**Fig. 3.** Displacement and velocity charts

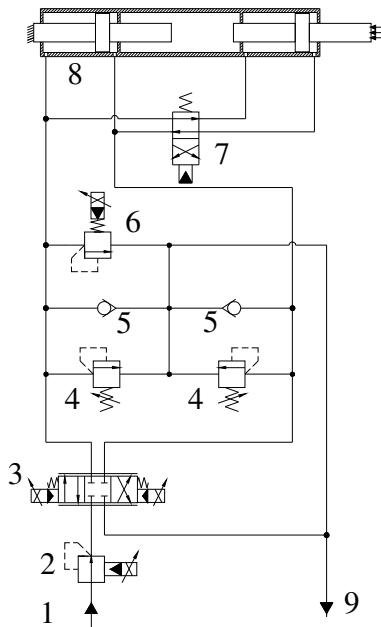
As the manipulator is huge, high frequent start-stop of the manipulator leads to great impact on the walking system. The working condition of the walking system is extremely terrible when the forging frequency is higher than 30 times per minute. The velocity of the clamp  $V_2$  is the sum of the manipulator walking velocity  $V_0$  and the horizontal buffer cylinder velocity  $V_1$ .

It is possible to design a feeding mode with a constant manipulator walking velocity and a variable horizontal buffer cylinder velocity. According to the requirement, the clamp and the work piece must feed and stop at regular intervals to coordinate with the press machine. The velocity charts of the horizontal buffer cylinder and the clamp are shown in Fig. 3(f). The horizontal buffer cylinder must stay in the initial position after times of forging.

## 2 Horizontal Buffer System

Different from the traditional buffer system, the horizontal buffer system in this paper has two main functions: negative buffer function and positive feed function. As shown in Fig. 4, two working functions can be achieved by electromagnetic direction valve 7 and proportional direction valve 3. When the direction valve 7 is energized,

the system works on positive feed function and the velocity of the cylinder can be controlled accurately by the proportional direction valve 3. When the direction valve 7 is de-energized and the proportional direction valve 3 keeps working on the left position, the system works on negative buffer function and the velocity of the cylinder is due to the impact force from the clamp. At buffer function, the buffer force of the cylinder can be regulated by proportional relief valve 6 and the restoring force can be regulated by proportional pressure reducing valve 2.

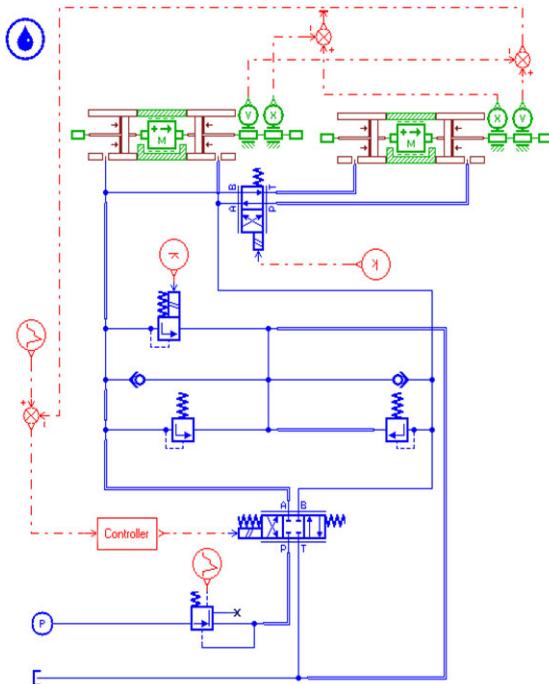


1) Pressure oil, 200bar	2) Proportional pressure reducing valve
3) Proportional direction valve	4) Safety valve
5) Check valve	6) Proportional relief valve
7) Direction valve	8) Horizontal buffer cylinder
9) Return oil, 3bar	

**Fig. 4.** Schematic diagram of the horizontal buffer system

### 3 Simulation Model and Controllers Design

Simulation model of the horizontal buffer circuit presented in section 2 is built in the AMESim software environment, which is shown in Fig. 5. All of the following simulations are based on this model.



**Fig. 5.** Simulation model of the buffer system

At quick-forging mode, the raising system moves along with the forging press and the horizontal buffer system moves along with the walking system. The ideal velocity of the horizontal buffer cylinder is shown in Fig. 3(f). It moves forward at velocity  $V_1$  during stage AD and moves backward at velocity  $V_0$  during stage DE. The velocity of the horizontal buffer cylinder must be controlled precisely in order to avoid the impact from the forging press. However, frequently change of the velocity direction leads to high difficulty in velocity control.

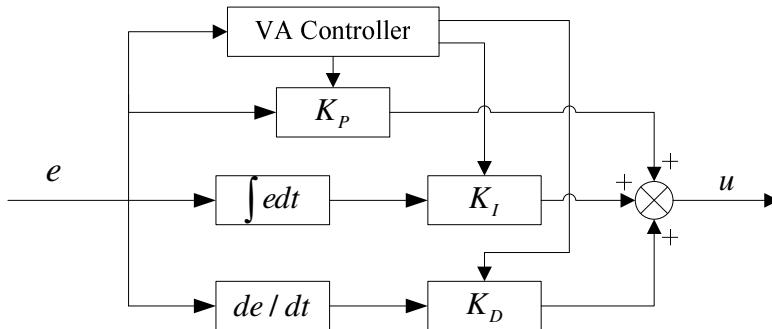
In order to optimize the controller of the velocity control, three controllers are designed based on different control algorithms for the forging manipulator: traditional PID closed-loop velocity control, velocity planning PID closed-loop displacement control and VAPID velocity control. PID control is widely used in displacement and velocity control systems. Velocity planning PID displacement control is a controller that calculating the displacement from the ideal velocity and comparing it with the value detected by the displacement sensor.

VAPID control is an evolution from traditional PID control. The parameters of VAPID control change dynamically with the velocity error during the working of the system [5] [6]. Fuzzy control is adopted to control the parameters of VAPID in this paper. The change rate of the velocity error is ignored in the fuzzy control module in order to simplify the controller, so the membership functions of the fuzzy control are univariate functions of velocity error. Then the VAPID controller can be written as:

$$u = K_{P0}f_P(e)e + K_{I0}f_I(e)\int edt + K_{D0}f_D(e)\frac{de}{dt}, \quad (1)$$

In equation (1),  $K_{P0} = 1500$ ,  $K_{I0} = 2000$  and  $K_{D0} = 6$ .

The control block diagram is shown in Fig. 6.

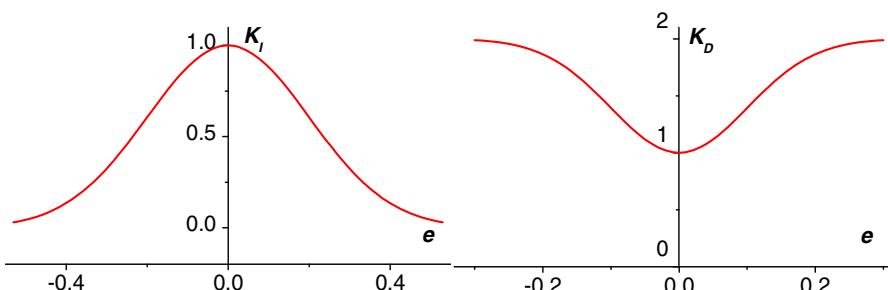


**Fig. 6.** Block diagram of the variable argument PID control

Gaussian membership function is adopted to control parameters  $K_I$  and  $K_D$ , as shown in Fig. 7. Parameter  $K_P$  is designed to be a constant as its variation has little influence on the control property. The membership functions of  $K_I$  and  $K_D$  are given in equation (3) and (4):

$$f_I(e) = \exp(-50e^2), \quad (2)$$

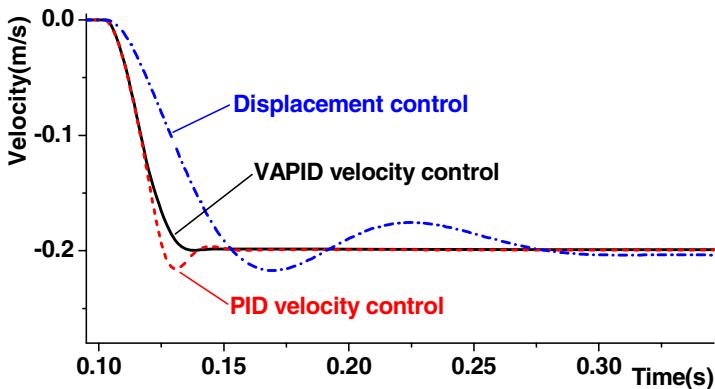
$$f_D(e) = 2 - \exp(-50e^2). \quad (3)$$



**Fig. 7.** Membership diagrams of  $K_I$  and  $K_D$

## 4 Results and Discussion

Serial simulations for velocity control of the horizontal buffer system are carried out based on the model built in section 3. By applying different control algorithms to the system, three different velocity diagrams of the horizontal buffer cylinder are achieved, as shown in Fig. 8. The VAPID velocity control works best as it reaches target value quickly enough and has no overshooting.



**Fig. 8.** Velocity response curves of different control algorithms

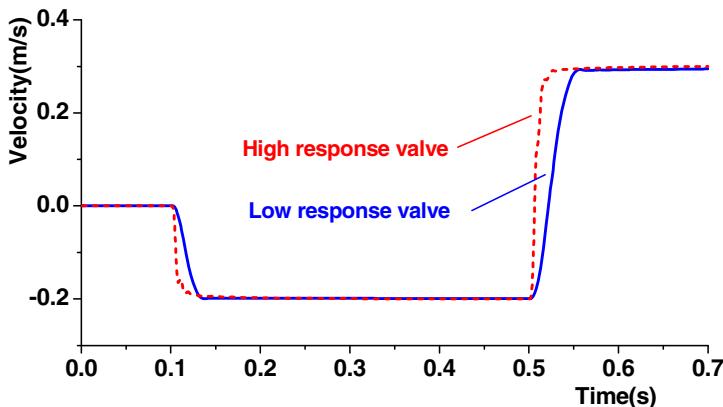
There are several parameters to evaluate the performance of the controller: rise time  $t_r$  which refers to the time required for a actuator moving from 10% to 90% of the target value, overshoot  $M_p$ , adjusting time  $t_s$  of error range  $\pm 2\%$ . The values of the evaluation parameters of different controllers are given in Table. 1.

**Table 1.** Performance index of three controllers

Controller	$t_r$ (s)	$M_p$	$t_s$ (s)
VAPID	0.020	-	0.032
PID	0.016	7.9%	0.037
Displacement	0.034	8.5%	-

There are two main limits which make it hard to reduce the rise time and adjusting time. One is the large mass of the clamp and the work piece, as it limits the acceleration of the horizontal buffer cylinder. The other is the response frequency of the proportional direction valve. In a proportional control hydraulic system, especially that of large flow, the proportional valve always has the lowest response frequency in the whole circuit. So the response time of the hydraulic system couldn't be too short as it takes time to change its direction for the proportional direction valve. Simulations of the velocity control are carried out to study the influence of the proportional valve. The velocity comparison between high response valve and low response valve is

shown in Fig. 9. It is obvious that the dynamic characteristics of the system with high response valve are much better. To a valve, large flow means low response, so multi-valve of small flow in parallel can be used to replace one large flow valve to get a better velocity control.



**Fig. 9.** Velocity response curves with different valves

## 5 Conclusion

(1) Closed-loop velocity control is better than velocity planning closed-loop displacement control, and the velocity control based on VAPID is better than traditional PID. Overshooting is eliminated and adjusting time is reduced by controller based on VAPID algorithm.

(2) The response frequency of the proportional direction valve has great influence on the velocity control. The adjusting time could be greatly reduced by replacing the valve with a higher response one. The hydraulic circuit can be improved by replacing one large flow valve with multiple low flow valves.

**Acknowledgments.** The work described in this paper is supported by National Basic Research Program of China (973 Program, No. 2006CB705400) and Zhejiang Provincial Natural Science Foundation of China (R105008).

## References

1. Jiaoyi, H., Wei, W., Hua, Z., Xin, F., Huayong, Y.: Simulation Study of Quick-Forging for Heavy Load Forging Manipulator. In: 7th International Fluid Power Conference, Workshop Group A, Aachen, pp. 143–152 (2010)
2. Yunpeng, R., Chongshao, L., Qingkai, H., Tianxia, Z., Bangchun, W.: Simulated Comparison on Kinematics Properties of Two Typical Mechanisms of Forging Manipulator. In: 4th International Conference on Mechatronics and Information Technology: Mechatronics, MEMS, and Smart Materials, Japan (2007)

3. Chen, W., Zhang, X., Cui, Z.: Numerical Simulations of Open-Die Forging Process for Manipulator Design. In: Xiong, C.-H., Liu, H., Huang, Y., Xiong, Y.L. (eds.) ICIRA 2008. LNCS (LNAI), vol. 5315, pp. 650–658. Springer, Heidelberg (2008)
4. Chen, G., Wang, H., Lin, Z., Liu, G.: Performance Analysis of a Forging Manipulator Based on the Composite Modeling Method. LNCS. Springer, Heidelberg (2008)
5. Zhong, Q., Xie, J., Li, H.: PID Controller with Variable Arguments. *Information and Control* 4, 27–277 (1999)
6. Yang, J., Zhao, Y., Wu, X., Yang, Q.: A Fuzzy Varying Parameters Control Method. *Process Automation Instrumentation* 9, 16–18 (2004)

# Performance Optimization of Forging Manipulator during the Whole Forging Stroke

Deshi Liu<sup>1</sup>, Gang Li<sup>1</sup>, Xinglin Guo<sup>1</sup>, Yonggang Shang<sup>2</sup>, and Dehao Liu<sup>2</sup>

<sup>1</sup> State Key Laboratory of Structural Analysis of Industrial Equipment,  
Dalian University of Technology, Dalian, 116024, China

<sup>2</sup> Huadong Enginnering machinery co., Ltd,  
Qingdao, 266071, China

Ligang@dlut.edu.cn, Liudeshi@dl.cn

**Abstract.** The dynamic behavior of the forging manipulator during the whole forging stroke is investigated. The focus is on the dynamic force on the forging manipulator at the vertical direction. The effects of the resilient system parameters are studied, including the initial height of air chamber of the accumulator, the orifice area of the resilient cylinder. Results show that the dynamic force exerted to the clamp increases with the decreasing of initial height of air chamber of the accumulator and orifice area. On the other hand, the working efficiency of forging manipulator needs to be guaranteed. An approximately analytical relationship between working efficiency and resilient parameters is established by design of experiment (DOE) based response surface model. An optimization procedure by adaptive simulated annealing algorithm is executed. The goal of the optimization is to reduce the force transmitted to the clamp without losing working efficiency.

**Keywords:** Forging manipulator, Open die forging, Compliance control.

## 1 Introduction

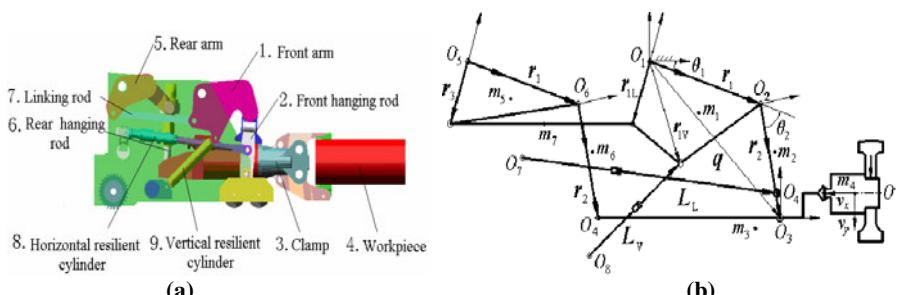
The automation of forging process makes the forging manipulator necessary to improve efficiency and safety. One of the most important functions of the manipulator is to faithfully follow the press motions during the metal shaping period .A resilient system is provided for this purpose. An important task in the design of the manipulator resilient systems is to reduce the vertical loads transmitted to the clamp in which limited value is allowed. Active stragedies have been tried to accomplish this task. Lilly [1] used the efficient O(N) recursive algorithm for the dynamic simulation of the forging manipulator, and the neural network for the compliance control of the manipulator. Nye et al. [2] introduced a prediction model of real-time process characterization of workpiece, and provided a control strategy for open die forging automation. The active method is more versatile due to the high programmability. However, inaccuracies may arise from the inevitable uncertainties of the environment and measurement errors in the vibrant forging conditions. The Pahnke Company [3]

reported a semi-active method to make the manipulator to move downward with the upper die according to the time signal from the forging press. But the application is limited to the manipulator serving for the Pahnke made sinusoidal pump press so far. Nowadays, for the design of the manipulator serving for the hydraulic press, the passive resilient system is still popular in the forging industry due to the simplicity and reliability, which works through the spring damping components [4]-[6]. Wang [7] calculated the forces transmitted the clamp by combination of different commercial softwares. Liu and Li [8] studied the effects of the manipulator parameters on the bending moment transmitted to the forging clamp due to the metal forming. Chen et al. [9] simulated the force by the commercial FEM software Deform-3D. From the review of the preceding literature, most of the work focuses on the dynamic behavior of the forging manipulator during the down stroke of the hydraulic press, and little attention has been paid to the dynamic behavior of the manipulator during up stroke of the hydraulic press. An important issue in this period is that the forging clamp needs to rebound and cease to the original working position within a shorter duration after the leaving of the upper die, so the manipulator can go on to the next forging stroke as soon as possible. The rebound time could be used as an index of the working efficiency of manipulator.

The dynamic behavior of the forging manipulator with the passive resilient system during the whole forging stroke is investigated in this paper. The focus is on the vertical force transmitted to the manipulator, which is obtained by an inverse dynamic analysis of the manipulator mechanism. The actual pressure equation of resilient cylinder is derived, which is validated by the workshop experiment. The effects of the resilient system parameters on the dynamic behavior of the forging manipulator are studied, including the initial height of the air chamber of the accumulator, the orifice area of the resilient cylinder. An optimization procedure is executed to minimize the force transmitted to the manipulator without losing working efficiency.

## 2 Dynamic Modeling of the Forging Manipulator System

Fig. 1(a) is a practical forging manipulator, whose mechanism system is shown in Fig. 1(b). The workpiece (4) is gripped by the clamp (3), which is suspended on a moving



**Fig. 1.** (a) CAD model of the forging manipulator. (b) mechanism diagram of the forging manipulator.

carriage. The clamp is operated by means of a planar parallel linkage system (1), (2), (5), (7), (8), which keeps the clamp horizontal during the forging operation. The dynamic equation of manipulator is given in the Cartesian space as

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{G}(\mathbf{q}) = \mathbf{F}_{\text{ext}} + \mathbf{J}^T \mathbf{F}(\mathbf{q}, \dot{\mathbf{q}}). \quad (1)$$

where  $\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}}$ ,  $\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})$ ,  $\mathbf{G}(\mathbf{q})$  are the inertial matrix, Coriolis and centripetal effects, and gravity, respectively.  $\mathbf{q}$  is the vector of generalized coordination of point  $O_3$  as shown in Fig. 1(b).  $\mathbf{F}_{\text{ext}} = [F_x, F_y]^T$  represents the extra forces exerted to the workpiece by the hydraulic press.  $\mathbf{J}_T$  is the Jacobian transformation matrix between the joint space and the Cartesian space.  $\mathbf{F}(\mathbf{q}, \dot{\mathbf{q}})$  represents the reaction forces of the two resilient cylinders (8) (9). If the motion curve of the clamp  $(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$  during the compression period is given,  $\mathbf{F}_{\text{ext}}$  could be solved according to Eq. (1). During the up stroke period,  $\mathbf{F}_{\text{ext}} = \mathbf{0}$  because the upper die has left the workpiece and no extra forces occur. In order to measure the dynamic effects due to the press motion, the force ratio  $R_y$  is introduced as

$$R_y = F_y / m_4 g \quad (2)$$

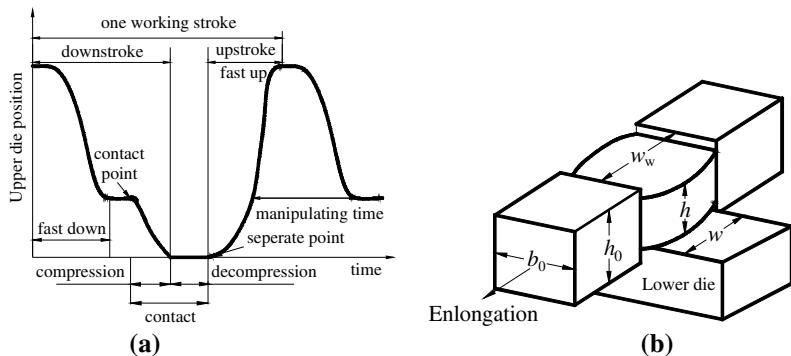
where  $F_y$  is the vertical force,  $m_4 g$  is the workpiece static weight.

### 3 Motion Curve of the Clamp during the Compression Period

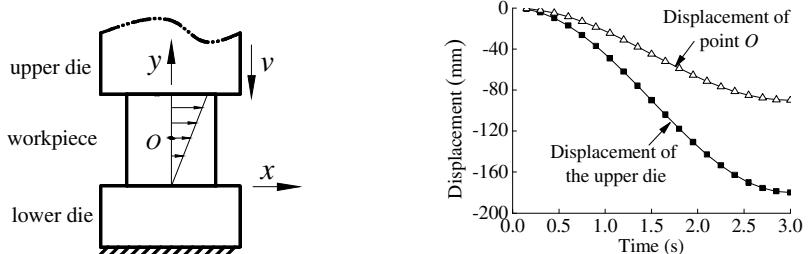
In this section, the motion curve of the forging clamp  $\mathbf{q}$  is derived to solve  $\mathbf{F}_{\text{ext}}$  during the compression period according to the forging art.

#### 3.1 Motion of the Clamp in the Vertical Direction

The whole working stroke for a conventional forging press includes four stages as shown in Fig. 2(a): the empty down period, compression period, decompression period and upstroke period. During the empty down period, suppose the workpiece is just touching the lower die and no contact force happens at the beginning. The upper die moves downward rapidly, until it nearly reaches the upper surface of the workpiece. Then the upper die slows down almost to zero velocity to reduce the impact on itself. During the compression period, the upper die applies pressure to the workpiece (4) towards the lower die. The pressing motion forces the forging manipulator to move downwards against the reaction forces of resilient cylinders (8) and (9). During the decompression period, the press stays at a prescribed position. Finally, during the upstroke period, the upper die moves upward rapidly and the manipulator goes through a free vibration.



**Fig. 2.** (a) Motion curve of the upper die during the whole forging stroke. (b) Geometry of the workpiece during the metal forming process.



**Fig. 3.** (a) Vertical velocity distribution along the  $y$  axis (right view of Fig. 1 (b)). (b) Motion curves of the upper die and point  $O$  in the vertical direction during the compression period.

The motion of upper die during the compression period is approximated by a damped sine function which is expressed mathematically as

$$u_y(t) = -\frac{d}{2} \left( 1 - \cos \frac{\pi t}{\tau} \right) \quad (3)$$

where  $\tau$  is the duration of the compression period and  $d$  is the displacement amplitude of the upper die. The lower die is supposed to be stationary. Due to the symmetry of the workpiece, the vertical velocity of point  $O$  which is on the centre line of the workpiece is half of the velocity of the upper die as shown in Fig. 3(a). Also, the displacement of point  $O$  is half of that of the upper die in the vertical direction as shown in Fig. 3(b), which is as same as the motion of point  $O_3$ .

### 3.2 Motion of the Clamp in the Horizontal Direction

In practice, elongation in the horizontal direction occurs with the decreasing height of the workpiece as shown in Fig. 2(b), which causes the horizontal movement of the

clamp. Starting with a rectangular workpiece of the initial height  $h_0$ , width  $b_0$  and length  $l_0$ , the width and length after one forging stroke could be expressed as

$$b = b_0 \left( \frac{h_0}{h} \right)^s \quad (4)$$

$$l = l_0 \left( \frac{h_0}{h} \right)^{1-s} \quad (5)$$

The value of the spread coefficient  $s$  was determined experimentally by Tomlinson and Stringer [10], who fitted the empirical model

$$s = 0.29 - 0.16 \frac{h}{h_0} + 0.343 \frac{w}{b_0} - 0.048 \left( \frac{w}{b_0} \right)^2 \quad (6)$$

The actual height of the workpiece is written as

$$h = h_0 + u_y(t) \quad (7)$$

Substituting Eqs (4) (6) and (7) into Eq. (5), the elongation of the workpiece is

$$\Delta l = l_0 \left( \frac{h_0}{h_0 + u_y(t)} \right)^{0.71+0.16 \frac{h_0+u_y(t)}{h_0} - 0.343 \frac{w}{b_0} + 0.048 \left( \frac{w}{b_0} \right)^2} - l_0 \quad (8)$$

Due to the symmetry of the workpiece geometry, the horizontal displacement of the clamp is half of  $\Delta l$ . So the motion of the clamp in the two directions  $q$  could be determined by Eq. (3) and Eq. (8).

## 4 Reaction Forces of Resilient Cylinders

### 4.1 Inside Pressure of the Resilient Cylinder

The detail of the vertical resilient system is shown in Fig. 4(a). Assume that the fluid flows from the resilient cylinder to the air accumulator. The mass continuity for incompressible fluid states that the flow rate.

$$A_v \dot{l}_v = A_{ov} \dot{x}_{ov} \quad (9)$$

where  $A_v$  is the area of the resilient cylinder,  $A_{ov}$  is area of the orifice,  $\dot{x}_{ov}$  is the flow rate through orifice,  $\dot{l}_v$  is the velocity of the piston, which can be solved from Eq. (9). Neglecting the gravity and height difference, the Bernoulli's equation could be expressed as

$$\frac{p_v}{\rho} + \frac{\dot{l}_v^2}{2} = \frac{p_{AV}}{\rho} + \frac{\dot{x}_{OV}^2}{2} \quad (10)$$

where  $p_v$  is the pressure in the resilient cylinder,  $p_{AV}$  is pressure in the air accumulator.

Substituting Eq. (10) into Eq. (9), one can obtain

$$\dot{x}_{OV} = \sqrt{\frac{(p_v - p_{AV})A_v^2}{\rho(A_v^2 - A_{OV}^2)}} \quad (11)$$

In a realistic flow situation, there is a flow rate loss due to flow restrictions in the orifice. This loss is empirically quantified by a discharge coefficient  $C_{dv}$ , which represents the percentage of the ideal flow that actually occurs. This coefficient, when multiplied by the ideal flow, yields

$$Q_r = C_{dv} A_{OV} \dot{x}_{OV} \quad (12)$$

Substituting Eqs. (10)(11) into Eq. (12), we have

$$Q_r = C_{dv} A_{OV} \sqrt{\frac{(p_v - p_{AV})A_v^2}{\rho(A_v^2 - A_{OV}^2)}} \quad (13)$$

For the incompressible fluid, the continuity yields

$$C_{dl} A_{OV} \sqrt{\frac{(p_v - p_{AV})A_v^2}{\rho(A_v^2 - A_{OV}^2)}} = A_v \dot{l}_v \quad (14)$$

The pressure of the resilient cylinder could be solved by

$$p_L = \frac{\rho \dot{l}_v (A_v^2 - A_{OV}^2)}{2C_{dv}^2 A_{OV}^2} + p_{AV} \quad (15)$$

where  $p_{AV}$  is the air pressure in the air accumulator, which could be solved as follows:

Nitrogen pressure changes by polytrophic gas law for a closed system is expressed as

$$p_{AV} = p_{AV}^0 \left( \frac{x_{AV}^0}{x_{AV}} \right)^\gamma \quad (16)$$

where  $x_{AV}^0$  is the initial height of nitrogen,  $x_{AV}$  is the actual height of nitrogen.  $p_{AV}^0$  is the initial pressure of the air accumulator,  $\gamma$  is the polytrophic gas constant.

Substituting Eqs. (14)(15) into Eq. (16), we obtain

$$p_v = \frac{\rho \dot{l}_v^2 (A_v^2 - A_{OV}^2)}{2C_{dv}^2 A_{OV}^2} + p_{AV}^0 \left( \frac{x_{AV}^0}{x_{AV}} \right)^\gamma \quad (17)$$

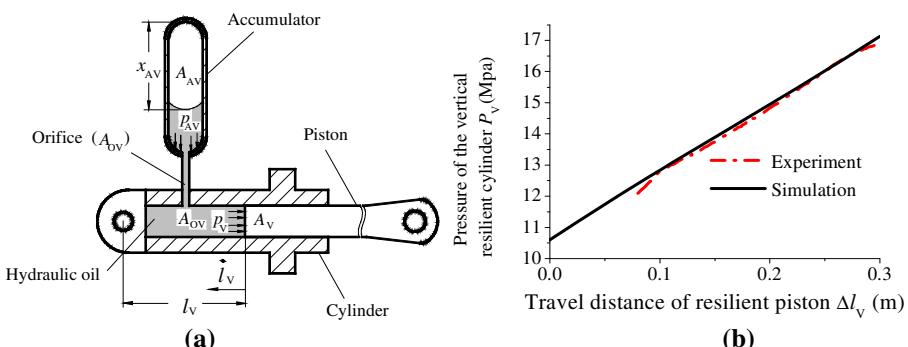
If the fluid flows from the air accumulator to the resilient cylinder, the pressure in the cylinder could be also solved by using the similar procedure. So the reaction forces could be solved by the given motion of the clamp, which has deduced in the previous section. Similarly, the reaction force of the horizontal cylinder could also be obtained because it has the same structure as the vertical cylinder.

## 4.2 Validation of the Pressure Equation

To verify the validation of the pressure equation, a forging manipulator of 20-ton manipulator was tested at the workshop. The resilient system parameters used for the experiment are listed in Table. 1. The orifice damping can be ignored in this experiment. In the vertical resilient cylinder, the liquid pressure and piston travel distance were measured by the displacement and pressure sensors, respectively. A good agreement could be seen in Fig. 4(b), which presents a linear stiffness character.

**Table 1.** Resilient system parameters of the forging manipulator for experiment

parameters	value
Initial height of vertical accumulator $x_{AV}^0$ (m)	0.933
Initial height of horizontal accumulator $x_{AL}^0$ (m)	0.733
Area of vertical accumulator $A_{AV}$ ( $\text{cm}^2$ )	753.4
Area of horizontal accumulator $A_{AL}$ ( $\text{cm}^2$ )	542.3
Area of vertical resilient cylinder $A_v$ ( $\text{cm}^2$ )	693
Area of horizontal resilient cylinder $A_L$ ( $\text{cm}^2$ )	628



**Fig. 4.** (a) Vertical resilient cylinder system of the manipulator. (b) Simulated and experimental pressure–distance curves of the vertical resilient cylinder of the 20-ton manipulator.

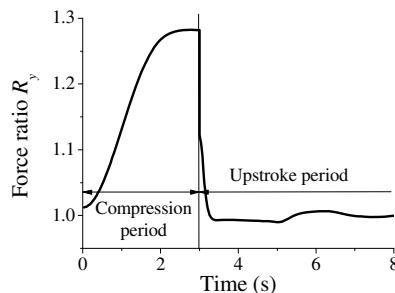
## 5 Dynamic Responses of the Forging Manipulator

The damping effects are taken into consideration of the simulation, with the area of vertical orifice  $A_{OV}$   $1.58 \text{ cm}^2$ , the area of horizontal orifice  $A_{OL}$   $1.72 \text{ cm}^2$ . The other

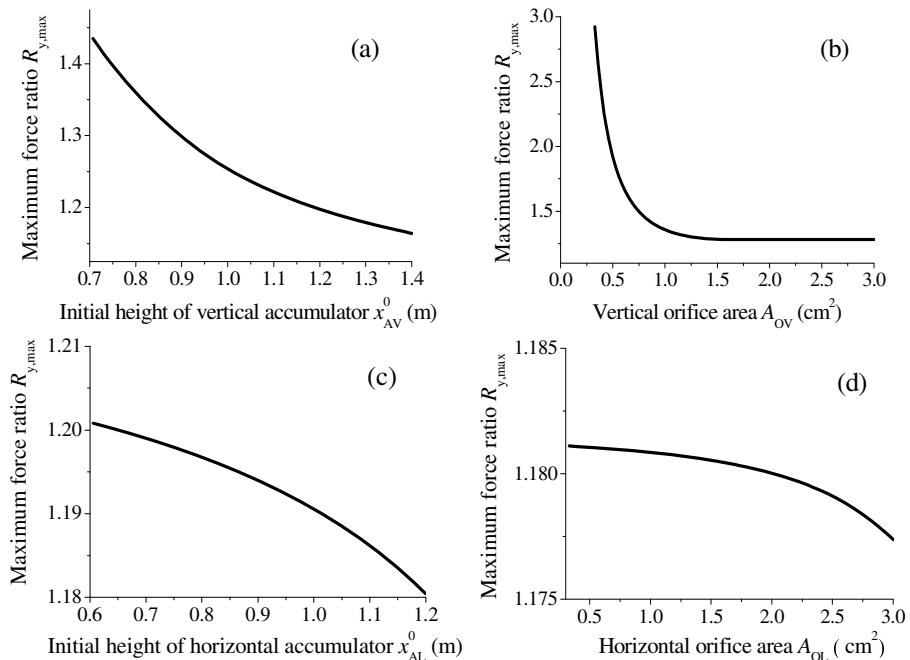
parameters are listed in Table 1. The decompression period is not considered because the manipulator always keeps static in this period.

### 5.1 Vertical Force Transmitted to the Clamp

The vertical force variation during the whole forging stroke is plotted in Fig. 5. It can be seen that vertical force increases with the increasing depth of penetration in the compression period, the maximum force ratio  $R_{y,\max}$  reaches to 1.29 at the end of shaping. Then the manipulator system goes through a free vibration after the upper die left, the dynamic force will reduce to the workpiece weight eventually.



**Fig. 5.** Time history of force ratio  $R_y$  during the whole forging stroke



**Fig. 6.** Influence of resilient parameters to  $R_{y,\max}$ . (a)  $x_{AV}^0$  . (b)  $A_{OV}$  . (c)  $x_{AL}^0$  . (d)  $A_{OL}$

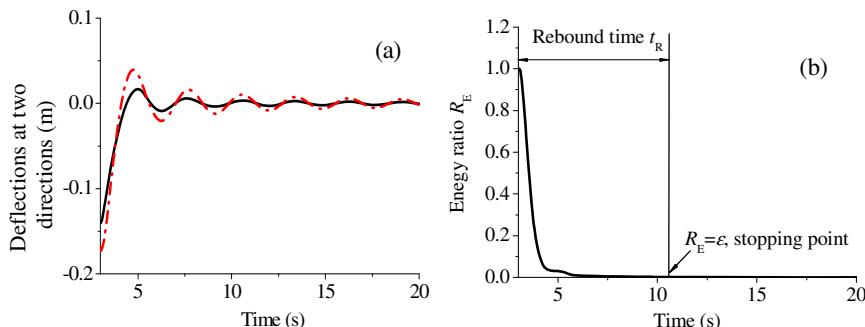
The effects of the vertical resilient parameters on the maximum force in the time domain are shown in Fig. 6(a) and 6(b). It can be seen that the vertical force transmitted to the clamp increases with the initial level. Small orifice produces higher vertical force when  $A_{OL} \leq 1.41\text{cm}^2$ , otherwise the damping effect of orifice can be neglected. The similar trends could be seen for the influence of the horizontal resilient parameters in Fig. 6(c) and 6(d), due to the coupling character of the manipulator mechanism. But the contribution is minor compared with that of the vertical resilient parameters.

## 5.2 Energy Dissipation of the Forging Manipulator during the Upstroke Period

The deflection responses of the clamp in the two directions during the upstroke period are plotted in Fig. 7(a). It can be seen that the deflections of clamp decrease fast at the two directions, eventually, and the clamp will stop vibrating over time because of the energy dissipation of orifices. The time point for stopping vibration is defined as that when the energy ratio is less than  $\varepsilon$  ( $\varepsilon=0.002$  in present analysis), which is written as

$$R_E = E_F / E_P \quad (18)$$

where  $E_P$  is the energy at the end of the metal shaping,  $E_F$  is the actual energy during the forging operation. It can be seen the energy ratio of the system decreases monotonously after the upper die left. The duration from the end of the metal shaping to the vibration stopping point is defined as the rebound time  $t_R$ , which is regarded as the index of the working efficiency of the forging manipulator.



**Fig. 7.** (a) Displacement responses of the clamp during the upstroke period. (b) Variation of system energy during the upstroke period.

The analytical relationship between rebound time  $t_R$  and four design variables  $X=[x_1 \ x_2 \ x_3 \ x_4] = [A_{OV} \ x_{AV}^O \ A_{OL} \ x_{AL}^O]$  is approximated following the factorial design of experiments and response surface methodology. A  $5^4$  full factorial design was used in order to get the output data uniformly distributed all over the ranges of the input parameters. In this way 625 experiments were carried out with different combinations of the levels of the input parameters. The full second-order model in four variables from the DOE analysis is obtained:

$$t_R = 0.71 - 8.92x_1 + 17.49x_2 - 1.92x_3 + 0.7x_4 + 1.42x_1^2 - 7.9x_2^2 + 2.35x_3^2 + 0.18x_4^2 + 2.78x_1x_2 + 2.6x_1x_3 + 0.13x_1x_4 - 3.21x_2x_3 - 0.14x_2x_4 - 0.35x_3x_4 \quad (19)$$

It can be noted the horizontal parameters have more influence on the rebound time than that on the maximum force transmitted to the clamp.

## 6 Performance Optimization of the Forging Manipulator

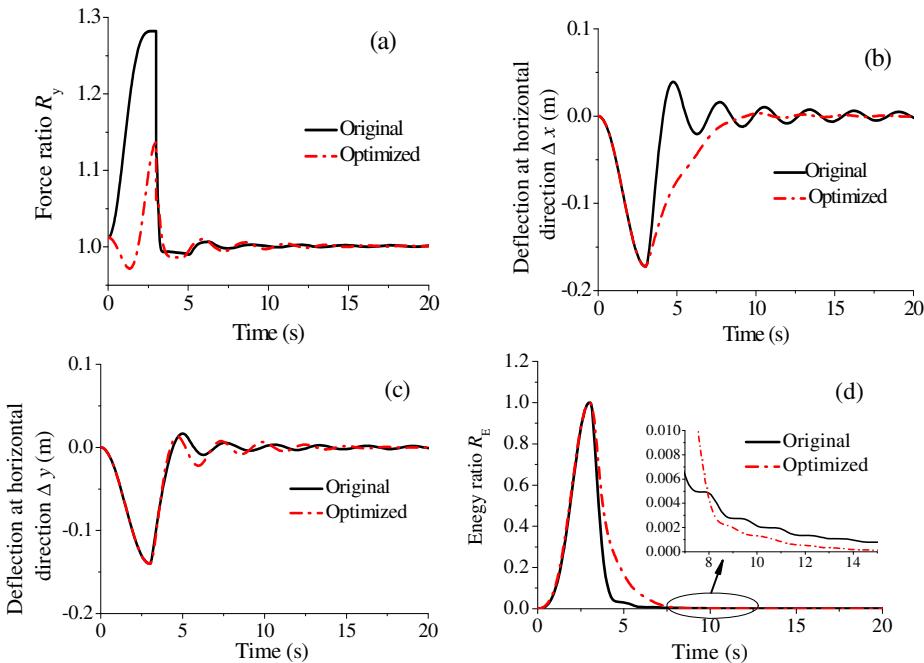
All the design parameters play a conflicting role for the manipulator performance. A parameter optimization is executed to reduce the maximum force exerted to the manipulator. The design parameters should also satisfy the working efficiency constraint ( $t_R \leq 10.2$  s) as well as some physical constraints (e.g.  $x_{iL} \leq x_i \leq x_{iU}$  for  $i=1,2,3,4$ , where  $i$  is the number of design parameter,  $x_{iL}$  and  $x_{iU}$  are lower bound and upper bound of each parameter, respectively). Adaptive simulated annealing algorithm (ASA) [11] is used in this work to solve the formulated resilient system optimization.

The comparison of the optimized design and original design are listed in Table. 2. It can be seen that the optimized design presents a lower stiffness and damping of the vertical cylinder, and a higher stiffness and damping of the horizontal cylinder.

The variations of vertical force for both designs during the whole forging stroke are plotted in Fig. 8(a). It can be noted that the maximum force ratio  $R_{y,max}$  of the optimized design is 1.16, which is 91.3% of the maximum force ratio  $R_{y,max}$  of the original design 1.27. Without considering the dead load of the workpiece weight, the actual extra force induced by forging motion is reduced by 40.7%. A significant improvement is obtained. The vertical force decreases with the increasing penetration depth at the beginning of forging because of the softening of the vertical resilient cylinder. The displacement responses of the clamp in the two directions are plotted in Fig. 8(b) and 8(c). It can be seen that optimized design presents increasing vibration amplitude at the vertical direction and decreasing vibration amplitude at the horizontal direction, which means more energy dissipation transferred from the vertical cylinder to the horizontal cylinder. The total energy ratio variation is plotted in Fig. 8(d). It can be observed that the energy dissipation of the optimized design is faster than the original design when  $t \geq 7.6$  s, although its energy dissipation is much lower than the original design at beginning. As a result, the rebound time of optimized design is reduced to 8.4s, which is 82.4% of that of the original design 10.2s, as shown in Fig. 8(d).

**Table 2.** A comparison of the originally designed and optimized resilient systems

	$A_{OV}$ (cm <sup>2</sup> )	$x_{AV}^O$ (m)	$A_{OL}$ (cm <sup>2</sup> )	$x_{AL}^O$ (m)
Original design	1.58	0.933	1.721	0.733
Optimized design	2.52	1.4	0.303	0.632



**Fig. 8.** (a) Vertical force transmitted to the clamp. (b) Displacement response at the horizontal direction. (c) Displacement response at the vertical direction. (d) Variation of system energy.

## 7 Conclusions

Forging press produces the high force on the clamp of the forging manipulator at the vertical direction. Reducing the force transmitted to the clamp is a major concern for this widely used auxiliary metal forming equipment. At the same time, the manipulator needs to cease to the original position as soon as possible to guarantee the working efficiency. In this paper, dynamic behaviour of the forging manipulator is investigated. The focuses are on the vertical force exerted to the clamp and the working efficiency. The resilient parameters may play the cruel roles for the manipulator performance. A sensitivity study on the vertical force with respect to the given resilient parameters is carried out. Results show that the vertical force increases with the decreasing of the initial height of accumulator as well as the orifice area. A parameter optimization is executed to reduce the vertical force with the constraint of the rebound time of the clamp. The dependence of rebound time on the resilient parameters is approximated by design of experiments based the response surface model. The force is reduced by 8.7% based on the formulated design optimization which demonstrates acceptable force reducing capability.

**Acknowledgments.** This study is supported by the National Basic Research Program of China (No.: 2006CB705403).

## References

1. Lilly, K.W., Melligeri, A.S.: Dynamic simulation and neural network compliance control of an intelligent forging center. *J. Intell. Robot. Syst. Theory & Applications* 17, 81–99 (1996)
2. Nye, T.J., Elbadan, A.M.: Real-time process characterization of open die forging for adaptive control. *J. Eng. Mat. and Tech.* 123, 511–516 (2001)
3. Westermeyer, W.: Process and apparatus for operating a press unit. US Patent 5454262 (1995)
4. Kutz, H.: Workpiece manipulator for use in a forging press. US Patent 4031736 (1977)
5. Kost, E., Folta, W.: Gripper feed and Gripper resilience cylinders on forging manipulators. US Patent 4420287 (1983)
6. Eberhard, B.: Forging manipulator. Germany Patent DE10128451(A1) (2002)
7. Wang, W.R., Zhao, K., Lin, Z.Q., et al.: Evaluating interactions between the heavy forging process and the assisting manipulator combining FEM simulation and kinematics analysis. *Int. J. Advanced Mat. and Tech.* (2009), doi:10.1007/s00170-009-2285-3
8. Liu, D.S., Li, G.: Dynamic behavior of the forging manipulator under the press motions. In: 2nd International Conference on Advanced Design and Manufacturing, pp. 417–420. Trans Tech Publications Ltd., Switzerland (2010)
9. Chen, W., Zhang, X.X., Cui, Z.S.: Numerical Simulations of Open-Die Forging Process for Manipulator Design. In: Xiong, C.H., Liu, H.H., Huang, Y.G., Xiong, Y.L. (eds.) ICIRA 2008. LNCS (LNAI), vol. 5315, pp. 650–658. Springer, Heidelberg (2008)
10. Tomlinson, A., Stringer, J.D.: Spread and elongation in flat tool forging. *J. of Iron Steel Institute* 193, 157–162 (1959)
11. Ingber, L.: Very fast simulated re-annealing. *Math. Comput. Model.* 12(8), 967–973 (1989)

# Structural Optimization for Wall Frame Design of a Forging Manipulator

Kai Zhang<sup>\*</sup>, Gang Li, and Zhan Kang

Department of Engineering Mechanics

State Key Laboratory of Structural Analysis of Industrial Equipment,

Dalian University of Technology, Dalian, 116024, China

[zhang\\_kai@mail.dlut.edu.cn](mailto:zhang_kai@mail.dlut.edu.cn), [ligang@dlut.edu.cn](mailto:ligang@dlut.edu.cn)

**Abstract.** The optimization designs are performed to the structural wall frame of the forging manipulator. The stress conditions of the wall frame are obtained based on the results of the dynamic analysis of the forging manipulator under the multiple load cases. The topology optimization and sizing optimization are carried out to the classic types of heavy-load manipulators. The global stiffness of the wall frame is maximized subject to the volume constraint for the topology optimization; and the sizing optimizations, in which the objective is the minimizing the maximum displacements with the different design variables, is carried out under the mass constraint and the maximum Mises stress constraints. Results show the improvement of the performance of the wall frame in the forging manipulator.

**Keywords:** Forging manipulator, wall frame, topology optimization, sizing optimization.

## 1 Introduction

There exist three different aspects of the structural design problem, which are the sizing, shape, and topology optimization. With the purpose of the optimal material layout of load-bearing structural components, topology optimization has been focused on both in theoretical research and practical applications for the past two decades. Topology optimization based on the Finite Element Method (FEM) is usually served as a rational method for the innovative structural designs due to its conceptual design characteristic. Generally speaking, continuum topology optimization problems can be stated as to minimize the structural weight (or material volume) or to optimize the structural performance (compliance, fundamental frequency, etc.) under the material volume constraint. The topology optimization methods of the continuum structure can be classified into two categories: [1,2] macroscopic approaches and microscopic approaches. The representative method of the former are the bubble method and the recently developed level set method[3]; the latter mainly includes the varied thickness

---

\* Corresponding author.

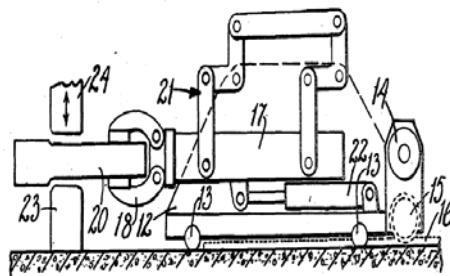
method, the homogenization method[4], the evolutionary structural optimization method[5]and the SIMP method which introduces the artificial material model[6]. The SIMP method is popular and effective, in which the relation between the elastic stiffness and the relative density of the material is assumed to obey a power law, so that the intermediate density can be suppressed. It has been investigated by Bendsoe and Sigmund[7] that this assumption can be physically meaningful and further strengthens the theoretical background of the SIMP method. Some numerical techniques such as the filter techniques [8,9], the perimeter control method[10,11] have been developed to deal with the mesh dependency and checkerboard problems.

In applications regarding component design of aerospace and aeronautical structures, topology optimization has been employed for improving the structural and multi-disciplinary behaviors such as the overall stiffness, the aerodynamic performance and the fundamental frequencies under the specified structural weight constraint. Krog and Tuck[12] applied the topology optimization techniques to the optimal design of aircraft wing-box ribs. Maute and Allen[13] presented a material density-based approach for the optimal placement of stiffeners in a flexible wing subject to aerodynamic forces. The material distribution formulation has also been used by Maute and Reich[14]in the optimal design of actuator locations and the inner structures of adaptive wings. Afonso et al. [15] combined the topology optimization and the shape optimization for the design of stiffened shell structures, which have been widely used in the aerospace industry. Kang et al. [16] achieved the topology optimization results of the space vehicle structures considering attitude control effort. Chen et al. [17] applied topology optimization techniques in the layout design of stiffeners in a storage tank of a space craft for minimum weight under stiffness and strength constraints.

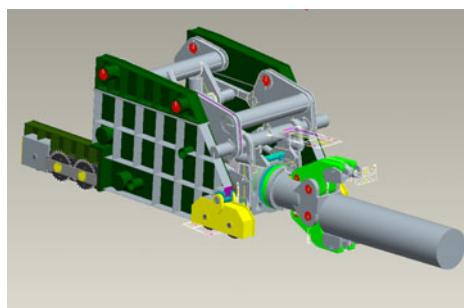
A forging manipulator is a vehicle for supporting and manipulating a workpiece during a forging operation on the workpiece at a specified speed. As the heavy open forging press is upgraded in speed and precision, the automation of forging process with the forging manipulator becomes necessary to improve efficiency and safety, and to save cost and natural resources. The forging process is in the inhospitable and vibrating environment. Most researches on the vibration problems in forging industry were focused on the dynamic behavior of the foundation, workshop and forging press. However, there were few researches on the dynamic analysis and structural optimization of the forging manipulator. In order to improve the global stiffness of the wall frame of the forging manipulator, the topology and sizing optimization for the wall frame design is focused on in this paper. In the topology optimization the optimal material distribution could instruct the locations of the ribs between the wall frames compared to the regular rib locations of the original wall frames. The sizing optimization [18] always minimizes (or maximizes) a physical index such as the mean compliance (external work), peak stress, structural response, subject to the required constraints, which could determine the sizes of the ribs to minimize the max deflection of the structure under the constraints from the stress and mass of the side wall frame.

## 2 FEM Model of the Forging Manipulator

A forging manipulator is a vehicle of a rail-bound carriage, carrying a peel which has a pair of jaws for gripping the workpiece. The peel can be moved horizontally, vertically and rotationally [19]. Fig. 1. is a schematic drawing showing the forging manipulator. The carriage 12 of the manipulator is mounted on wheel 13 running on rails; the carriage carries a hydraulic drive motor 14 driving a pinion 15 meshing with a rack 16 locates on the floor. The carriage 12 also supports a peel 17 having a pair of jaws 18 which grip an end of the workpiece indicated at 20. The peel is supported by a linkage 21 and can be driven in the longitudinal direction, i.e., parallel to the carriage movement on the rails, by actuating means in the form of a piston and cylinder unit 22, of which the cylinder is shown as secured to the carriage and the piston to the peel. The wall frame is a part of the carriage in the manipulator. The wall frame in the carriage 12 is the main supporting component in the manipulator, whose stiffness and strength will be improved in the optimization design. Fig. 2. is the CAE model for the forging manipulator, and Fig. 3. shows the true wall frame of a forging manipulator. The FEM model of the wall frame with the regular rib distribution is shown in Fig. 4.



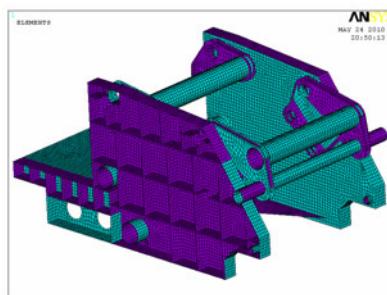
**Fig. 1.** Schematic drawing of the manipulator (adopted from [19])



**Fig. 2.** CAE model of the manipulator



**Fig. 3.** Wall frame of the manipulator

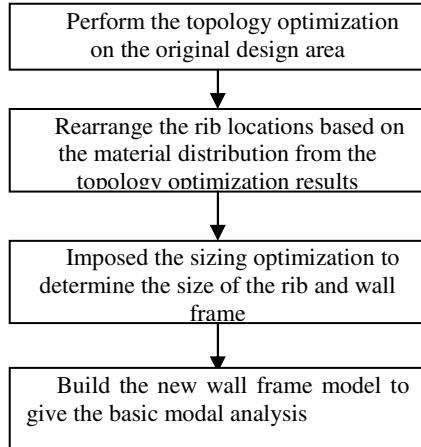


**Fig. 4.** FEM model of the wall frame with ribs

The wall frame of the forging manipulator is the key structural component, which is the main supporting part of the forging manipulator. The uplifting motion of this mechanism is the common and relatively dangerous load case to the wall frame, in which the forging manipulator grips the hot workpiece from a lower position to a higher position. The up lifting hydraulic cylinders produce a high press to the connecting hinges on the wall frame due to the acceleration variation of the movement. So the stiffness of the wall has to be considered to resist the pressure under this load case, and the maximum displacement is used as the objective function in the optimization design with the constraints of the stress and mass.

### 3 Structural Optimization of the Wall Frame

The proposed structural optimization for the wall frame includes two parts, topology optimization and sizing optimization. The general flowchart of the optimization procedure is shown in Fig. 5.:



**Fig. 5.** Flowchart of the optimization

### 3.1 Topology Optimization of the Wall Frame

Conventionally, topology optimization of continuum structures in the multiload cases is formulated as a minimum mean compliance problem under a material volume constraint. In the SIMP formulation, an artificial isotropic material model is introduced and element-wise material density variables are used for describing the structural layout. To suppress the intermediate densities in the resulting optimum design, a penalized interpolation scheme between Young's modulus E and the material density x is assumed as Eq.(1)

$$E = x^0 E_0 \quad (1)$$

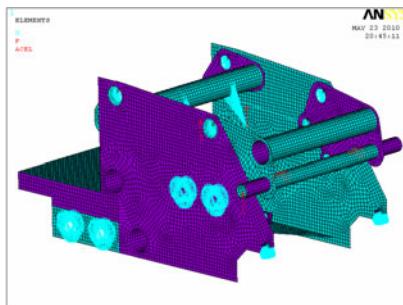
where  $p > 1$  is the penalty factor,  $E_0$  is the Young's modulus of the fully solid material, which has a unit density.

In this study, the topology optimization problem with a volume constraint is expressed as Eq. (2):

$$\begin{aligned} \min_x \quad & C(\mathbf{X}) = \sum_i^g \frac{1}{g} c_i(\mathbf{X}) = \sum_i^g \frac{1}{g} \mathbf{U}_i^T \mathbf{K}_i \mathbf{U}_i = \sum_i^g \sum_{e=1}^N \frac{1}{g} ((x_e)^p \mathbf{u}_e^T \mathbf{k}_e^0 \mathbf{u}_e)_i \\ \text{subject to} \quad & \frac{V(x)}{V_0} = f_v \\ & 0 < x_{\min} \leq x_e \leq x_{\max} = 1, e = 1, 2, \dots, N \end{aligned} \quad (2)$$

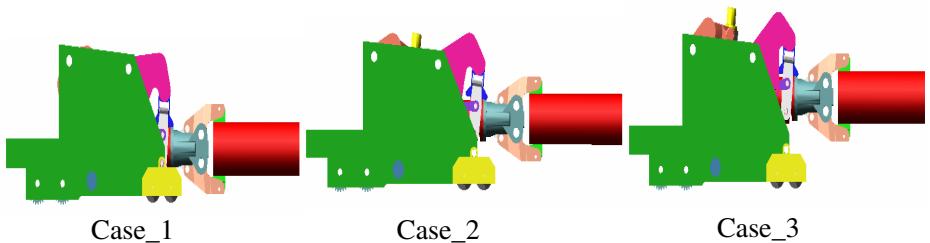
where  $c$  represents the strain energy,  $g$  means the number of the load cases,  $x_e$  denotes the density of the  $e$ -th element,  $\mathbf{x}$  is the vector of elemental design variables,  $\mathbf{K}$  is the global stiffness matrix,  $\mathbf{U}$  is the vector of nodal displacement,  $\mathbf{u}_e$  denotes the elemental displacement vector,  $\mathbf{k}_e^0$  is the stiffness matrix of the  $e$ -th element when it is made of the solid material,  $N$  is the total number of elements,  $V$  is the material volume,  $V_0$  is the design domain volume,  $f_v$  is the prescribed volume fraction. In order to avoid the numerical difficulties caused by non-positive definite stiffness matrix, a lower-bound limit  $x_{\min} > 0$  is imposed on the material density.

In the topology optimization, the wall frame is simplified as a 3-dimensional area which is showed in Fig. 6., whose stress boundary conditions can be obtained by the dynamic analysis with ADAMS of the whole manipulator, and the displacement boundary is assumed as fixed on the place between the wall frame and the wheel running on rails. In the topology optimization the 3-dimensional model is used for analysis, and the 2-dimensional area of the frame wall is taken as the design zone for topology optimization. The shell element is selected to simulate the wall frame. The Sequential Quadratic Programming is employed to maximize the global stiffness under the constraint of the 30% volume of initial value.



**Fig. 6.** The boundary conditions of the wall frame

Three load cases of the stress conditions obtained from the ADAMS means three positions of the lifting arms of the manipulator which can be seen from Fig. 7. and Fig. 8. showing where the revolute Joint.33 and Joint.61 are.



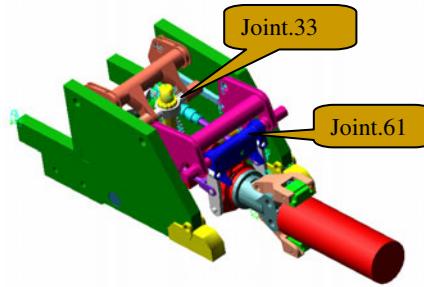
**Fig. 7.** Three load cases positions of the manipulator when lifting the workpiece

**Table 1.** The information of the three load cases of the Joint.33 and Joint.61

Unit(N)	Joint.33_X	Joint.33_Y	Joint.33_Z	Joint.61_X	Joint.61_Y	Joint.61_Z
Case_1	9.8204E4	1.6537E6	117	-4E5	-3.85E6	-4.8504E4
Case_2	1.6924E5	1.2656E6	152	-4.7785E5	-2.9874E6	-6.3149E4
Case_3	1.16E5	1.013E6	133	-2.80303E5	-2.4057E6	-3.0469E4

### 3.2 Sizing Optimization of the Wall Frame

In this paper, the basic dynamic performance based sizing optimization is considered as followed Eq. (3):

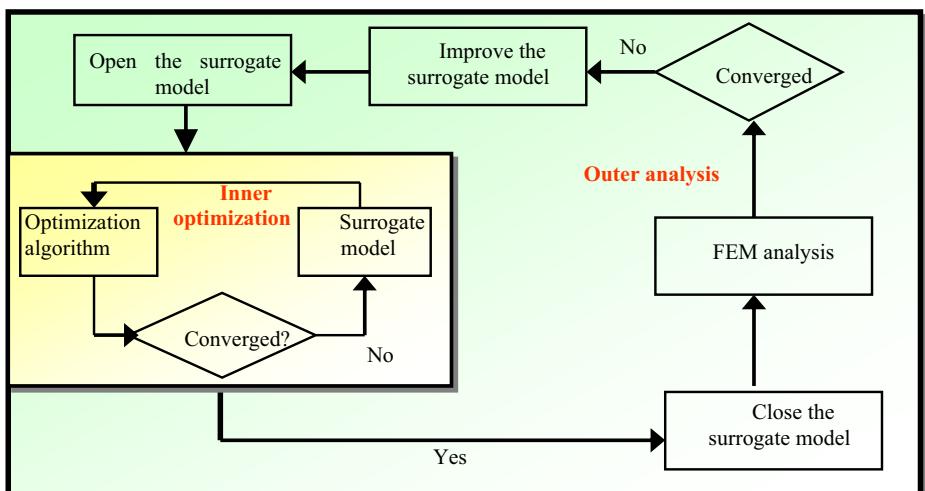


**Fig. 8.** The positions of revolute joint.33 and revolute joint.61

$$\begin{aligned} & \max_x d_{\max} \\ & \text{subject to } 0 < M(\mathbf{X}) \leq M_0 \\ & \quad 0 < s_{\max} \leq s_{\max_0} \end{aligned} \quad (3)$$

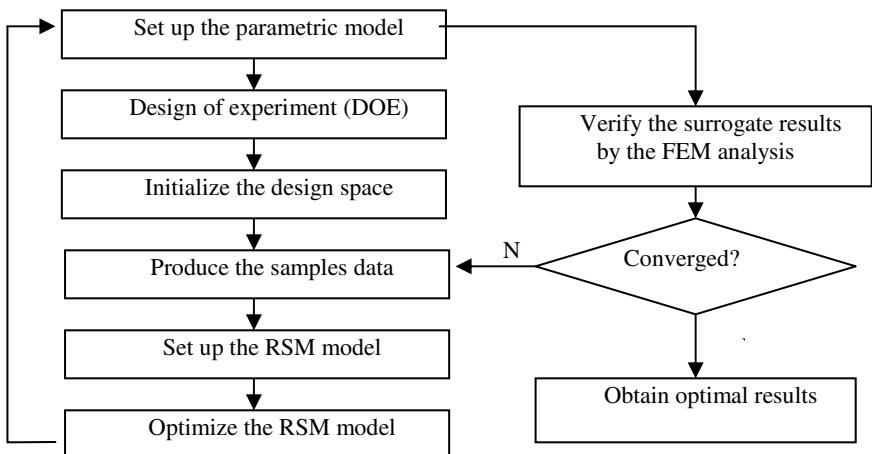
where  $d_{\max}$  is the maximum displacements of the wall frame,  $\mathbf{X}$  is the vector of the design variables of the sizes of the rib and the thickness of the wall frame,  $M(\mathbf{X})$  is the mass of the wall frame and the  $M_0$  is the mass of the original design,  $s_{\max}$  is the maximum Mises stress of the wall frame and  $s_{\max_0}$  is the maximum stress of the original design.

Considering the expensive computational efforts of 3D FEM analysis, the double-level optimization technique has been adapted, which includes the inner optimization based on the design of experiment (DOE) and the surrogate model (such as response surface method, RSM), and the outer FEM analysis that can update the results from the inner optimization. The flowchart is as follows



**Fig. 9.** The flowchart of the double-level optimization

The double-level sizing optimization contains the following steps: 1) Obtain the design space by the design of experiment (DOE) with the optimized Latin Hypercube Sampling method; 2) Set up the second-order RSM surrogate model according to the above samples; 3) Apply the Hooke-Jeeves algorithm to find the local optimization, which could be verified and updated by the outer FEM analysis. The calculation is converged if the relative error between the results from the surrogate and the ones from the FEM analysis is satisfied, otherwise the data from the surrogate is updated by the FEM analysis and another inner optimization search is performed for the surrogate model. Consequently, the auto sizing optimization flowchart can be depicted as followed:

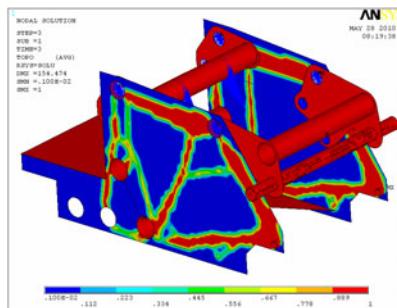


**Fig. 10.** The flowchart of the auto sizing optimization

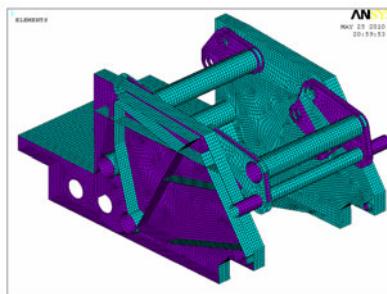
## 4 Optimal Results

### 4.1 Results of Topology Optimization

The result of the topology optimization in the multiload cases is showed as Fig. 11., based on which the new 3D wall frame with the optimal rib distribution is built up, considering the convenience of manufacture, shown in Fig. 12.. In 3D FEM model with the regular and optimal ribs distribution, more details of the forging manipulator were considered such as the hydraulic station, cylinders between the walls, the triangles connected to the wall etc., which were simulated by the shell elements with different thickness.



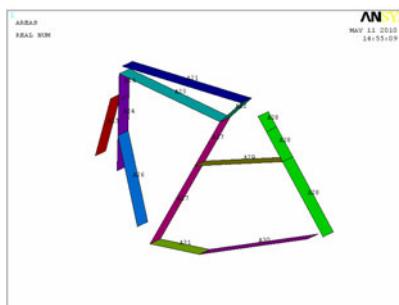
**Fig. 11.** Result of topology optimization



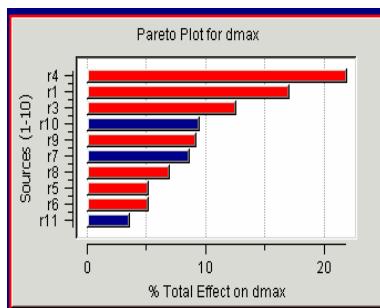
**Fig. 12.** FEM wall frame with the new ribs location

#### 4.2 Results of Sizing Optimization

In order to carry out the sizing optimization of the wall frame, the initial value of the design variables is used as 16 mm for the sizes of the ribs which are used as the  $r_1, r_2 \dots r_{11}$  and showed in the Fig. 13. So the optimized Latin hypercube sampling method is performed for the DOE, and the Pareto solution is obtain as Fig. 14. from which we can get that the  $r_4$  contributes the response  $d_{max}$  most.



**Fig. 13.** The design variables of ribs

**Fig. 14.** The pareto solution of the DOE

The optimal results are obtained after the double-level optimization, listed in Table 2. The comparison between the original model and the optimized model among the three load cases are depicted in.

Table 3. which compares the maximum Mises stress and the maximum displacement of the wall frame to show the improvement of the wall frame structure of the manipulator. Stress\_C1 and Disp\_C1 stand for the maximum Mises stress and the maximum displacement in load case 1, respectively.

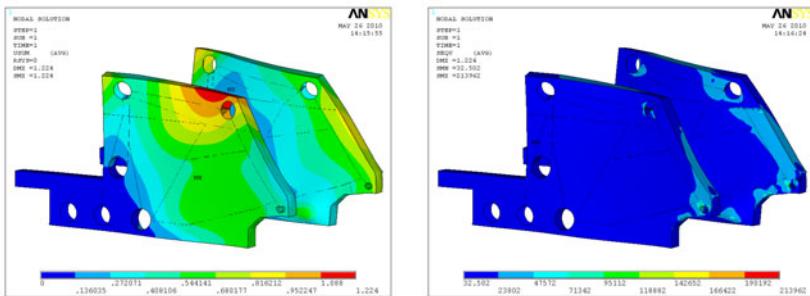
**Table 2.** The optimal design results

Design Variables	Type	Lower Bound	Current Value	Upper Bound
r1	REAL	0.01	25.25	100.0
r2	REAL	0.01	11.72	100.0
r3	REAL	0.01	28.26	100.0
r4	REAL	0.01	31.25	100.0
r5	REAL	0.01	23.68	100.0
r6	REAL	0.01	23.25	100.0
r7	REAL	0.01	6.75	100.0
r8	REAL	0.01	24.84	100.0
r9	REAL	0.01	26.10	100.0
r10	REAL	0.01	6.75	100.0
r11	REAL	0.01	9.38	100.0

**Table 3.** Comparison between the initial and optimal design among the three load cases

Models	Original model	Topology optimized model	Percentage improved (%)
Ribs Mass(Kg)	5073.6	3606.2	28.92
Stress_C1(Mpa)	264.644	213.962	19.15
Disp_C1(mm)	1.284	1.224	4.67
Stress_C2(Mpa)	203.893	173.498	14.91
Disp_C2(mm)	1.199	1.161	3.1
Stress_C3(Mpa)	173.774	140.682	19.04
Disp_C3(mm)	0.825507	0.786861	4.68

The results of the optimized design of the wall frame of the first load case in 3D can be shown in Fig. 15. General speaking, the mass of the ribs is reduced from 5073.6Kg to 3606.2Kg by 28.92%. For all the load cases, the maximum Mises stresses are reduced to 213.962Mpa, 173.498Mpa and 140.682Mpa with the improvement of 19.15%, 14.91% and 19.04%, which are lower than the strength limit 235Mpa of the material. Especially for the load case 1, the maximum Mises stress of the original design exceeds the material strength limit 235Mpa while the optimized result is obtained in a normal stress level. The maximum displacements in these load cases are also improved.



**Fig. 15.** Results of the displacement and the Mises stress of the optimized design in load cases 1

## 5 Conclusions

In this paper, the topology optimization and the sizing optimization are performed to improve the global stiffness of the wall frame of the forging manipulator, to obtain the conceptual design of the wall frame by 2D topology optimization for the multiload cases and to determine the design variables by the sizing optimization to minimize the maximum displacement of the wall frame structure of the manipulator. The results showed the improvement of the optimal design, with the decrease of the ribs mass of the wall frame from 5073.6Kg to 3606.2 Kg (28.92 % lower) and the maximum Mises stress of the structure are reduced from 264.664Mpa, 203.893Mpa and 173.774Mpa to 213.962Mpa(19.15% lower), 173.498Mpa(14.91) and 140.682Mpa(19.04% lower) under three load cases, respectively.

## Acknowledgment

The support of the National Basic Research Program of China (No.: 2006CB705403) is much appreciated.

## References

1. Rozvany, G.: Aims, scope, method, history and unified terminology of computer- aided topology optimization in structural mechanics. Structural Optimization 21(2), 90–108 (2001)
2. Eschenauer, H.A., Olhoff, N.: Topology optimization of continuum structures: a review. Applied Mechanics Reviews 54(4), 331–389 (2001)

3. Wang, M.Y., Wang, X., Guo, D.: A level set method for structural topology optimization. *Computer Methods in Applied Mechanics and Engineering* 192, 227–246 (2003)
4. Bendsøe, M.P., Kikuchi, N.: Generating optimal topologies in structural design using a homogenization method. *Computer Methods in Applied Mechanics and Engineering* 71, 197–224 (1988)
5. Xie, Y.M., Steven, G.P.: *Evolutionary Structural Optimization*. Springer, Berlin (1997)
6. Bendsoe, M.P.: *Optimization of Structural Topology, Shape and Material*. Springer, Berlin (1997)
7. Bendsoe, M.P., Sigmund, O.: Material interpolation schemes in topology optimization. *Archives of Applied Mechanics* 69(9-10), 635–654 (1999)
8. Sigmund, O., Petersson, J.: Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima. *Structural Optimization* 16, 68–75 (1998)
9. Bourdin, B.: Filters in topology optimization. *International Journal for Numerical Methods in Engineering* 50, 2143–2158 (2001)
10. Haber, R.B., Jog, C.S., Bendsoe, M.P.: A new approach to variable-topology shape design using a constraint on parameter. *Structural Optimization* 11(1), 1–12 (1996)
11. Fernandes, P., Guedes, J.M., Rodrigues, H.: Topology optimization of three-dimensional linear elastic structures with a constraint on “perimeter”. *Computers and Structural* 73, 583–594 (1999)
12. Krog, L.A., Tuck, K.: Topology optimization of aircraft wing box ribs. In: AIAA, pp. 2004–4481 (2004)
13. Maute, K., Allen, M.: Conceptual design of aeroelastic structures by topology optimization. *Structural and Multidisciplinary Optimization* 27(1-2), 27–42 (2004)
14. Maute, K., Reich, G.W.: Integrated multidisciplinary topology optimization approach to adaptive wing design. *Journal of Aircraft* 43(1), 253–263 (2006)
15. Afonso, S.M.B., Sienz, J., Belblidia, F.: Structural optimization strategies for simple and integrally stiffened plates and shells. *Engineering Computations* 22(4), 429–452 (2005)
16. Kang, Z., Wang, X.M., Wang, R.: Topology optimization of space vehicle structures considering attitude control effort. *Finite Elements in Analysis and Design* 45, 431–438 (2009)
17. Chen, B.S., Kang, G.J., Li, Y.P.: Design optimization of stiffened storage tank for spacecraft. *Structural and Multidisciplinary Optimization* 36(1), 83–92 (2008)
18. Bendsoe, M.P., Sigmund, O.: *Topology Optimization: theory, methods and applications*, Berlin, Heidelberg, New York (2003)
19. Harrison: Forgin manipulator, United States Patent, No.3696651,10.10 (1972)

# Gripping Force Model and Influence Factors of Gripping Device for Heavy Forging Manipulators

Lixin Yang\*, Hua Deng, and Zhen Xu

School of Mechanical and Electrical Engineering, Central South University, Changsha,  
410083, China  
yanglixinsailing@sina.com

**Abstract.** A mathematic model of point contact friction force between gripper jaw and workpiece was established based on the wrench theory and the principle of force equivalence. The eccentric of workpiece was studied and its effect on the gripping force of the jaw was analyzed. This model revealed the factors influencing the variation of gripping force under heavy load conditions and provided a theoretical support for studying the load-carrying capability of stable gripping.

**Keywords:** Forging manipulator; Wrench theory; Eccentric of workpiece; Gripping force; Gripper.

## 1 Introduction

Gripping devices play a very important role in large scale heavy-load manipulators [1]. The structure and driving system of gripping devices are designed primarily based on the gripping forces holding the workpiece. To grip the forging workpiece stably, the gripping forces should be designed properly. At present, several models have already been established to study the gripping forces. The mechanical model between the driving force and the tong force considers the effect of friction moment, inertia force and radius of workpiece when the jaws are located on the vertical and horizontal positions [2]. The point contact model considers friction of jaw [3, 4]. Because the workpiece gripped by the gripping device of heavy-load manipulator is very large, the eccentric of the workpiece is unavoidable. However, the influence of workpiece eccentric is not considered in current models. Based on the wrench theory and the principle of force equivalence, a mathematic model of the point contact friction force between the jaw and workpiece is established, which considers the eccentric effect of workpiece on the gripping device of forging manipulators. The model will provide a theoretical support for studying the load-carrying capability of stable gripping and revealing its rules under the heavy load conditions.

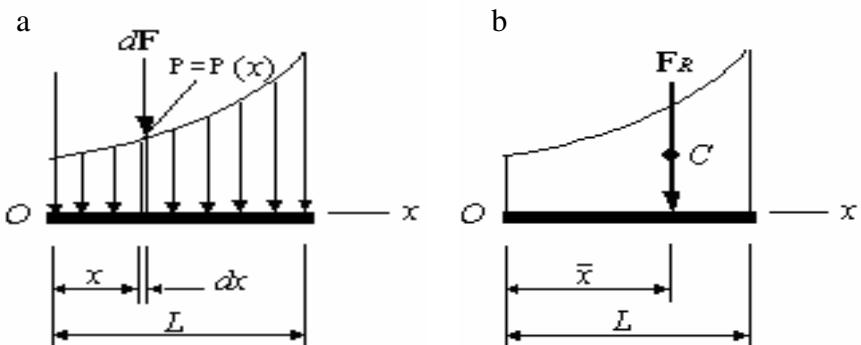
---

\* Master student, research area is modeling and simulation of mechanical and electrical systems.

## 2 Modeling of Gripping Force

### 2.1 Simplification of Contact Force

Take round shaft forging as an example. Assume the jaw and workpiece are rigid contact and there is no deformation in the contact area. The contact between V-shaped jaw and workpiece is line contact and the contact lines are the tangents between the planes of the upper or lower jaws and the cylindrical surface of the workpiece. The contact force distribution along the longitudinal axis  $x$  of the jaw is schematically shown in Figure 1a. The contact force applied on the contact line  $L$  can be considered as the total force of infinite number of parallel forces applied on infinitesimal lengths. Assume that the distributed force is a function  $p(x)$  of the jaw length  $x$  and this distributed force system can be equivalent to a concentrated force  $F_R$  applied at location  $\bar{x}$  as shown in Figure 1b.



**Fig. 1.** Distribution (a) and simplification (b) of contact forces along the longitudinal axis of jaw

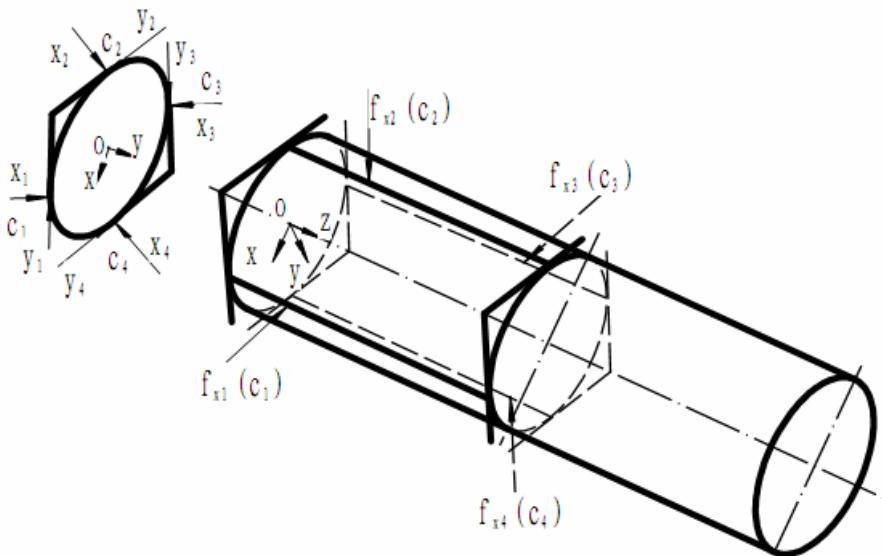
The equivalent law of the force and moment gives:

(1) The total force  $F_R$  is equal to the sum of distributed forces,  $F_R = \int_L p(x)dx$ ;

(2) The moment  $M_{RO}$  about the origin  $O$  of total force  $F_R$  is also equal to the moments of distributed forces  $M_O$ ,

$$M_{RO} = \sum M_O \Rightarrow \bar{x}F_R = \int_L xp(x)dx \Rightarrow \bar{x} = \frac{\int_L xp(x)dx}{\int_L p(x)dx}$$

When the distribution of contact force is unknown, it can be equivalently replaced by a concentrated force applied on a certain point of the contact line in order to study the relations among the external force, contact force, and joint driving force. Therefore, the gripping contact model can be simplified to a point contact model.



**Fig. 2.** Coordinate system and model of contact force

## 2.2 Coordinate System

The gripping equilibrium equations describe the balance among the applied force on the forge piece, the contact force between the forge piece and the jaw, and the driving force of the gripping cylinder. An effective method, gripping matrix is used to describe this relation. That is, the contact between the jaw and the forge piece is considered as a mapping between the force on the contact point applied by the jaw and the total force screw on a certain reference point of the forge piece. Therefore, the coordination system should be determined first to analyze the contact forces. As shown in Figure 2, drawing a line to connect the two pins of the V-shaped jaws, the line intersects with the forge piece and forms an intersection plane, take its center as the forge piece coordinate origin  $o$ . A series of coordinate systems  $C_i$  fixed at each contact point as the reference coordinate of the contact force and the forge piece are established. For the sake of convenience, set the  $x$  axis direction along the normal direction of forge piece surface at the contact point when choosing the contact coordinate system, and then define the contact location relative to the forge piece coordinate origin  $o$ .

## 2.3 Force Balance Equation

When the grip head of the large scale heavy load forge manipulator is in the vertical state, the reaction force applied on the jaw causes the jaw to rotate about the pin axis because the jaw bears higher gripping moment, then the forge piece drops and rotates certain angle accordingly. This reaction force leading to the rotation of jaw should be

balanced by the friction force. Therefore, friction force cannot be ignored when establishing the grip contact force model. The contact force at each contact point can be represented by a normal compressive force and two tangent friction forces

$F_{ci} = (f_{xi}, f_{yi}, f_{zi})$ . Based on the Coulomb friction model, the friction force applied along the shaft surface tangent direction can be considered as the function of the normal force. In addition, the normal force applied on the forge piece can only be compression force instead of tension force. Therefore, each component of the contact force  $F_{ci}$  should satisfy the cone of friction condition [5]:

$$FC_{ci} = \begin{cases} \sqrt{f_{yi}^2 + f_{zi}^2} \leq \mu f_{xi} & (i=1,2,3\dots k) \\ f_{xi} \geq 0 \end{cases} \quad (1)$$

Generally, wrench-based  $B_{ci} \in R^{6 \times mi}$  and cone of friction  $FC_{ci}$  are two of the widely used methods to represent the contact model. The dimension of wrench-based  $mi$  means the number of dimensions of wrench-based. The set of contact forces allowed for a given contact mode is

$$F_{ci} = B_{ci} f_{ci} \quad f_{ci} \in FC_{ci} \quad (2)$$

As  $GF = -w$  [7], and considering the friction of grip jaw, then the contact force model can be expressed as

$$(G_1 \quad G_2 \quad G_3 \quad G_4)(f_{c1} \quad f_{c2} \quad f_{c3} \quad f_{c4})^T = w \quad f_{ci} \in FC_{ci} \quad (3)$$

Where  $G$  is grip matrix,  $w$  is external wrench on workpiece.

$$G = (G_1 \quad G_2 \quad G_3 \quad G_4) \quad (4)$$

$$G_i = \begin{pmatrix} R_{oci} & 0 \\ \hat{P}_{oci} R_{oci} & R_{oci} \end{pmatrix} B_{ci} \quad (5)$$

$$B_{ci} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}^T \quad (6)$$

$$w = (f_x \quad f_y \quad f_z \quad m_x \quad m_y \quad m_z)^T \quad (7)$$

Assume the opening angle of jaws in Figure 2 is  $2\beta$ , then the pose of coordinate  $C_i$  at contact point  $i$  relative to the forge piece coordinate  $O$  is

$$R_{oc1} = \begin{pmatrix} -\cos\beta & -\sin\beta & 0 \\ \sin\beta & -\cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad P_{oc1} = \begin{pmatrix} r\cos\beta \\ -r\sin\beta \\ z_1 \end{pmatrix} \quad (8a)$$

$$R_{oc2} = \begin{pmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad P_{oc2} = \begin{pmatrix} -r\cos\beta \\ -r\sin\beta \\ z_2 \end{pmatrix} \quad (8b)$$

$$R_{oc3} = \begin{pmatrix} -\cos\beta & \sin\beta & 0 \\ -\sin\beta & -\cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad P_{oc3} = \begin{pmatrix} r\cos\beta \\ r\sin\beta \\ z_3 \end{pmatrix} \quad (8c)$$

$$R_{oc4} = \begin{pmatrix} \cos\beta & \sin\beta & 0 \\ -\sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad P_{oc4} = \begin{pmatrix} -r\cos\beta \\ r\sin\beta \\ z_4 \end{pmatrix} \quad (8d)$$

Where  $P_{oci}$  denotes the coordinates of the contact coordinate system origin  $o$  relative to the forge piece coordinate system  $C_i$ , and  $R_{oci}$  denotes the rotation matrix of contact coordinates relative to the forge piece coordinates.

Define the parameters used in the model as follow, eccentric distance  $e$ , eccentric angle  $\gamma$ , mass of the workpiece  $m$ , rotation angle of the workpiece  $\theta$ , falling angle of the workpiece  $\Phi$ , and distance between the center of gravity and the end of jaw  $L$ . Figure 3 shows the condition of workpiece when the grip system is in upright position. The external force screw  $w$  can be obtained from different eccentric angles.

$$f_x = -mg \sin \theta \quad (9a)$$

$$f_y = mg \cos \theta \cos \Phi \quad (9b)$$

$$f_z = mg \sin \Phi \quad (9c)$$

$$m_x = -mg \cos \theta \cos \Phi L - mg \sin \Phi \cos \gamma e \quad (10a)$$

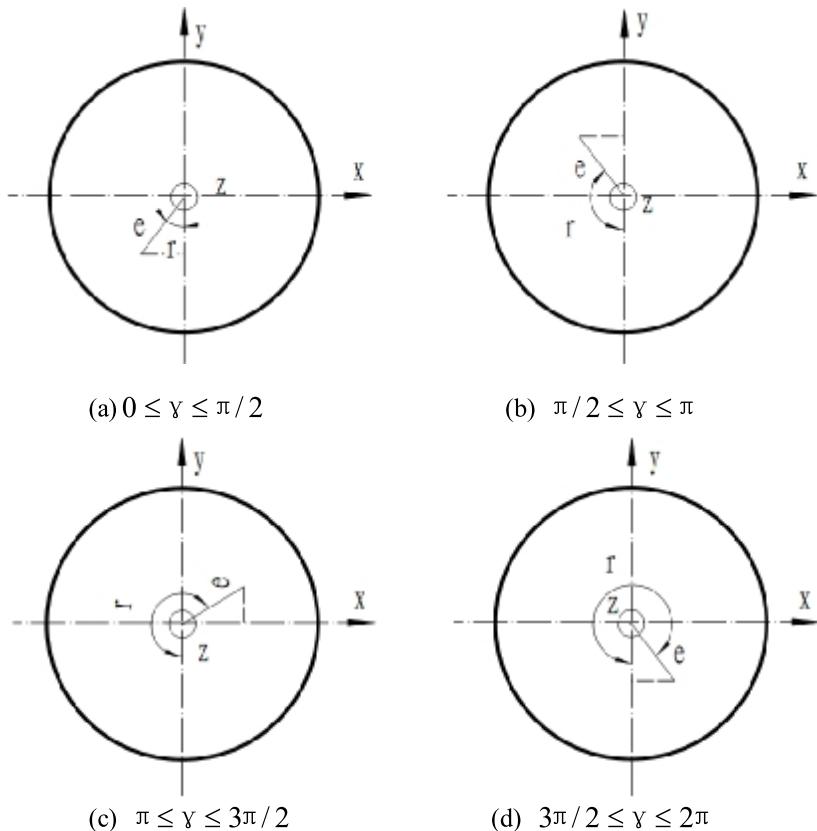
$$m_y = -mg \sin \theta L + mg \sin \Phi \sin \gamma e \quad (10b)$$

$$m_z = -mg \sin \theta \cos \Phi e - mg \cos \theta \cos \Phi \sin \gamma e \quad (10c)$$

After the contact force  $F$  applied on the jaw is obtained, then the gripping force of jaw  $P_{jc}$  can be calculated using the following formula,

$$P_{jc} = R_1 + R_2 = (f_{x1} + f_{x2})\sin\beta + (f_{x3} + f_{x4})\sin\beta \quad (11)$$

where  $R_1$  and  $R_2$  denote the grip forces of the upper and lower jaws, respectively.

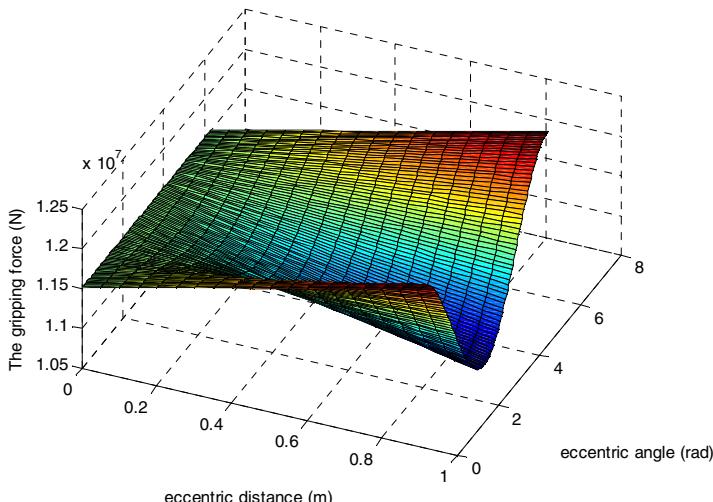


**Fig. 3.** Different conditions of eccentric forge piece

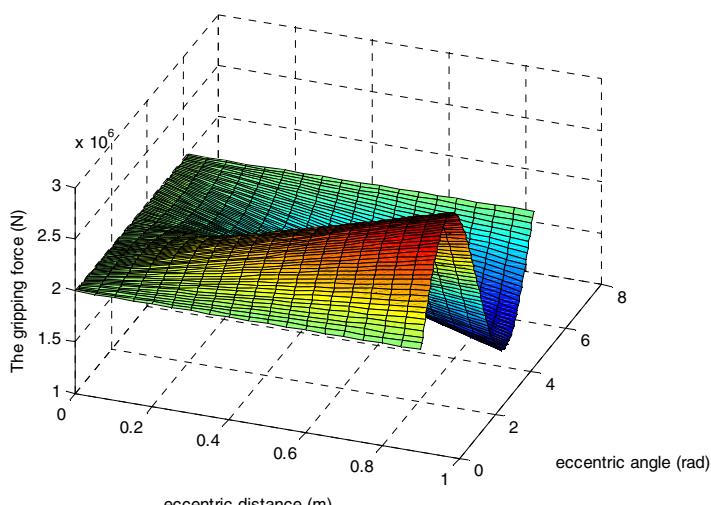
### 3 Analysis of Influence Factors

The required contact force at any position can be found by employing the stable gripping force closure principle and simplifying the nonlinear cone of friction constraint to linear constraint optimization with Lagrangian multipliers method [7-9]. The contact force at any given contact point is optimized using the linear constraint gradient flow method to obtain the minimal contact force at different contact points. Assume there is no falling angle produced by the forge piece, and the jaw angle is  $130^\circ$ , the radius and weight of forge piece are 1.13 m and 250 tons, respectively. The grip force of jaw varies with the eccentric distance and angle. When the jaw is in the vertical position

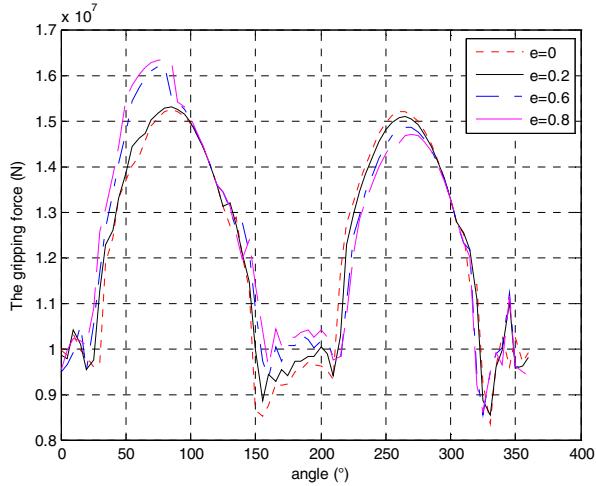
(Figure 4), with the increase of eccentric angle, the grip force decreases first in the range of 0 to 180°, and then increases in the range of 180 to 360°; with the increase of eccentric distance, the grip force increases when the eccentric angle is in the range of 0 to 90° and 270 to 360°, but decreases in the range of 90 to 270°. When the jaw is in the horizontal position (Figure 5), with the increase of eccentric angle, the grip force increases in the range of 0 to 90° and 270 to 360°, but decreases in the range of 90 to 270°; with the increase of eccentric distance, the grip force increases when the eccentric angle is in the range of 0 to 180°, but decreases when the eccentric angle is in the range of 180 to 360°.



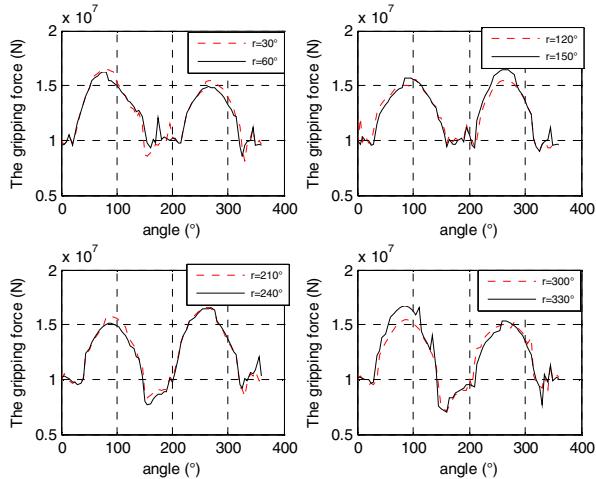
**Fig. 4.** Variation of gripping force with eccentric distance and angle (in the vertical position)



**Fig. 5.** Variation of gripping force with eccentric distance and angle (in the horizontal position)



**Fig. 6.** Gripping force at any position of forge piece with different eccentric distances



**Fig. 7.** Gripping force at any position of forge piece with different eccentric angles

Suppose the eccentric angle is  $60^\circ$ . The gripping force at any position of forge piece with different eccentric distances is shown in Figure 6. When the jaw is located in the ranges of  $30^\circ\sim 210^\circ$ , the gripping force increases with the increase of eccentric distance, because the contact point approaches the center of pin axis gradually; while the jaw is located in the range of  $210^\circ\sim 320^\circ$ , the gripping force decreases with the increase of eccentric distance, because the contact point departs the center of pin axis gradually. The variation of eccentric distance effect on the gripping force fluctuates when the jaw

is located in the ranges of  $0^\circ$  to  $30^\circ$  and  $320^\circ$  to  $360^\circ$ . When the jaw is located in the ranges of  $100^\circ$  to  $120^\circ$  and  $280^\circ$  to  $310^\circ$ , the effect of eccentric distance on the gripping force is small.

Suppose the eccentric distance is 0.6 m. The gripping force at any position of forge piece with different eccentric angles is shown in Figure 7. The trends of grip force with the rotation angle and the eccentric angle are summarized in Table 1. Table 1 indicates the trend of gripping force is completely opposite after the jaw being rotated  $180^\circ$ , this can be explained by the eccentric angles in the first and second quadrants are corresponding to the eccentric angles in the third and fourth quadrants before rotation.

**Table 1.** Trends of gripping force with the position of forge piece and eccentric angle

	1 <sup>st</sup> quadrant	2nd quadrant	3rd quadrant	4th quadrant
$0^\circ-90^\circ$	↓	↓	↑	↑
$90^\circ-180^\circ$	↑	↓	↓	↑
$180^\circ-270^\circ$	↑	↑	↓	↓
$270^\circ-360^\circ$	↓	↑	↑	↓

First row: the quadrant where the eccentric angle is located in.

First column: rotation angle, namely, the position of the jaw.

↑ and ↓ represent the gripping force increase and decrease with the rotation angle increases, respectively.

Theoretically, when the jaw is located in the range of  $180^\circ$  to  $210^\circ$ , the gripping force should decrease with the increase of eccentric angle. However, the calculated result (Figure 7) is contradicted to the theoretically predicted trend in this range. In addition, the curves in Figure 6 and 7 are not smooth when the jaw is located at certain angles due to the computational algorithm. Using optimization method of Lagrange multipliers, the nonlinear cone of friction constraint was simplified to a linear constraint to search the required contact forces at any position. The contact force at any given contact point was optimized by linear constraint gradient current method [7-9]. The searching process is stopped when the contact force satisfying the force closed condition is found during searching the position of contact point. As a result, there are some discontinuities in the curves. We are improving the algorithm to solve these problems. At the same time, experimental verification is in progress.

## 4 Conclusions

The gripping forces of forging manipulator are analyzed. A mathematic model with point contact was established to describe the friction force and gripping force between the jaws and forge piece. The eccentric of workpiece was considered in this model, and the influence of eccentric on the gripping force is studied. This model provides theoretical foundation to develop and evaluate the load-carrying capability of gripping devices under large scale heavy load conditions.

## Acknowledgement

This work was supported by the National Basic Research Program of China under Grant 2006CB705400.

## References

1. Wan, S., Wang, Y.: Mechanization and automation of Forging. China machine press, Beijing (1983) (in Chinese)
2. Liu, S., Deng, H., He, J.: Dynamic model of the tong force and analysis of load-carrying ability of the large scale heavy-load tong device. Modern Manufacturing Engineering 1, 107–111 (2010) (in Chinese)
3. Li, Q., Gao, D., Deng, H.: Analysis of Contact forces for Forging Manipulator Grippers. In: Proceedings of IEEE CIS\_RAM 2008, Chengdu, pp. 716–719 (2008)
4. Li, Q., Gao, D., Deng, H.: Influence of Contact Forces on Stable Gripping for Large-Scale Heavy Manipulator. In: Xiong, C.-H., Liu, H., Huang, Y., Xiong, Y.L. (eds.) ICIRA 2008. LNCS (LNAI), vol. 5315, pp. 839–847. Springer, Heidelberg (2008)
5. Zuo, B., Qian, W.: A nonlinear programming algorithm for force-closure analysis in multi-fingered robotic grasping. Chinese Journal of Mechanical Engineering 2, 205–215 (1999) (in Chinese)
6. Zhang, Y., Li, J., Li, J.: Robot hands. China machine press, Beijing (2007) (in Chinese)
7. Muway, R.M., Li, Z.x., Shankar Sastry, S.: A Mathematical Introduction to Robotic Manipulation. CRC Press, Inc., Boca Raton (1994)
8. Li, Q., Gao, D., Deng, H.: Calculation of Contact Forces of Large-Scale Heavy Forging Manipulator Grippers. Key Engineering Materials 419–420, 645–648 (2010)
9. Buss, M., Hashimoto, H., Moore, J.B.: Dextrous hand grasping force optimization. IEEE Transactions on Robotics and Automation 12, 406–418 (1996)

# A Contact/Impact Analysis of Rigid Bodies Based on Newmark Direct Integration Method

Zhen Xu and Hua Deng

School of Mechanical and Electrical Engineering  
Central South University  
Changsha, 410083, China  
xuzhen21@163.com

**Abstract.** An analytical approach for contact bodies of heavy-manipulator is presented in this paper based on the derivation of an entirety motion equation in the contact region and an entirety analysis method for 3D contact/impact problem is proposed. Using this method, contact force is derived and solved under the common motion status and is used to be the criteria for evaluating contact conditions. This method is used to simulate the collision of two bodies and its validity is verified.

**Keywords:** Contact/Impact problem; Entirety analysis method; Newmark direct integration; Heavy-manipulator.

## 1 Introduction

The movement of large mechanism with heavy load is usually obtained by contact restraint. Among various motions of mechanism, contact motion is a complex process involving multi-body engagement, multi-point contact and particularly nonlinear characteristics of discontinuous contact. To analyze the contact status of heavy mechanism during motion and to enhance accurate movement of mechanism and reliability, contact force and motion state of contact mechanism are important parts in mechanism design.

In recent years the study on the static and dynamic contact problems of mechanism has been reported by many researchers [1–4]. Some researchers developed different formulations to solve the differential equations of motion for frictional dynamic contact/impact problems. Other researchers have been working on finite element based methods for dynamic contact/impact problems and particularly their applications to gear driven systems [5,6].

In this paper, a simple method for the analysis of contact and motion of multi-bodies is proposed, which can be used for the design of gear driven systems and the coordination of heavy-manipulators. In the following sections, the basic theoretical model used for the analysis of contact/impact problem is firstly presented. This is followed by the solution of entirety motion formulation of dynamic contact/impact problems based

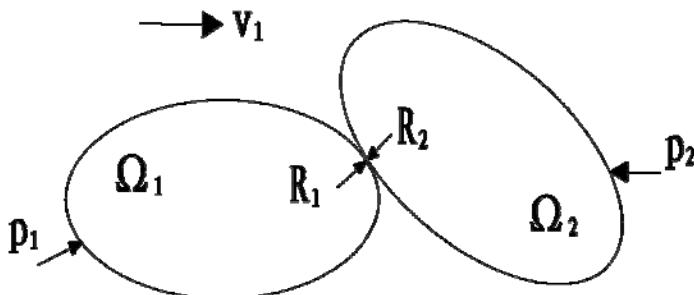
on newmark direct integration method. Finally, a collision of two bars is simulated to verify the method and the conclusions are given at the end of the paper.

## 2 Contact/Impact Problem for Two Bodies

The dynamic contact model of two bodies is shown in Fig 1, in which body  $\Omega_1$  is impacting toward body  $\Omega_2$  with the velocity  $v_1$ .  $p_i$  ( $i=1,2$ ) are the total external forces for two bodies respectively.  $R_i$  ( $i=1,2$ ) are the contact forces of the two bodies. In this contact process, they sometimes contact or separate each other. Thus, there are nonlinear characteristics of discontinuous motion. In terms of the d'Alembert principle, the dynamic contact problem of two bodies may be described as follows [7]:

$$m_i a_i(t) + c_i v_i(t) + k_i u_i(t) = p_i(t) + R_i(t) \quad (i \in \Omega_1, \Omega_2) \quad (1)$$

Where  $m_i$ ,  $c_i$ ,  $k_i$  are the mass, damping and stiffness of the two bodies, respectively.  $p_i(t)$ ,  $R_i(t)$  are the applied load and the contact force of the two bodies.  $u_i(t)$ ,  $v_i(t)$ ,  $a_i(t)$  are the displacement, velocity and acceleration of the two bodies respectively.



**Fig. 1.** Dynamic contact model of two bodies

## 3 Newmark Direct Integration Based Solution of Entirety Motion Equation

The Newmark direct integration method which is widely used for numerical solution in dynamic equations is adopted to solve Eq.(1). In this method, it is assumed that: ( $t_{n+1} = t_n + \Delta t_n$ )

$$u(t_{n+1}) = u(t_n) + v(t_n) \Delta t + [(\frac{1}{2} - \beta)a(t_n) + \beta a(t_{n+1})] \Delta t^2 \quad (2)$$

$$v(t_{n+1}) = v(t_n) + [(1 - \gamma)a(t_n) + \gamma a(t_{n+1})]\Delta t \quad (3)$$

where  $\beta$  and  $\gamma$  are adjustable parameters depending on the integration accuracy and stability. When  $\beta \geq \frac{1+\gamma^2}{2} / 4$  and  $\gamma \geq \frac{1}{2}$ , it is confirmed that the Newmark method has unconditional stability [1,6]. In this research,  $\beta = \frac{1}{4}$  and  $\gamma = \frac{1}{2}$  are used. Substituting Eqs.(2) and (3) into Eq.(1) yields the effective stiffness equation of the dynamic contact/impact bodies, in which displacement  $u_i(t_{n+1})$  in the equation can be solved[8]:

$$\bar{k}_i u_i(t_{n+1}) = \bar{p}_i(t_{n+1}) + R_i(t_{n+1}), (i \in \Omega_1, \Omega_2) \quad (4)$$

Where

$$\bar{k}_i = k_i + \frac{1}{\beta \Delta t^2} m_i + \frac{\gamma}{\beta \Delta t} c_i \quad (5)$$

$$\begin{aligned} \bar{p}_i(t_{n+1}) = & p_i(t_{n+1}) + m_i \left[ \frac{1}{\beta \Delta t^2} u_i(t_n) + \frac{1}{\beta \Delta t} v_i(t_n) + \left( \frac{1}{2\beta} - 1 \right) a_i(t_n) \right] \\ & + c_i \left[ \frac{\gamma}{\beta \Delta t} u_i(t_n) + \left( \frac{\gamma}{\beta} - 1 \right) v_i(t_n) + \left( \frac{\gamma}{2\beta} - 1 \right) a_i(t_n) \Delta t \right] \end{aligned} \quad (6)$$

In reference [8], in order to solve the displacement  $u_i(t_{n+1})$ , the flexibility equation are used to solve contact force  $R$  and then the displacement of the bodies is also solved through the Eq.(4), so the acceleration and velocity are obtained in the Eqs.(2) and (3). Moreover the contact force  $R$  is firstly solved and the displacement  $u$  is secondly obtained [8]. Because the first solution of contact force are inconvenient through introducing the flexibility equation. Therefore the displacement  $u$  is firstly solved in this research by means of Eq.(4), the acceleration and velocity are also obtained through Eqs.(2) and (3), and the contact force  $R$  is lastly obtained by using Eq.(1).

To justify the contact condition, it is assumed that the friction is not considered when the two bodies contact each other. Two possible contact conditions are defined as stick and separation.

Two bodies will have the same displacement, velocity and acceleration due to the sticking of the two bodies, so the two bodies can be considered as one body in the stick condition. Hence the contact force  $R(t)$  becomes the internal force of two bodies. Then equation (1) can be formed as

$$m_{12}a(t) + c_{12}v(t) + k_{12}u(t) = p_{12}(t) \quad (7)$$

where  $a(t)$ ,  $v(t)$ ,  $u(t)$  are the common acceleration, velocity and displacement of the two bodies.  $m_{12}$  is the mass of the two bodies, i.e.  $m_{12} = m_1 + m_2$ ;  $c_{12}$  is the minimum damping of the two bodies, i.e.  $c_{12} = \min\{c_1, c_2\}$ ;  $k_{12}$  is the minimum stiffness of the two bodies, i.e.  $k_{12} = \min\{k_1, k_2\}$ .  $p_{12}(t)$  is the whole external force of the two bodies.

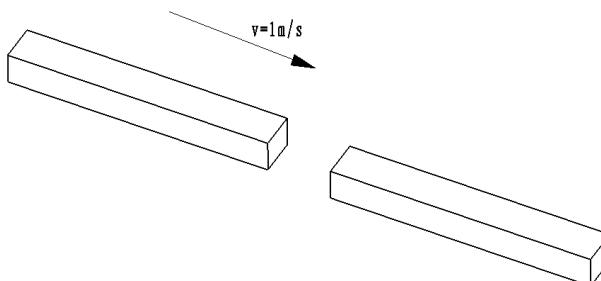
The separation condition indicates that the two contact bodies are separated. If the displacement, velocity and acceleration of the two bodies are not identical, it means that the contact force is zero ( $R(t) = 0$ ). Therefore the Eq. (1) may be rewritten as

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} a_1(t) \\ a_2(t) \end{bmatrix} + \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} + \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = \begin{bmatrix} p_1(t) \\ p_2(t) \end{bmatrix} \quad (8)$$

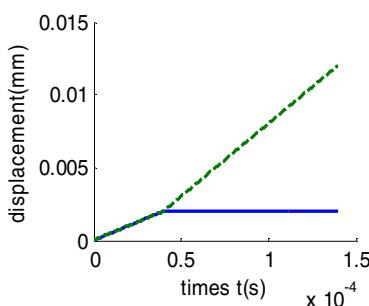
In order to confirm the change of contact conditions, the contact force is used to judge the stick condition or separation condition. If  $R(t)$  is not equal to zero, it states that the contact is in the stick condition. If  $R(t)$  is equal to zero, the two bodies already separate each other. Meanwhile the equations under different contact conditions are also changed in the computational iterations.

#### 4 Numerical Example of Dynamic Contact Problem

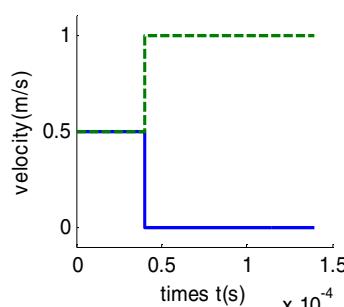
In order to verify the proposed method, simulations of a benchmark problem [8] considering the center-to-center collision between two identical bars is performed as shown in Fig. 2. A bar is striking the other bar with a velocity of 1m/s. The time step is  $dt=2\times10^{-7}$ s. Young's Modulus is  $E=2.1\times10^{11}$ N/m<sup>2</sup> and Density is  $\rho=7.8\times10^3$ kg/m<sup>3</sup>.



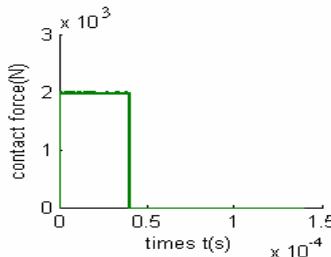
**Fig. 2.** Collision of two bars



**Fig. 3.** Time domain of displacement



**Fig. 4.** Time domain of velocity



**Fig. 5.** Time domain of dynamic contact

The results of numerical solution for two bars in time domain of the displacement and velocity on the contact bodies are shown in Fig.3 and Fig.4, where the solid lines represent the first body and the dashed lines represent the second body. Fig. 5 shows the contact force on the contact bodies. It can be seen that the results solved by the proposed method in this research have a good agreement with the results of reference [8]. However, the idea of entirety computation method is more simple than the finite element method for the solution of contact forces, so the computation time of this method is shorter than the method in reference [8].

## 5 Conclusions

In this paper, an entirety computation method is proposed for 3D dynamic contact/impact problems. This method is based on the derivation of the entirety motion equation of two contact bodies, which has the same movement of displacement, velocity and acceleration. The contact force is considered as the computational criteria for the change of contact conditions, which is very efficient for the solution of contact bodies. A numerical computation example is simulated through the proposed methods, which verify that this method is validity in the simulation of static and dynamic contact analysis.

## Acknowledgement

This work was supported by the National Basic Research Program of China under Grant 2006CB705400.

## References

1. Hughes, T.J.R., Taylor, R.L., Sackman, J.L., Curnier, A., Kanok-nulchai, W.: A finite element method for a class of contact-impact problems, *Comput. Methods Appl. Mech. Engrg.* 8, 249–276 (1976)
2. Armero, F., Petocz, E.: Formulation and analysis of conserving algorithms for frictionless dynamic contact/impact problem. *Comput. Methods Appl. Mech. Engrg.* 158, 269–300 (1998)

3. Bittencourt, E., Creus, G.J.: Finite element analysis of three-dimensional contact and impact in large deformation problems. *Comput. Struct.* 69, 219–234 (1998)
4. Demkowicz, L., Bajer, A.: Conservative discretization of contact/impact problems for nearly rigid bodies, *Comput. Methods Appl. Mech. Engrg.* 190, 1903–1924 (2001)
5. Lin, T.: Numerical Simulation and Experiment Research on Nonlinear Impact Vibration for Gear System, Ph.D. Dissertation, Chongqing University, Chongqing (1999)
6. Lin, T., Li, R., Chen, B., Tang, Q.: Nonlinear dynamic contact analysis of meshing gears. In: *Proc. Int. Conf. Mech. Transmission*, pp. 248–251 (2001)
7. Bathe, K.J.: *Finite Element Procedures*. Prentice-Hall, New Jersey (1996)
8. Lin, T., Ou, H., Li, R.: A finite element method for 3D static and dynamic contact/impact analysis of gear drives. *Comput. Methods Appl. Mech. Engrg.* 196, 1716–1728 (2007)

# RFID Enabled Remote Structural Health Monitoring for Heavy Lifting Appliances

Xubing Chen and Fan Zhou

School of Mechanical and Electrical Engineering, Wuhan Institute of Technology,  
Wuhan 430073, P.R. China  
bluegif@gmail.com

**Abstract.** To realize remote structural health monitoring for heavy lifting appliances, e.g., 500t gantry cranes, technologies of RFID wireless sensor network, LabVIEW virtual instrument and health diagnosis expert system are employed to establish an integrated system. The system is divided into three layers in logic structures: the data acquisition layer, the network communication layer and the system service layer. In the data acquisition layer, RFID sensors, RFID readers are deployed to collect the structural strains, vibrations, noises and other motion parameters for each monitored object. In the system service layer, monitoring and analysis services and a health diagnose expert system are employed to carry out real-time monitoring and safety evaluation tasks. The data acquisition points connect the centre platform through LabVIEW communication tunnels. As the result of a comparison with the common optics sensing system, the proposed system has merits of wireless, easy to deploy, passive and save energy.

**Keywords:** Heavy Lifting appliance, remote structural health monitoring, RFID sensor, LabVIEW, expert system.

## 1 Introduction

Structural health monitoring is currently a hot issue in the engineering and academic researches, and can be used to assess the structural integrity of the monitored objects accurately and reliably, reduce maintenance costs, extend service life and promise safety. Remote structural health monitoring systems employing a sensor-based quantitative assessment of in-service demands and structural condition are perceived as the future in long-term management programs. As reported, there are many remote structural monitoring applications in bridge and mechanical fields. Matthew deployed a large-scale wireless sensor network for ambient vibration testing of a single-span integral abutment bridge to derive in-service modal parameters remotely [1]. Pan et al. realized a remote online machine condition monitoring system built up in the architecture of both the Borland C++ Builder (BCB) software-developing environment and Internet transmission communication [2]. Hameed reviewed different techniques, methodologies and algorithms for monitoring the performance of wind turbine as well

as for an early fault detection to keep away the wind turbines from catastrophic conditions due to sudden breakdowns [3].

RFID, Radio Frequency Identification, is one of the key enablers of the Internet of Things and has caught the eyeballs of structural monitoring engineers for its wide applications and merits of anti-pollution, durable, reusable, small size, diversity and security [4].

Start from the applications of Wal-Mart and DOD (United States Department of Defense), RFID has played a significant role in supply chain, logistics and retail businesses. Alejandro et al. suggested a methodology to turn the MT, a Spanish company developed innovative and ecological packaging and transport unit for the grocery supply chain, into an intelligent product platform by embedding active RFID tags [5]. Chow et al. proposed a real-time knowledge support framework for the development of an RFID-multi-agent based process knowledge-based system which has the ability to solve dynamic logistics process management problems [6]. Lee et al. developed an RFID-based reverse logistics framework and optimization of locations of acquisition points which allow economically and ecologically reasonable recycling [7]. Zhou et al. emerged RFID as the hottest information tracing technology in supply chain management with its inherent ability to reveal item-level product information [8].

Besides of supply chain, logistics and retail businesses, mechanical and engineering are also the major battle fields of RFID applications. Soungho et al. introduced the application of RFID technology to prevention of collision accidents with heavy equipment such as hydraulic excavators and cranes [9]. In recognition of the need for more effective and efficient factory system integration solutions, Qiu proposed a corresponding framework to enable the instant delivery of pertinent data and information on a uniquely identifiable job/product at point-of-need across factories and presented a small-scale manufacturing application to show how the proposed approach can be implemented on the shop floor [10]. Chen et al. proposed a multi-agent system framework called agent-based manufacturing control and coordination (AMCC) system, an agent-based framework using ontology and RFID technology to monitor and control dynamic production flows and also to improve the traceability and visibility of mass customization manufacturing processes [11]. Huang et al. presented an affordable solution to these problems by using RFID wireless manufacturing (WM)—an emerging advanced manufacturing technology (AMT) for implementing the concept of Just-In-Time (JIT) manufacturing to reduce the shop-floor work-in-progress (WIP) inventories and smoothening their flows through real-time information visibility and traceability [12]. Chuong et al. discussed a framework (called Radio-Mama) using a FRID technology for real-time management of mobile assets [13]. Ko provided a web-based RFID building maintenance system to improve facility and equipment maintenance efficiency [14] and proposed an RFID 3D sensing algorithm to assist construction managers in locating the positions of relevant personnel, equipment, apparatus and materials [15].

In the bottom layer of applications in environment and structural monitoring, many different RFID sensors and systems are reported. Abarca et al. described the design and manufacture of a prototype (currently in use) of an intelligent RFID sensor for the tracking and monitoring of blood temperature over the entire process [16]. Abad et al. demonstrated the validation of a RFID smart tag developed for real-time traceability

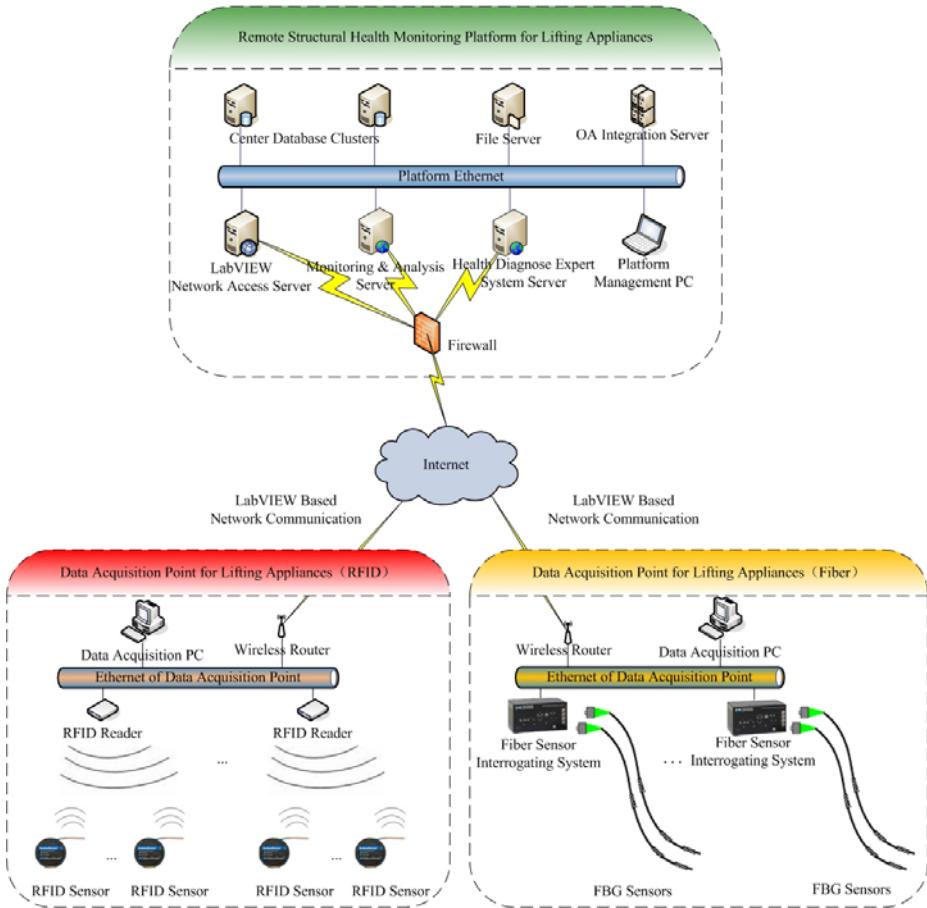
and cold chain monitoring for food applications [17]. Alexandru et al. produced capacitive humidity sensors with intelligent RFID tags from polymer foils and/or thin films and successfully tested them on experimental platforms well reproducing the real operation conditions [18]. Radislav et al. demonstrated an attractive approach to adapt conventional radio frequency identification (RFID) tags for multi-analyze chemical sensing and evaluated the capability for such position-independent analyze quantification using multivariate analysis tools [19]. Lin et al. presented a MEMS (micro electro-mechanical system) capacitive-based pure bending strain sensor for use in wireless spinal fusion monitoring and developed an analytical model for the sensor mounted to a cantilever test bar and compared to experimental results of actual devices [20]. Wang proposed an RFID-based quality inspection and management (RFID-QIM) system, which functions as a platform for gathering, filtering, managing, monitoring and sharing quality data of constructions [21].

The objective of this research is to establish a remote structural health monitoring system for heavy lifting appliances based on some advanced technologies to achieve remote and real-time structural health monitoring, and discover early warning forecasts for the structural safety of lifting appliances to assess the security conditions. The rest of the paper is organized as follows. First of all, the general system schema is introduced in Section 2. Next, the key technologies, such as the RFID wireless sensor network, the LabVIEW virtual instrument and the health diagnosis expert system are described in Section 3. In Section 4, a comparison has been provided between the proposed system schema and the commonly used optics sensing schema, and some application problems are discussed. Finally, the paper is concluded in Section 5.

## 2 General System Schema

The proposed remote structural health monitoring system provides key parameters of lifting appliance structures during their operations to analyze the structures by safety evaluation and structural remaining life analysis, recommend preventive maintenance programs, capture structural failures before accident bursting out, and discover potential risks to prevent sudden catastrophic accidents, heavy casualties and property losses. The general system schema is composed of data acquisition points and a center platform, as shown in Fig. 1.

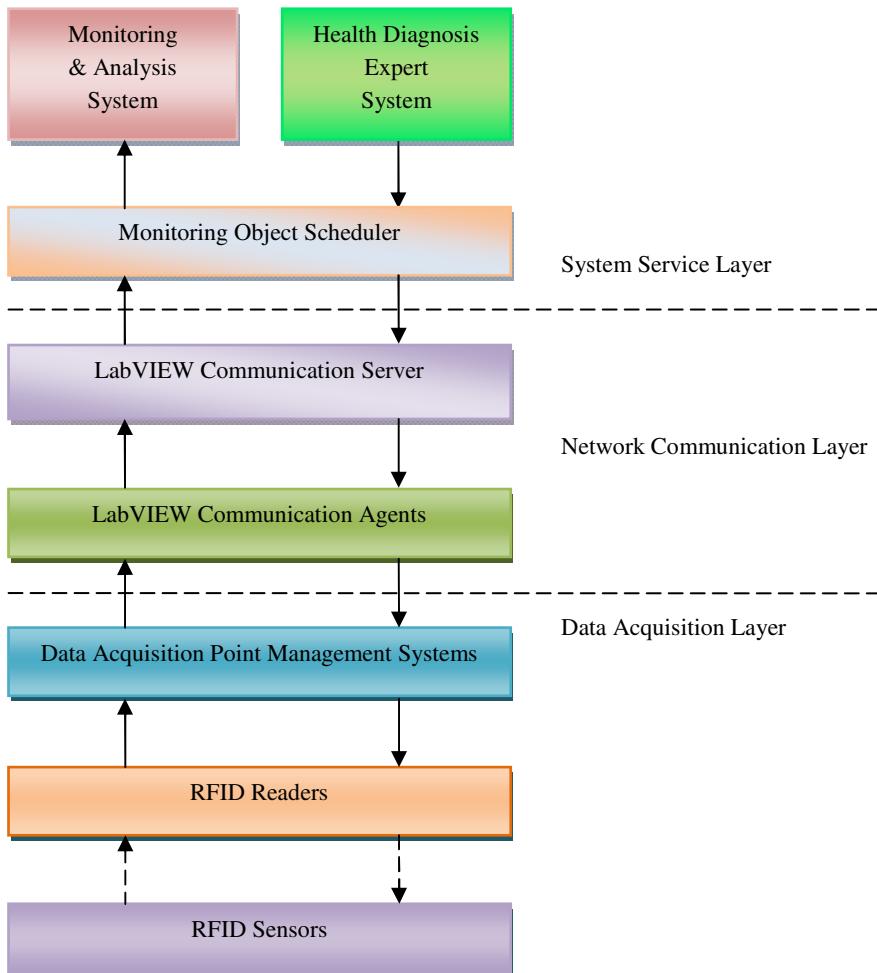
In Fig. 1, the data acquisition system is installed in or near each heavy lifting appliance according to the monitoring requirements. The system is including of several RFID wireless sensors deployed in the structures of the heavy lifting appliance, a proper quantity of RFID readers set nearby, and a data acquisition terminal computer installed with a wireless router. RFID readers read RFID sensors to collect the structural strains, vibrations, noises and other motion parameters and report these parameters to the data acquisition terminal. After pretreatment, data acquisition terminal reports the pretreated monitoring data to the remote structural health monitoring platform by wireless access network. The platform collects all parameter formations to describe monitoring and analysis curves, while the health diagnosis expert system detects the health state of lifting appliances and recommends preventive maintenance programs.



**Fig. 1.** Network structure of the RFID enabled remote structural monitoring system

Besides the RFID data acquisition points, there may be some fiber data acquisition points existed in engineering applications. To meet the compatible requirements, the proposed remote structural health monitoring system has to support both the two kinds of data acquisition points at the same time. In the fiber case, FBG (Fiber Bragg Grating) sensors are constructed in a short segment of optical fiber that reflects particular wavelengths of light and transmits all others. Fiber sensor interrogating systems are employed to carry out the photoelectric conversions. All the data acquisition points are connected to the remote structural health monitoring platform for lifting appliances by wireless networks, such as GPRS, CDMA, 3G, and so on.

As shown in Fig. 2, the system software structure is divided into three layers: the data acquisition layer, network communication layer and the system service layer. Therefore, the proposed RFID enabled remote structural monitoring system has formed the framework of a distributed system: distributed acquisition and centralized diagnosis.



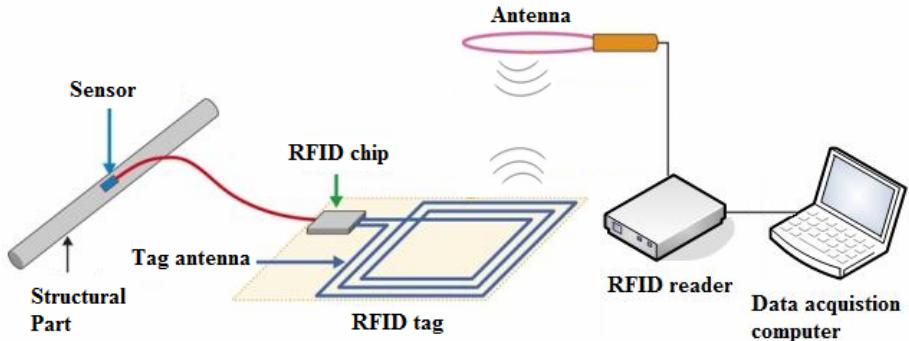
**Fig. 2.** Three-layer structure of the remote health monitoring system

### 3 Key Technologies

#### 3.1 RFID Wireless Sensor Network

Anti-metal RFID sensors can be mounted on the metal surfaces directly for their merits of excellent resistances to electromagnetic interferences, and outstanding performance of suppressing typical noises in industrial environments. The passive RFID sensors acquire energy by the radio frequency activation of readers, and have long time stabilities. Just because this property, a disconnected RFID reader will access many RFID sensors in its working area with the help of anti-collision algorithms. All the RFID sensors in a monitored lifting appliance are forming an RFID wireless sensor network, and maintenance wireless communications with the help of the

formation of a multi-hop and self-organization network, which aims to collaborate in perception, acquisition and processing of network coverage area to prepare perceptual object information for observers.



**Fig. 3.** Basic working principle of RFID sensors

The basic working principle of RFID sensors is described in Fig. 3. An RFID sensor is composed of a sensor and an RFID tag. And they are connected by a short electrical wire. The tag acquires energy from the radio frequency waves of an RFID reader and drives the sensor to work. After that, the returned monitoring parameter is emitted by the RFID chip and obtained by the RFID reader again. Therefore, the RFID wireless sensor network keeps silence at most case, but changes to active by the frequent scanning of RFID readers. The merits of this mechanism are wireless, easy to deploy, passive and save energy.

The technical parameters of the proposed RFID sensors are listed in Table 1.

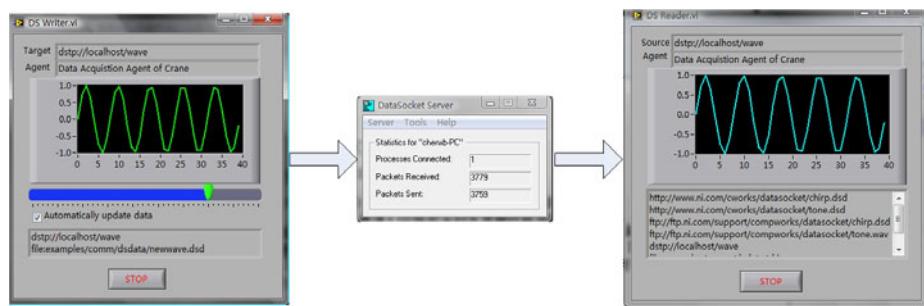
**Table 1.** Major technical parameters of the proposed RFID sensors

Index	Description
<b>Measurement parameters</b>	pressure, strain, temperature, vibration or noise
<b>Working frequency</b>	920MHz-925MHz
<b>Reading and writing distance</b>	0m-20m
<b>Data keep time</b>	>10 years
<b>Anti-collision technology</b>	for multi-tag reading
<b>Flooding test</b>	20°C, 24 hours, the depth of 2m (does not affect the read and write performance)
<b>Chemical tests</b>	bleach / caustic soda (20°C/10 hours, does not affect the appearance and read-write performance)
<b>Equilibrium pressure</b>	20MPa 5 minutes (look no cracks or broken)
<b>Axial stress</b>	1KPa 10 seconds (look no cracks or broken)
<b>Vibration test</b>	vibration frequency of 32Hz, vibration intensity 50mm / 24 hours (look no cracks or broken)
<b>Operating temperature</b>	-40°C to +100°C

### 3.2 LabVIEW Virtual Instrument

Laboratory Virtual Instrument Engineering Workbench (LabVIEW) of National Instrument (NI) is the most applied software in the fields of data acquisition, process monitoring and automatic testing, which is the de facto industry standard in the field of testing and measurement. The development of network measurement technologies has boosted the network based virtual instruments, which decompose the data acquisition module, data processing module and the data display module into different computers to form remote monitoring systems distributed in wide operation locations. LabVIEW has powerful data acquisition and network functions, and has the merits of a program tool for perfect remote monitoring systems. The major merits of LabVIEW are described as follow.

1. Virtual instrument technology employs high-performance modular hardware and efficient and flexible software to realize various test, measurement and automation applications. Compared with traditional instruments, it has four major advantages: high extensible, less developing time and perfect integration ability.
2. Both integrated DAQ (Data Acquisition) hardware and automatic test plans are provided. Many DAQ devices have direct memory access capability, and transfer the collected data directly through the PCI bus to the RAM, while CPU time is not consumed and high-speed data throughput is guaranteed.
3. Versatile and powerful methods are available for developing distributed applications: TCP/IP, Internet Toolkit, VI Server, Front Panel Web Publishing, Remote Data Acquisition (RDA), DataSocket, and so on.
4. With the help of DataSocket, data can be published and received from client applications easily without worrying about data formats and network protocols. And LabVIEW applications can easily share live data with a variety of clients, including of Visual Basic applications, Web browsers, Visual C++, Microsoft Excel, LabWindows/CVI, and other LabVIEW applications.



**Fig. 4.** Data flow of a LabVIEW DataSocket system

As shown in Fig. 4, with the help of a DataSocket Server, lightweight and stand-alone writers can broadcast live measurement data at high rates across the Internet to multiple remote readers concurrently. Because the DataSocket Server is a stand-alone

component, it simplifies network (TCP/IP) programming by automatically managing connections to clients and automatically converting measurement data to and from the stream of bytes sent across the network. By this means, parsing code is not needed. And as the DataSocket Server can run on any machine on a network, performance and security are also promised by isolating the web connections from the acquisition application.

### 3.3 Health Diagnose Expert System

There are many different kinds of characteristic parameters needed to be detected by sensors in the structural health monitoring process, such as strain, vibration, noise, temperature, and motion parameters. Sensor fusion combines sensory data to get better results than these data are employed individually. With the help of sensor fusion technology, the periodic or continuous changes in the trend of device state law are monitored continuously to estimate the remaining life of the lifting appliance and predict potential catastrophes. The diagnosis methods for typical parameters are listed in Table 2.

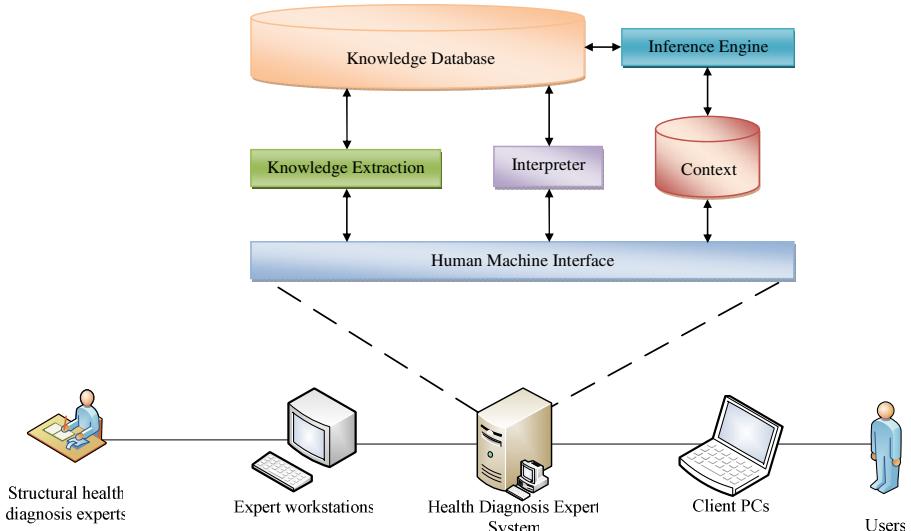
**Table 2.** General diagnosis technologies and their diagnosis parameters

Diagnosis method	Diagnosis parameter
Vibration diagnosis	Mechanical vibration, shock, mechanical mobility and modal parameters
Acoustic diagnosis	Noise (sound pressure and sound intensity), noise resistance, ultrasonic, acoustic emission
Temperature diagnostic	Temperature, temperature difference, temperature field, thermal imaging
Diagnosis of pollutants	Leakage, residues, gas, liquid, solid particle content changes
Optical diagnosis	Brightness, spectrum and radiation effects
Performance trend diagnosis	Main performance of mechanical equipment
Strength diagnosis	Force, stress, torque
Pressure diagnosis	Pressure, differential pressure and pressure fluctuation
Diagnosis of electrical parameters	Current, voltage, resistance, power and other signals and magnetic properties
Surface morphology diagnostics	Deformation, cracks, spots, pits, color, etc.

Generally, it is very hard to discover the required result from a mass of and continuous sensor data flow. Valuable experiences of experts are required to decrease the discovering difficulty and improve the preciseness of health analysis. A health diagnose expert system is scheduled to play the role of experts here. The self-learning diagnostic expert system takes advantage of the knowledge concluded by human experts and solves health diagnosis problems in similar ways to analyze the health status of lifting appliances and propose preventive maintenance programs. Moreover, the proposed health diagnose expert system is not only a simple computer-aided diagnosis

system, but also an artificial intelligence computer system full of valuable experiences extracted from a group of device management experts. It employs the knowledge and experiences of experts to summarize the diagnosis rules, constitute a knowledge base, simulate expert reasoning by the inference engine, resolve the various complex issues, and finally provide correct suggests and operational guidance to users.

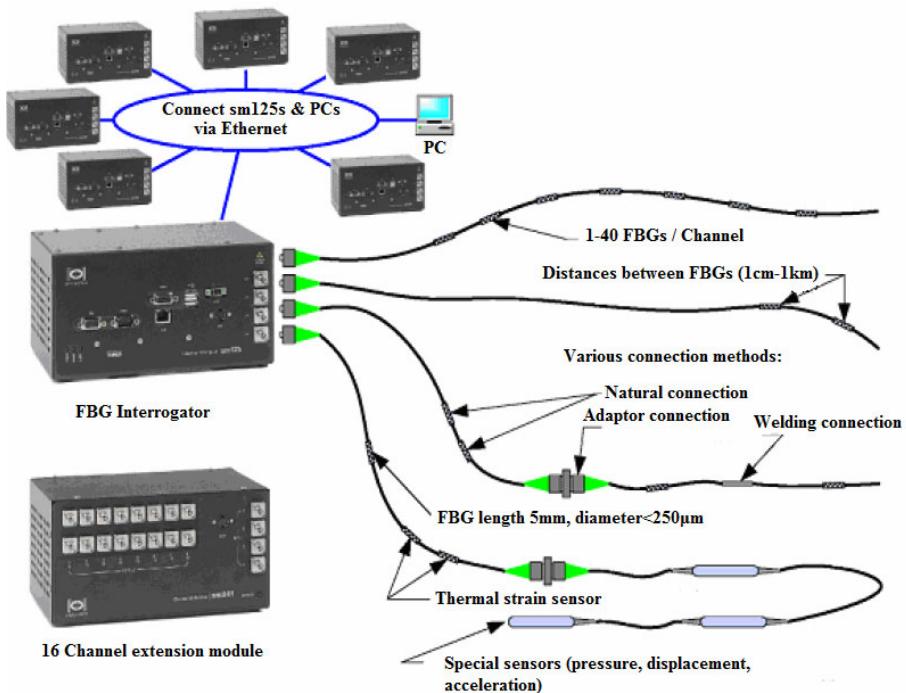
As shown in Fig. 5, an expert system usually consists of six components: knowledge database, context (i.e. the global database), inference engine, knowledge acquisition mechanism, explanation mechanism and human-machine interface. It applies artificial intelligence and computer technology, according to several experts in this field of knowledge and experience, reasoning and judgments, simulation of human decision-making process of experts to address the needs of human experts who deal with complex problems. Meanwhile, in the long-term applications, the health diagnosis expert system accumulates experiences and enhances its capabilities of analyzing and solving problems by self-learning.



**Fig. 5.** Software structure of the health diagnosis expert system

## 4 Discussions

Compared with the common used health monitoring schema of lifting appliances, which is based on the fibre Bragg grating (FBG) technology as shown in Fig. 6, the proposed schema is based on passive RFID technology, the sensor is adhered in the structures of heavy lifting appliances individually, and no optic fibre is needed to connect sensors in sequence to report the detected parameters to FGB interrogators. The tasks of deploying sensors in the structures of heavy lifting appliances are easy and convenient. What is more, as the passive RFID sensors are worked without power supply, the proposed system has merits of wireless, easy to deploy, passive and save energy.



**Fig. 6.** Micron optics sensing schema

But there are also demerits existed in the realization of the proposed schema. As there often are plenty of electric disturbances and metal structures in the working places of heavy lifting appliances, the working distances of RFID sensors are decreased rapidly. So as to RFID readers are needed to be deployed near RFID sensors to acquire suitable reading performances. Therefore, the burdensome and tedious wire deploying tasks are partial existed, and there is also wire deployment task of RFID readers. In a way, the deployed structural health monitoring system is a tree with connected trunks and free lived leaves.

## 5 Conclusions

Aiming at the realization of remote structural health monitoring for heavy lifting appliances, e.g., 500 t gantry cranes, RFID wireless sensor network, LabVIEW virtual instrument and health diagnosis expert system are employed to establish an integrated system. The system is divided into three layers in logic structures: the data acquisition layer, the network communication layer and the system service layer. In data acquisition layer, RFID sensors, RFID readers and a data acquisition computer are deployed to collect the structural strains, vibrations, noises and other motion parameters for each monitored heavy lifting appliance. In the system service layer, monitoring and analysis service and health diagnose expert system are deployed to carry out real-time

monitoring and safety evaluation tasks. The two layers are connected by LabVIEW communication tunnels. As the result of a comparison with the common optics sensing system, the proposed system has merits of wireless, easy to deploy, passive and save energy.

However, there are some limitations with the proposed methodology and future work is required. Currently, heavy lifting appliances have huge sizes, but the accessing distances of RFID sensors are limited, there are partial wire deploying tasks remained in the structures. The reading distanced will be extended by developing multi-hop sensor networks to decrease the length of deployed wires in the further work. Another aspect is that the evolution of health status will induce the change of sensor locations. Finite Element Method (FEM) analysis and health diagnosis expert system will construct a closed loop to predict the movement of sensor locations.

## Acknowledgments

The work is supported by National Natural Science Foundation of China under Grant No. 50835004 and 50905131, and the State Key Laboratory of Digital Manufacturing Equipment and Technology under Grant No. DMETKF2009007. The authors thank to Director Keqin Ding of China Special Equipment Inspection and Research Center (CSEI) for his useful discussions and suggestions. The authors are also grateful to the editors and the anonymous reviewers for helpful comments.

## References

1. Matthew, J.W., Michael, V.G., Kerop, D.J., Ratneshwar, J.: Real-time wireless vibration monitoring for operational modal analysis of an integral abutment highway bridge. *Engineering Structures* 31(10), 2224–2235 (2009)
2. Pan, M.C., Li, P.C., Cheng, Y.R.: Remote online machine condition monitoring system. *Measurement* 41(8), 912–921 (2008)
3. Hameed, Z., Hong, Y.S., Cho, Y.M., Ahn, S.H., Song, C.K.: Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renewable and Sustainable Energy Reviews* 13(1), 1–39 (2009)
4. Ngai, E.W.T., Karen, K.L., Frederick, J.R., Candace, Y.Y.: RFID research: An academic literature review (1995–2005) and future research directions. *International Journal of Production Economics* 112(2), 510–520 (2008)
5. Alejandro, S.M., Esteban, E., Felipe, G., Joan, G.: Tracking of returnable packaging and transport units with active RFID in the grocery supply chain. *Computers in Industry* 60(3), 161–171 (2009)
6. Chow, K.H., Choy, K.L., Lee, W.B.: A dynamic logistics process knowledge-based system—An RFID multi-agent approach. *Knowledge-Based Systems* 20(4), 357–372 (2007)
7. Lee, C.K.M., Chan, T.M.: Development of RFID-based Reverse Logistics System. *Expert Systems with Applications* 36(5), 9299–9307 (2009)
8. Zhou, W.: RFID and item-level information visibility. *European Journal of Operational Research* 198(1), 252–258 (2009)
9. Soungho, C., Tomohiro, Y.: Application of RFID technology to prevention of collision accident with heavy equipment. *Automation in Construction* 19(3), 368–374 (2010)

10. Qiu, R.G.: RFID-enabled automation in support of factory integration. *Robotics and Computer-Integrated Manufacturing* 23, 677–683 (2007)
11. Chen, R.S., Tu, M.: Development of an agent-based system for manufacturing control and coordination with ontology and RFID technology. *Expert Systems with Applications* 36(4), 7581–7593 (2009)
12. Huang, G.Q., Zhang, Y.F., Jiang, P.Y.: RFID-based wireless manufacturing for walking-worker assembly islands with fixed-position layouts. *Robotics and Computer-Integrated Manufacturing* 23(4), 469–477 (2007)
13. Chuong, S.V., Naveen, C., Seng, W.L., Torab, T.: Radio-Mama: An RFID based business process framework for asset management. *Journal of Network and Computer Applications* (2010) (in Press) (Corrected Proof)
14. Ko, C.H.: RFID-based building maintenance system. *Automation in Construction* 18(3), 275–284 (2009)
15. Ko, C.H.: RFID 3D location sensing algorithms. *Automation in Construction* 19(5), 588–595 (2010)
16. Abarca, A., Fuente, M., Abril, J.M., García, A., Pérez-Ocón, F.: Intelligent sensor for tracking and monitoring of blood temperature and hemoderivatives used for transfusions. *Sensors and Actuators A: Physical* 152(2), 241–247 (2009)
17. Abad, E., Palacio, F., Nuin, M., Zárate, A.G., Juarros, A., Gómez, J.M., Marco, S.: RFID smart tag for traceability and cold chain monitoring of foods: Demonstration in an intercontinental fresh fish logistic chain. *Journal of Food Engineering* 93(4), 394–399 (2009)
18. Alexandru, O., Nicolae, B., Udo, W., Bauersfeld, M.L., Dirk, E., Jürgen, W.: Capacitive humidity sensors on flexible RFID labels. *Sensors and Actuators B: Chemical* 132(2), 404–410 (2008)
19. Radislav, A.P., Henri, M., William, G.M.: Position-independent chemical quantitation with passive 13.56-MHz radio frequency identification (RFID) sensors. *Talanta* 75(3), 624–628 (2008)
20. Lin, J.T., Kevin, W.W., Douglas, J., Julia, A., Mark, C., John, F.N., William, P.H.: Development of capacitive pure bending strain sensor for wireless spinal fusion monitoring. *Sensors and Actuators A: Physical* 138(2), 276–287 (2007)
21. Wang, L.C.: Enhancing construction quality inspection and management using RFID technology. *Automation in Construction* 17(4), 467–479 (2008)

# Compact Robot Vision for Moving Points

Kikuhiro Kawasue, Satoshi Nagatomo, Masahiro Nishiyama, and Yuki Kawano

Mechanical System Engineering, 1-1 Gakuen Kibanadai Nishi, Miyazaki  
889-2192, Japan  
kawasue@cc.miyazaki-u.ac.jp

**Abstract.** Circular dynamic stereoscopy (CDS) has special advantages as it enables a 3-D measurement using a single CCD camera without a cumbersome setting. Spiral streaks are recorded using this system and the size of spiral streaks directly concerns to the depth from the CCD camera. That is, the size of streaks is inversely proportional to the depth from the CCD camera and the three-dimensional information can be measured automatically by image processing technique. In this paper, robot vision system using circular dynamic stereoscopy is introduced. The arm robot is controlled using the measured data of CDS. Experimental results demonstrate the feasibility of our method.

**Keywords:** Robot, vision, image processing, three dimensional measurement, CCD, multi spot laser, calibration.

## 1 Introduction

In recent years, the industrial robots are used in various fields, such as conveyance and assembly work, at the factory. Computer vision system is used in order to raise the flexibility in these robot uses. The automatic inference of depth information has been one of primary aims of this computer vision system. Stereovision method and slit ray projection method are used in many cases [1][2]. There are some limitations in implementing these systems. In stereovision system, finding matching pairs between frames causes problematic computer processing in the case that there are several possibilities for the choice of matching points. In implementing slit ray projection method, the target must be stationary during the measurement. These systems also need enough space to secure the triangulation and some cumbersome efforts are required to execute the measurement.

In order to cope with these problems, we developed a circular dynamic stereo system (CDS)[3][4][10][11], which uses a single CCD camera. Let CCD camera move sideward against the measuring point, the amount of displacement of the measuring point appeared on the image plane is directly proportional to the amount of displacement of CCD camera and inversely proportional to the depth from the CCD camera to the measuring point. That is, the distance between CCD camera and measuring point can be estimated by the amount of movement of CCD camera and the displacement of the measuring point appeared on the image plane. This method is well known for the monocular motion stereo system [5]-[8]. Our circular dynamic stereo system realizes

this monocular motion stereo system in compact. By setting coupled mirrors at a certain angle to the optical axis of the CCD camera lens, the image of a measuring point recorded on the image plane is displaced by the corresponding amounts related to the distance between the camera and the measuring point. When the coupled mirrors are rotated physically at high speed (3,600 r/min) during the exposure of the camera, the image of a measuring point draws an annular streak. Since the size of the annular streak is inversely proportional to the distance between the camera and the measuring point, the 3D position of the measuring point can be obtained by processing the streak. This system applied to the measurement of water flow and obtained satisfactory results [3][4].

Calibration is an important task in machine vision system since it influences the measurement accuracy. Generally, calibration process is complicated and is not unified in 3D measurement system. By utilizing of multi laser spots and calibration board that moves along the z-state, a suitable calibration method for our system could be developed. The method will be introduced in the paper.

In this paper, for one of the applications of our system, the robot vision system using CDS is introduced. Measuring point is moved by the motor and the robot arm followed the movement of the point by using the information obtained by CDS.

## 2 Circular Dynamic Stereoscopy

When the CCD camera moves from left to right by a length  $\Delta x$ , as shown in Fig. 1, the position of the measurement point on the image plane appears to shift from right to left. The displacement  $\Delta u = u_1 - u_2$  on the image plane is directly proportional to the displacement  $\Delta x$  of the CCD camera and is inversely proportional to the distance  $D$  of the measurement point from the focal point of the CCD camera. More precisely,

$$\Delta u = \frac{f \cdot \Delta x}{D} . \quad (1)$$

where  $f$  is the focal length of the camera.

The distance  $D$  can then be calculated from the displacement  $\Delta u$  on the image plane and the displacement of the CCD camera  $\Delta x$ . This is the technique of monocular motion stereoscopy [5]-[9].

To realize monocular motion stereoscopy in a compact physical design, we developed the following system that adds a circular shift to the image. A simplified diagram of the proposed imaging system is shown in Fig. 2. A beam splitter and a mirror coupled to the CCD camera lens are introduced such that one image of the measurement point is directly recorded by the CCD camera through the beam splitter and a second image is displaced by traveling a path to the coupled mirror and through the beam splitter and is then recorded by the CCD camera, as shown in Fig. 2. The magnitude of the displacement of the second image is related to the distance between the beam splitter and the coupled mirror, and the displacement of the object that appears on the CCD camera is related to the distance between the CCD camera and the

measurement point. That is, the displacement  $r$  in the image is inversely proportional to the distance  $D$  between the measurement point and the camera as:

$$D = \frac{f \cdot d}{r} . \quad (2)$$

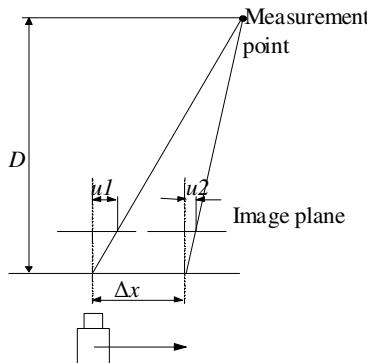
where  $f$  is the focal length of the camera and  $d$  is the magnitude of the image displacement caused by the beam splitter and the coupled mirror. When the beam splitter and the coupled mirror are rotated physically at high speed during the exposure of the CCD camera, an annular streak and a center point appears as an image for each measurement point, since the rotational shift is added to the image. Fig. 3a shows the photograph of the optical device with a beam splitter, and Fig. 3b shows the rotating optical device rotated by a non-shaft motor (Hollow motor) in front of the CCD camera. The CCD camera records the image through the hollow non-shaft motor, as shown in Fig. 3b.

Since the radius of the streak is inversely proportional to the distance of the measurement point from the camera, each annular streak contains 3D information of the measurement point. The position and the size of the annular streak in the image are related to the 3D location of the measurement point. Fig. 4a shows the multiple laser spots projected on the surface of objects without a circular shift, and Fig. 4b shows the image with circular shifts produced by the proposed system. Smaller annular streaks indicate laser spots on the surface of distant objects and larger annular streaks indicate laser spots on the surface of nearby objects.

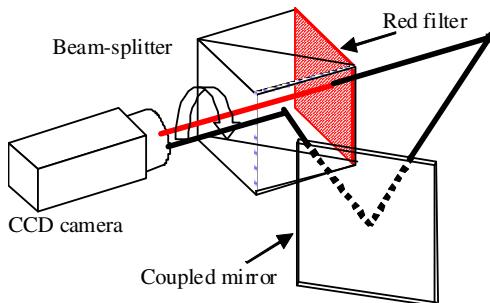
However, the spot inside the annular streak is not accurately located at the center of the annular streak. The spot is displaced slightly in the radial direction as it is displaced from the center of the image. The displacement of the point is caused by the refraction of the beam-splitter and the difference of the path-lengths between a direct path through a beam splitter and an indirect path through a combination of coupled mirrors. Therefore, the position of the spot can be compensated as follows:

$$\vec{p}' \equiv k\vec{p} . \quad (3)$$

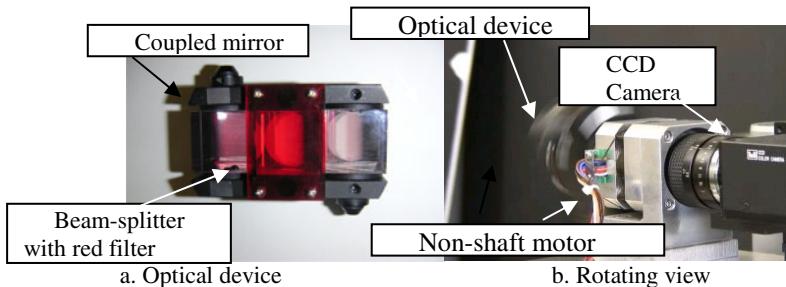
where  $\vec{p}'$  is a vector of the spot from the center in the image plane.



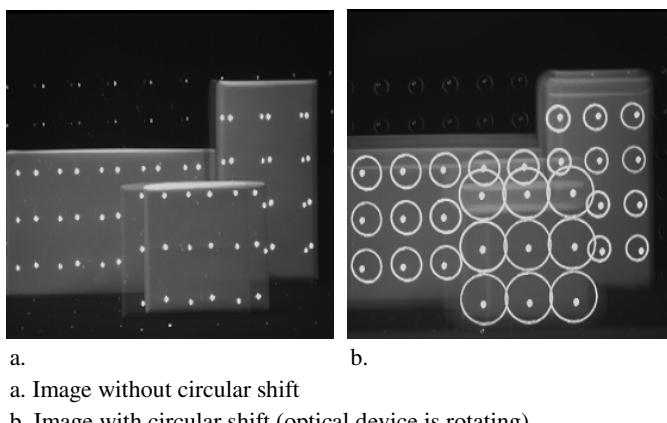
**Fig. 1.** Geometry of monocular motion stereoscopy



**Fig. 2.** Circular dynamic stereoscopy system with beam splitter



**Fig. 3.** Circular dynamic stereoscopy system



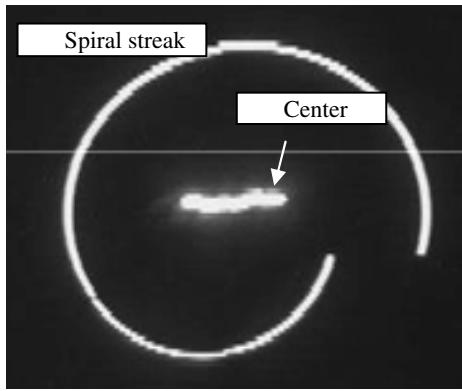
**Fig. 4.** Image obtained by circular dynamic stereoscopy

If a measurement point is moving from left to right, it traces out a spiral streak from left to right and the radius of the streak varies depending on the variation of the distance. An example of the spiral streaks is shown in Fig. 5.

Spiral streaks have a unified rotational direction according to the rotational direction of the beam-splitter and the coupled mirror. The location and the size of the spiral

streak in the image are related to the three-dimensional location of the particle, and the pitch and size variation of the streak is related to the three-dimensional velocity of the particle. The centerline of the streak is recorded simultaneously and helps computer image processing.

Image processing is relatively easy if the points being measured are relatively sparse. In cases where the points being measured are more densely grouped, for example tracer particles used in particle tracking velocimetry (PTV), the streaks in the image can easily overlap each other. In order to distinguish the center from overlapping annular streaks, the center is colored by placing a red filter in front of a beam splitter, as shown in Fig. 2. Each center can be extracted by color image processing.



**Fig. 5.** Example of a spiral streak with a center line

### 3 Calibration and Image Processing

The information required for 3D measurement of an image is the center position and the size of the annular streak. The center  $(u_c, v_c)$  and the diameter  $r_c$  of an annular streak are obtained to sub-pixel accuracy considering the pixel intensity. The parameters of the annular streak are converted to global coordinates  $(x_f, y_f, z_f)$  that are fixed on the focal point of the CCD camera by

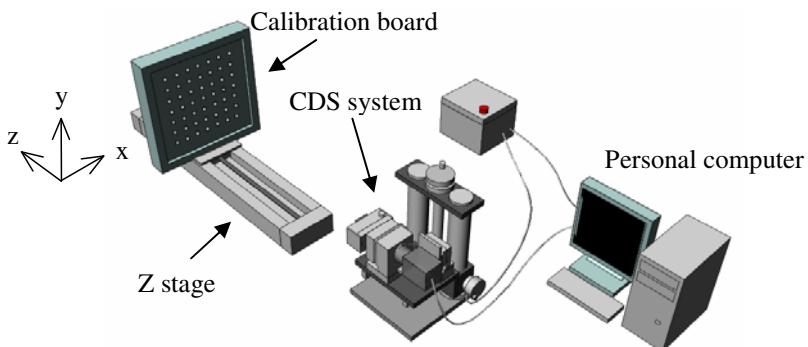
$$\begin{bmatrix} x_f \\ y_f \\ z_f \end{bmatrix} = \frac{d}{r_c} \begin{bmatrix} u_c \\ v_c \\ f \end{bmatrix}. \quad (4)$$

where  $d$  is the magnitude of the shift caused by the beam splitter and the coupled mirror,  $f$  is the focal length of the CCD camera, and  $r_c$  is the radius of the annular streak at this point. The values of  $f$  and  $d$  can be determined by sampling over two distinct non-coplanar points, the global coordinates of which are already known.

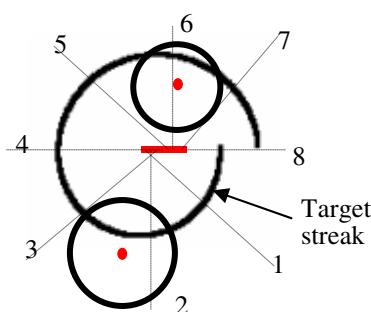
The system is setup for calibration as shown in Fig. 6. Suppose the global coordinate system is fixed on the focal point of the camera, the x-axis and y-axis are parallel to the image plane of the CCD camera, and the z-axis is along the optical axis of the

camera. A grid of dots in the xy-plane are drawn on the calibration board that can be moved along the z stage.

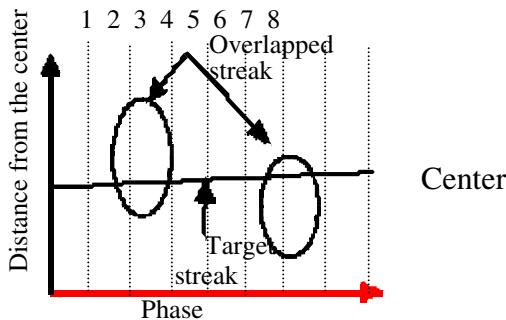
When the measurement points are moving, a spiral streak with a center line appears in the image, as shown in Fig. 7, because the movement of the measurement point is added to the rotational shift. As the center line is colored by the red filter, these streaks are easily distinguished. The rotating direction of the spiral streak is identical because it is determined by the physical rotational direction of the coupled mirror. The computer detects the centerline first by a labeling procedure, and moving information (velocity) can be estimated by analyzing the corresponding spiral streak. Fig. 8 shows the distance from the center line to the corresponding position of the spiral streak at each phase. The numbers allocated on the center line correspond between Fig. 7 and Fig. 8. Since the rotation of the coupled mirror is sufficiently faster than the movement of the measurement points and the velocity can be approximated as a constant in a short interval, the target streak in Fig. 8 becomes a straight line. The line can be analyzed using a Hough transform [12]. The distance from the center line is inversely proportional to the measurement point from the focal point according to (4) and the slope of the target streak in Fig. 8 is related to the movement information (depth direction). Image processing for Fig. 8 is executed for each streak, and the three-dimensional information is extracted.



**Fig. 6.** Setup for calibration



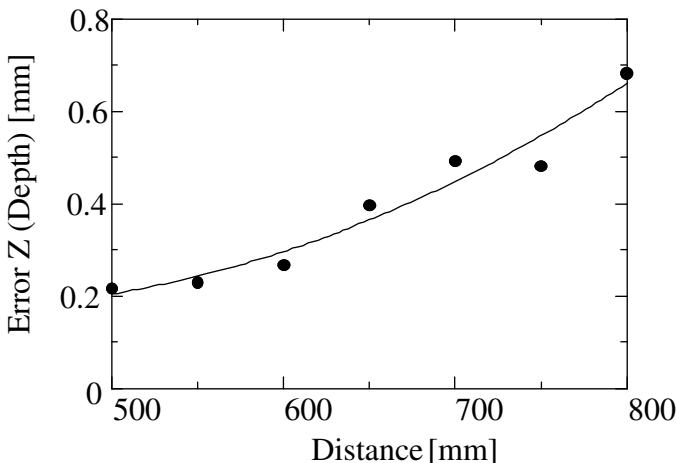
**Fig. 7.** Analysis of a spiral streak



**Fig. 8.** Distance of the spiral streak from the centerline at each phase

#### 4 Evaluation of Measurement Accuracy

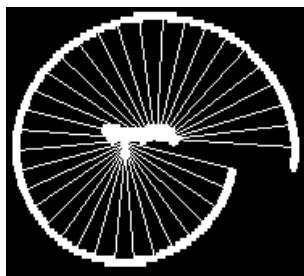
To evaluate the feasibility of the proposed system, the following experiment was conducted. A plane board was set parallel to the image plane of the CCD camera at a known distance. The measurement point was illuminated on the surface of the board by a laser spot beam. The reflected laser spot traced an annular streak in the proposed CDS system. The image data with a resolution of 1,280 pixels  $\times$  960 pixels was then stored in a computer and the position of the laser spot was calculated. The depth of



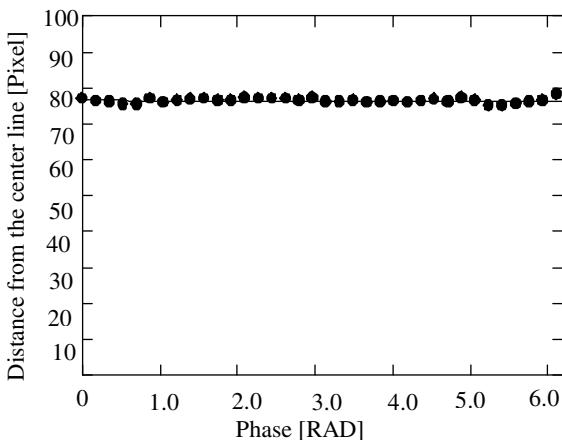
**Fig. 9.** Depth measurement accuracy for a stable point

the board was changed from 500 mm to 800 mm in 50 mm steps. The results of the experiment are shown in Fig. 9. The uncertainty in the measurement of the distance was approximately 0.1%.

In order to evaluate the measurement accuracy for a moving measurement point, a mirror positioning system was used to control the movement of a laser spot. The laser spot moved in the xy-plane. One example of the spiral streak recorded in this experiment is shown in Fig. 10, and the size analysis image explained in the previous chapter is shown in Fig. 11. The streak is divided into 36 phases and the distance from the center line to the spiral streak is measured. The lines at each phase are drawn computationally in Fig. 10. The result of the accuracy evaluation at a distance of 500 mm is shown in Fig. 12. The error increases as the measurement point moves faster but the error is within 0.3% under 200 mm/s.



**Fig. 10.** Spiral streak (moving laser)



**Fig. 11.** Distance from the center line to a spiral streak (moving laser)

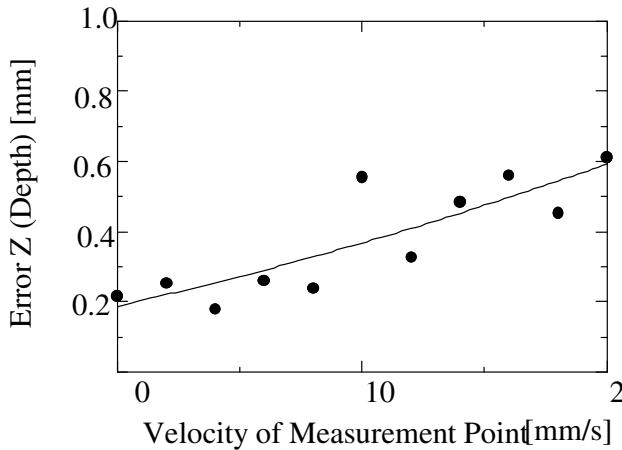


Fig. 12. Depth measurement accuracy for a moving point

## 5 Application for Robot Vision

CDS was attached to the arm robot and the information from the CDS was used for the robot control. The setup of the experiment is shown in Fig.13 and the CDS is attached to the robot hand as Fig.14. The measuring point (LED) was rotated in front of the Robot by a motor at 1 r/s and the continuous 30 frames are recorded by a CCD. Fig.15 shows the recorded spiral streak of the moving LED during one second. The spiral streak and the center line were recorded in a same image. Since the rotational frequency of the beam splitter and the coupled mirror was set at 10 r/s, 10 spirals were recorded in one rotation of the LED. The result of the analysis is shown in Fig. 16. The information of the measurement is used for the robot control.



Fig. 13. Setup for experiment

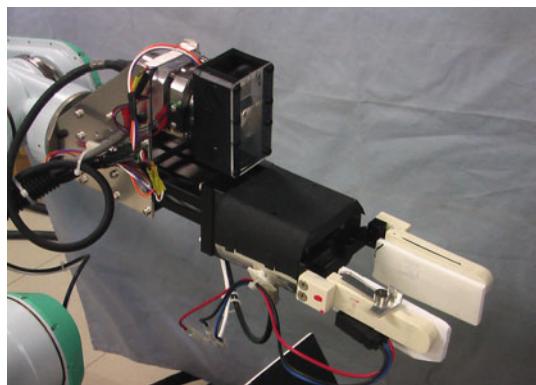
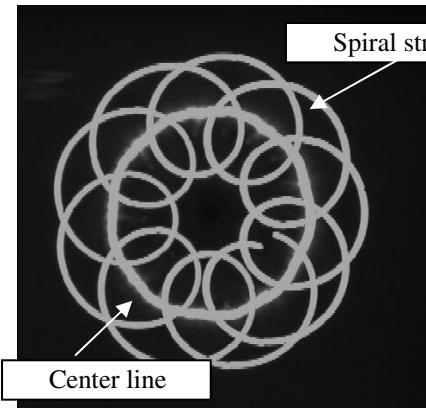
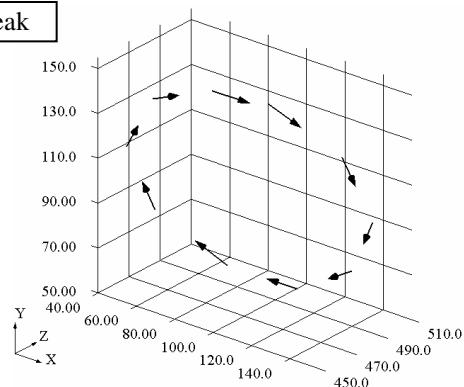


Fig. 14. Robot hand with CDS

**Fig. 15.** Recorded image**Fig. 16.** Result of analysis

## 6 Conclusion

We have introduced a new approach to obtain depth/distance information. A single camera and an image rotation apparatus record 3-D information on a single image. Spiral streaks recorded on the image plane relate directly to the 3-D positional information of the individual measuring points. The 3-D information of the measuring points is obtained by using an image processing technique.

Our system is compact and the setup is simple since it uses a single camera. The system is thus expected to be a useful tool for various fields such as a robot vision system. Experimental results demonstrate the feasibility of our system.

## References

1. Faugeras, O.: Three-dimensional computer vision, A geometric viewpoint (1996)
2. Springer-Verlag, Computer vision: Theory and industrial application (1992)
3. Kawasue, K., Shiku, O.: Circular Dynamic Stereo and its Application”, Recent advances in simulation, computational methods and soft computing. Electrical and computer engineering series, a series of reference books and textbooks, WSEAS, pp. 43–48 (2002)
4. Kawasue, K., Ohya, Y.: Circular dynamic stereo and its image processing. Advances in Multimedia, Video and Signal Processing Systems, Electrical and Computer Engineering Series, A Series of Reference Books and Textbooks, WSEAS, pp. 36–40 (2002)
5. Naavtia, R.: Depth measurement from motion stereo. Computer Graphics Image Process. 9, 203–214 (1976)
6. Williams, T.D.: Depth from camera motion in a real world scene. IEEE Trans. Pattern Anal. Machine Intell., PAMI 2, 511–516 (1980)
7. Sandini, G., Tistarelli, M.: Active tracking strategy for monocular depth inference over multiple frames. IEEE Trans. Pattern Anal. Machine Intell. 12, 13–27 (1990)
8. Black, M.J., Anandan, P.: A model for the detection of motion over time. In: Third international conference on computer vision, pp. 33–37 (1990)

9. Waxman, A.M., Shinha, S.S.: Dynamic stereo: Passive ranging to moving objects from relative image flow. *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI 8, 406–412 (1986)
10. Kawasue, K., Shiku, O., Ishimatsu, T.: Range finder using circular dynamic stereo. In: Proceedings of the 14th International Conference on Pattern Recognition, vol. 1, pp. 774–776 (1998)
11. Kawasue, K., Oya, Y.: Compact vision using circular dynamic stereoscopy with a beam splitter. *Journal of visualization* 9(2), 189–197 (2006)

# Motion Compensation Algorithm Based on Color Orientation Codes and Covariance Matching

Xuguang Zhang, Weili Ding, and Peifeng Niu

Key Lab of Industrial Computer Control Engineering of Hebei Province,  
Institute of Electrical Engineering  
Yanshan University  
Qinhuangdao, China, 066004  
xuguang.zhang78@yahoo.com.cn

**Abstract.** Regions extraction and matching are two key steps in image motion compensation. This paper applies the skills of covariance matching and information entropy to motion compensation to improve its performance. First, color orientation codes are employed to compute the information entropy of an image region and detect the sub-block regions automatically. Secondly, covariance matrices are used to match sub-blocks between current frame and previous frame. Finally, we gain the global motion parameters by affine motion model. Experimental results show that the proposed algorithm can detect and compensate global motion in indoor and outdoor environment and has outstanding result than traditional histogram matching.

**Keywords:** Motion Compensation; Color orientation codes; Information entropy; Covariance matching.

## 1 Introduction

In many fields of robot vision applications, the camera platforms may be mobile such as vehicle and airplane platform. Because of the camera movement, the images captured by the mobile platform will be changed in translation and rotation, which causes many difficulties in the applications such as image displaying, image understanding and target tracking. Therefore, detecting and compensating the global image movement is a very important task.

The common global motion compensation methods include: block matching method [1] [2], bit-plane matching [3] [4], and feature tracking method [5] [6]. Unfortunately, these methods have some disadvantages. Block matching can provide accurate results, but it is difficult to be applied to compensate the rotation motion, and it is difficult to divide the sub-blocks. The accuracy of bit-plane matching is low. Feature tracking can reduce the complexity of computation and improve the matching rate, but it deeply depends on the qualities of features extracted from two image frames, which limits its applications, such as occlusion occurred.

Regions extraction and features matching are two key steps in block matching method. If a background region has uniform color, it is difficult to match this region in another frame. Edge information is often used to judge whether a region has rich features. Unfortunately, some regions with rich edge information are unsuitable for this task such as step and edge of a street because the edge features of these regions are similar. Therefore, the region in which the distribution pattern of edge features is multiform is suitable to be extracted for background matching. Entropy is an effective tool to measure the degree of richness of information. Entropy has been applied to many image processing fields, such as similarity measure [7], entropy coding for data compression [8]. In this study, entropy is employed to measure the degree of richness of features for background regions extraction. The color and edge features are often selected to describe a background region in the previous frame. Then, this region will be matched in current frame. These features usually are integrated into a statistical model such as histogram. However, fusing several features through the histogram will increase the computation time exponentially. A covariance matrix (also named covariance descriptor) has been proposed to describe a target, which can fuse multiple features [9]. Covariance matrices had been successfully utilized in object tracking [10], human and license detecting [11], [12] and face recognition [13].

This paper focus on the regions extraction and matching to achieve an effective approach of motion compensation. First, we extract the regions with rich features from the reference frame. Secondly, block matching is used to match these blocks in consecutive frames. Finally, we calculate and compensate the global motion of the current image using the locations of corresponding regions.

The remainder of the paper is organized as the follows: Section 2 presents the principle of extracting the background regions. Section 3 introduces the sub-block matching method by the descriptor of covariance matrices. In section 4, the global motion is computed and compensated. In section 5, the proposed algorithm is evaluated on real video sequences. Finally, Section 6 concludes the paper.

## 2 Extracting Sub-block Regions

In order to obtain a stable matching result, the choice of sub-block is essential. The feature richness blocks could help us to get an accurate block matching. Based on the idea of physics, feature richness regions should have the higher entropy [14], so we can consider the entropy to determine whether the sub-block is rich enough. In this study, we use the entropy of color orientation codes to extract the feature richness sub-blocks.

### 2.1 Color Orientation Codes

For a color image, any  $(x, y)$  can be described by a vector  $\mathbf{c}(x, y) = \begin{bmatrix} R(x, y) \\ G(x, y) \\ B(x, y) \end{bmatrix}$  in the RGB color space. Let  $r, g$  and  $b$  are the unit vectors along the R, G and B axis of RGB color space, a pair of vector can be defined below according to [15]:

$$\mathbf{u} = \frac{\partial R}{\partial x} \mathbf{r} + \frac{\partial G}{\partial x} \mathbf{g} + \frac{\partial B}{\partial x} \mathbf{b} \quad (1)$$

and

$$\mathbf{v} = \frac{\partial R}{\partial y} \mathbf{r} + \frac{\partial G}{\partial y} \mathbf{g} + \frac{\partial B}{\partial y} \mathbf{b} \quad (2)$$

The gradient angle around point  $(x, y)$  can be obtained,

$$\theta_{xy} = \frac{1}{2} \tan^{-1} \left[ \frac{2g_{xy}}{g_{xx} - g_{yy}} \right] \quad (3)$$

where  $g_{xx}$ ,  $g_{yy}$  and  $g_{xy}$  are the dot products of vector  $\mathbf{u}$ ,  $\mathbf{v}$ :

$$g_{xx} = \mathbf{u} \cdot \mathbf{u} = \left| \frac{\partial R}{\partial x} \right|^2 + \left| \frac{\partial G}{\partial x} \right|^2 + \left| \frac{\partial B}{\partial x} \right|^2 \quad (4)$$

$$g_{yy} = \mathbf{v} \cdot \mathbf{v} = \left| \frac{\partial R}{\partial y} \right|^2 + \left| \frac{\partial G}{\partial y} \right|^2 + \left| \frac{\partial B}{\partial y} \right|^2 \quad (5)$$

$$g_{xy} = \mathbf{u} \cdot \mathbf{v} = \frac{\partial R}{\partial x} \frac{\partial R}{\partial y} + \frac{\partial G}{\partial x} \frac{\partial G}{\partial y} + \frac{\partial B}{\partial x} \frac{\partial B}{\partial y} \quad (6)$$

Given an interval of the gradient angle  $\Delta\theta$ , the color orientation codes for a pixel location  $(x, y)$   $c_{xy}$  can be represented as:

$$c_{xy} = \begin{cases} \left[ \frac{\theta_{xy}}{\Delta\theta} \right] & g > \Gamma \\ m & otherwise \end{cases} \quad (7)$$

where the parameter is denoted by  $g = \left| \frac{\partial R}{\partial x} \right| + \left| \frac{\partial G}{\partial x} \right| + \left| \frac{\partial B}{\partial x} \right| + \left| \frac{\partial R}{\partial y} \right| + \left| \frac{\partial G}{\partial y} \right| + \left| \frac{\partial B}{\partial y} \right|$ .

The orientation codes are quantified as  $m-1$  ranks.

Though the feature of the orientation codes represents edge information of an image region, rich edge information could not be extracted via the extended orientation codes for low contrast regions. A threshold value  $\Gamma$  is defined to evaluate the region features of the orientation codes base on certain contrast. In this paper, the color orientation codes are defined as an integer range from 0 to 15 (i.e.,  $\Delta\theta = \pi/32$  is set as

radians). The threshold value  $\Gamma$  is set as 60 by experience. It is to say that any pixel whose feature represented by the color orientation codes is less than 60 will be discarded.

## 2.2 Evaluating Feature Richness

In order to select the reasonable sub-block regions, it is necessary to evaluate the feature of any sub-block richness. Color orientation codes are employed to represent the edge information of an image grid region, the richness of a region is evaluated via entropy for the features of color orientation codes. The higher the entropy the more unordered distribution of the color orientation codes, the richer edge information is included in this sub-block region.

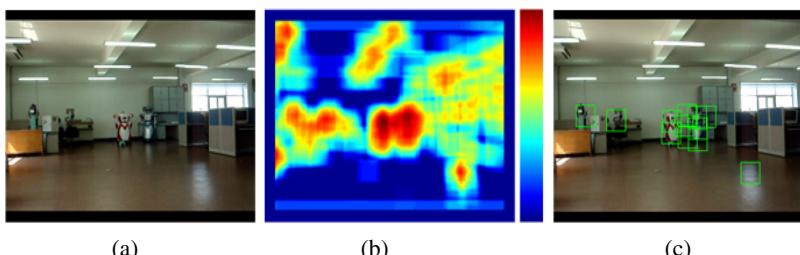
For an image region, the frequency of appearance of  $i$ th orientation code can be described by  $f(i)$  ( $i=0,1,\dots,m-1$ ). The number of effective orientation codes  $n$  is calculated to satisfy the relation  $g > \Gamma$  as shown in equation 7, the normality frequency of each orientation codes can be calculated by formula (8):

$$p(i) = \frac{f(i)}{n} \quad (i = 0, 1, \dots, m-1) \quad (8)$$

Furthermore, the entropy of a grid region is obtained as,

$$E = \sum_{i=0}^{m-1} p(i) \log_2[p(i)] \quad (9)$$

Figure 1 shows the result of entropy filter, in which the size of a sub-block is set as  $24 \times 24$  pixels; the entropy of each region reflects whether the features in this sub-block are rich. We select 12 sub-blocks which have the maximum entropy to match in next frame. In order to avoid the overlapping phenomenon of neighboring regions, we empty the entropy in the vicinity after getting one sub-block. Finally the sub-blocks we chosen are shown in Figure 1.



**Fig. 1.** Extraction of sub-blocks: (a) original image; (b) the result of entropy filter of (a); (c) extracted background regions

### 3 The Sub-blocks Matching

The stability of the calculation of global motion depends on the accuracy of sub-block matching. Using a single feature for block matching is too difficult to maintain the stability in the whole image sequence. Fusing multiple features can provide more stable matching. The covariance matrices descriptor proposed by O.Tuzel has played a good performance in image matching [9]. It also has been applied in the field of target tracking and license plate recognition [10] [12]. In this paper, covariance matrix is used as the feature model for sub-block matching.

#### 3.1 Feature Vectors Extraction

For an image region, the color feature of a pixel  $(x, y)$  can be described as  $R(x, y)$ ,  $G(x, y)$  and  $B(x, y)$ , the  $I_x(x, y)$  and  $I_y(x, y)$  are the gradient features of the image region, the  $I_{xy}(x, y)$  is the second-order gradient information, the  $d(x, y)$  is a spatial feature.  $d(x, y)$  can be calculated by formula (10):

$$d(x, y) = \sqrt{(x'^2 + y'^2)}, (x', y') = (x - x_0, y - y_0) \quad (10)$$

where  $(x_0, y_0)$  is the center coordinates of the region,  $(x', y')$  are the relative coordinates. We use color, gradient, second-order gradient and distance information to form a feature vector  $f_k$ :

$$f_k = [R(x, y) \ G(x, y) \ B(x, y) \ d(x, y) \ I_x(x, y) \ I_y(x, y) \ I_{xy}(x, y)] \quad (11)$$

#### 3.2 Construction of Covariance Matrix

If a background R is represented by  $M \times N$  rectangular regions, the feature vector of individual regions can be represented by covariance matrix  $\mathbf{C}_R$  as below,

$$\mathbf{C}_R = \frac{1}{MN} \sum_{k=1}^{MN} (\mathbf{f}_k - \boldsymbol{\mu}_R)(\mathbf{f}_k - \boldsymbol{\mu}_R)^T \quad (12)$$

where  $\boldsymbol{\mu}_R$  is the vector of the means of the corresponding features for the points in the region R.

#### 3.3 Distance Measure of Covariance Matrices

In order to match the sub-blocks between frames, the distances between covariance matrices are needed to be measured in two consecutive frames. However, the covariance matrices do not lie on Euclidean space, so subtraction of two matrices can not be used to measure the distance. The distance between the covariance matrices can be obtained by Förstner's contribution [16]. The sum of the squared logarithms of the

generalized eigenvalues is defined to compute the dissimilarity between two covariance matrices, the measure is calculated as

$$\rho(\mathbf{C}_i, \mathbf{C}_j) = \sqrt{\sum_{k=1}^d \ln^2 \lambda_k(\mathbf{C}_i, \mathbf{C}_j)} \quad (13)$$

1) where  $\{\lambda_k(\mathbf{C}_i, \mathbf{C}_j)\}$  are the generalized eigenvalues of  $\mathbf{C}_i$  and  $\mathbf{C}_j$ , computed from  $|\lambda\mathbf{C}_i - \mathbf{C}_j| = 0$ .

The process of block matching can be described as searching the smallest distance between sub-block's covariance matrix and the candidate region's covariance matrix.

## 4 Compensation of Global Motion

After obtaining the matching locations of 12 regions, the global motion parameters of the image are needed to be calculated by these locations to compensate the global motion of the current frame. Suppose the camera rotating around the optical axis, the variation between two successive frames can be described by the affine transformation model [17]:

$$\begin{cases} x_1 = a_{11}x_0 + a_{21}y_0 + b_1 \\ y_1 = a_{12}x_0 + a_{22}y_0 + b_2 \end{cases} \quad (14)$$

where  $(x_0, y_0)$  and,  $(x_1, y_1)$  are two corresponding points in two consecutive image frames,  $a_{11}, a_{21}, a_{12}, a_{22}, b_1, b_2$  are the parameters of an affine transform. In this paper, we only compensate the rotation and translation motion. Therefore, the scaling parameter of the affine model is set as 1.

Based on the principle of affine transformation, every three corresponding points can work out a set of transformation parameters. Then, 12 groups matching points can work out 220 groups parameters of affine transformation. Because most of the sub-blocks can be matched accurately by covariance matrices, the frequency of correct parameters is the highest in the 220 set of parameters. So, we can assign a parameter interval to devide these parameter into several levels. Then we compute the frequency of each level for the 220 set of parameters, using the average of the highest frequency parameters as the final result of global motion.

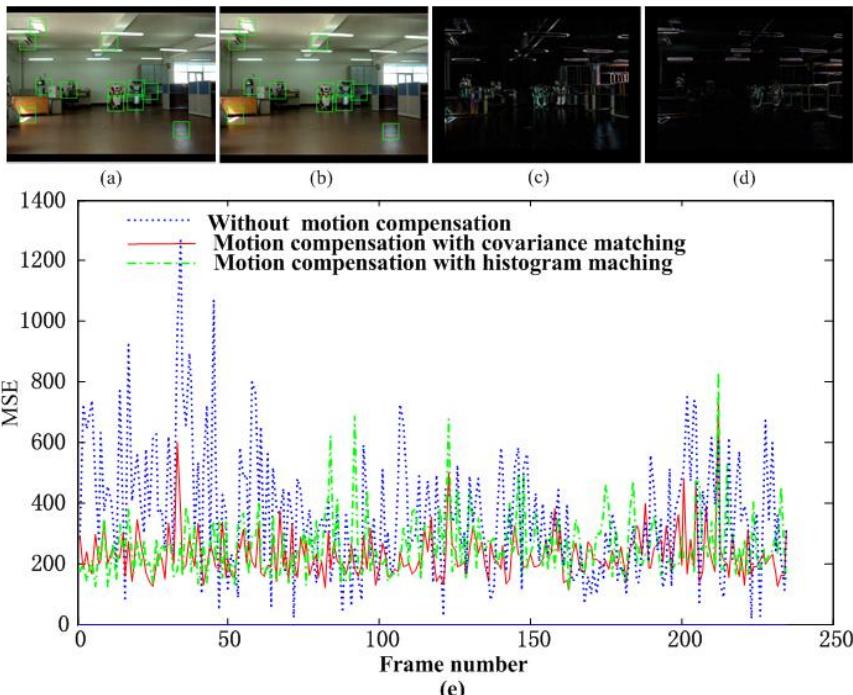
## 5 Experimental Results

The proposed algorithm had been evaluated in a variety of indoor and outdoor environment, and the image size is 320×240 pixels, the searching window is set as 20×20 pixels for covariance matching. To evaluate the effect of the proposed motion compensation algorithm, we calculate the Mean Square Error (MSE) of current frame and previous frame as fomular (15).

$$MSE = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H [(R_k(x,y) - R_{k-1}(x,y))^2 + (G_k(x,y) - G_{k-1}(x,y))^2 + (B_k(x,y) - B_{k-1}(x,y))^2] \quad (15)$$

where  $R_k(x, y)$ ,  $G_k(x, y)$  and  $B_k(x, y)$  are the color values of the k frame,  $R_{k-1}(x, y)$ ,  $G_{k-1}(x, y)$  and  $B_{k-1}(x, y)$  are the color values of the k-1 frame. W, H is the width and height of the image. Because the image boundary exists the new entry and exit part, we reduce the W and H by 20 pixels respectively during the calculation of the MSE. The smaller the MSE value is the better effect of the motion compensation gets.

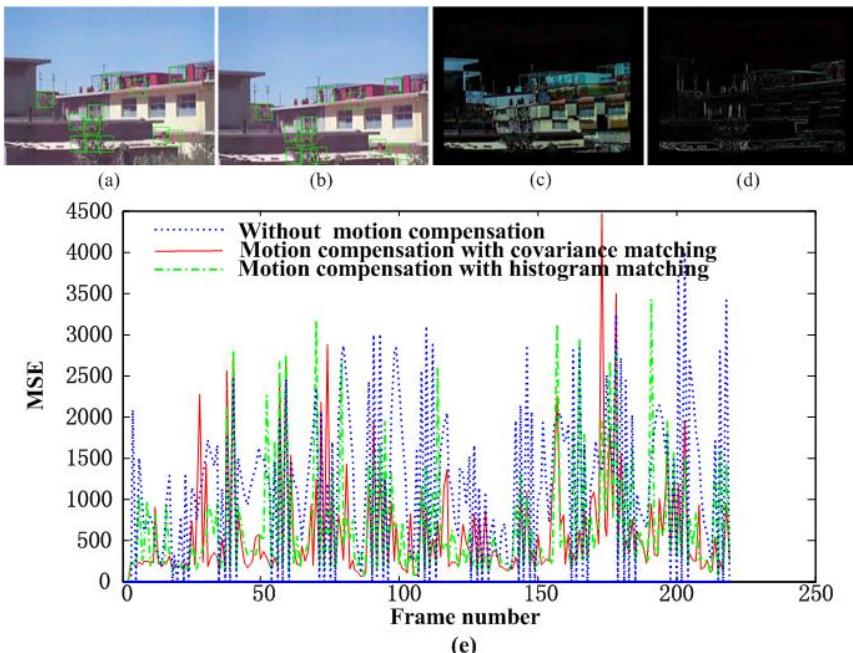
We choose two sequences from indoor and outdoor environment respectively to show the performance of our method. Sequence 1 is captured from indoor environment, a total of 235 frames. There are translation and rotation motion in this sequence. Figure 2 shows the effect of motion compensation: (a) shows the sub-blocks extracted in frame 65, the sub-block regions have rich features; (b) shows the matching results in frame 66; (c) shows the subtraction of frame 65 and frame 66 before motion compensation; (d) shows the subtraction of frame 65 and frame 66 after motion



**Fig. 2.** Comparison of MSE between current frame and reference frame of sequence 1: (a) frame 65 and the extracted regions; (b) frame 66 and the matching results; (c) subtraction of frame 65 from frame 66 before motion compensation; (d) subtraction of frame 65 from frame 66 after motion compensation.

compensation. It is clear that after motion compensation the differences between images are reduced significantly; (e) compares the MSE under three situations: motion compensation with covariance matching, motion compensation with histogram matching and without motion compensation. It is obviously that after motion compensation, the MSE is reduced considerably, and the proposed algorithm is excellent than histogram matching.

Sequence 2 is obtained from outdoor environment by a hand-held camera. In order to display an obvious effect of motion compensation we shake the camera dramatically. There are 219 frames in sequence 2. From Fig. 3(a) we can see the sub-block regions extracted from frame 98 are all with rich features which ensure us to get a good matching result. The effect of motion compensation is displayed by comparing the subtraction between frame 99 and frame 98 with or without motion compensation just as Fig. 3 (d) and Fig. 3 (c) shown. It is clear that after motion compensation the differences between two consecutive frames are reduced significantly. The MSE of these two frames with and without motion compensation is compared in Fig. 3 (e). It is obviously that after motion compensation, the MSE is reduced considerably. We also compare the MSE after motion compensation by covariance matching and histogram matching. Results show that the effect of proposed algorithm outweighs histogram matching.



**Fig. 3.** Comparison of MSE between current frame and reference frame of sequence 2: (a) frame 98 and the extracted regions; (b) frame 99 and the matching results; (c) subtraction of frame 98 from frame 99 before motion compensation; (d) subtraction of frame 98 from frame 99 after motion compensation.

## 6 Concluding Remarks

In order to achieve an effective performance of motion compensation, this study applies the entropy of orientation codes and covariance matrices to sub-blocks extraction and matching. Color orientation codes are used to compute the entropy of an image region. The sub-block regions are selected automatically by judging the value of the entropy. Then, covariance matrices are employed to match the sub-blocks between two consecutive frames. Experimental results show that the proposed algorithm can compensate the image global motion well under indoor and outdoor environment. In the future work, we plan to apply the method of random sampling to improve the computing speed of covariance matching.

**Acknowledgments.** This research was supported by National Natural Science Foundation of China (Grant No. 60805045, 60774028). The authors also thank Prof. Jianbin Jiao for providing the image sequences.

## References

1. Vella, F., Castorina, A., Mancuso, M.: Digital image stabilization by adaptive block motion vectors filtering. *IEEE Trans. on Consumer Electronics* 48(3), 796–801 (2002)
2. Morimoto, C., Chellappa, R.: Fast electronic digital image stabilization. In: *IEEE ICPR 1996*, pp. 284–288 (1996)
3. Sung, J.K., Sung, H.L., Seung, W.J.: Fast digital image stabilize based on gray-coded bit-plane matching. *IEEE Trans. on Consumer Electronics* 45(3), 598–603 (1999)
4. Pourreza, H.R., Rahmati, M., Behazin, F.: Weighted Multiple Bit-Plane Matching, a Simple and Efficient Matching Criterion for Electronic Digital Image Stabilizer Application. In: *WSCG 2002*, pp. 37–40 (2002)
5. Shi, J.B., Tomasi, C.: Good points to track. In: *IEEE CVPR*, pp. 593–600 (1994)
6. Censi, A., Fusielo, A., Roberto, V.: Image stabilization by featurestracking. Technical report, University of Udine, Machine Vision Laboratory, Dept. of Mathematics and Informatics (1998)
7. Gholipour, A., Kehtarnavaz, N., Yousefi, S., Gopinath, K., Briggs, R.: Symmetric deformable image registration via optimization of information theoretic measures. *Image and Vision Computing* 28(6), 965–975 (2010)
8. Cagnazzo, M., Antonini, M., Barlaud, M.: Mutual information-based context quantization. *Signal Processing: Image Communication* 25(1), 64–74 (2010)
9. Tuzel, O., Porikli, F., Meer, P.: Region covariance: A fast descriptor for detection and classification. In: *European Conf. Computer Vision*, Graz, Austria, vol. 2, pp. 589–600 (2006)
10. Porikli, F., Tuzel, O., Meer, P.: Covariance tracking using model update based on Lie algebra. In: *IEEE Conf. Computer Vision and Pattern Recognition*, New York, vol. 1, pp. 728–735 (2006)
11. Tuzel, O., Porikli, F., Meer, P.: Pedestrian detection via classification on Riemannian manifolds. *IEEE Trans. Pattern Anal. Machine Intelligence* 30, 1713–1727 (2008)
12. Porikli, F., Kocak, T.: Robust License Plate Detection Using Covariance Descriptor in a Neural Network Framework. In: *IEEE Conf. Advanced Video and Signal Based Surveillance*, p. 107 (2006)

13. Pang, Y., Yuan, Y., Li, X.: Gabor-Based Region Covariance Matrices for Face Recognition. *IEEE Trans. Circuits and System for Video Technology* 18(7), 989–993 (2008)
14. Domae, Y., Kaneko, S., Tanaka, T.: Robust tracking based on orientation code matching under irregular conditions. In: Proc. of SPIE, pp. 60510S-1–60510S-9 (2005)
15. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 2nd edn., pp. 335–339. Prentice Hall, Englewood Cliffs (2002)
16. Förstner, W., Moonen, B.: A metric for covariance matrices. Technical report, Dept. of Geodesy and Geoinformatics, Stuttgart University (1999)
17. Kang, E., Cohen, I., Medioni, G.: Robust Affine Motion Estimation in Joint Image Space Using Tensor Voting. In: 16th International Conference on Pattern Recognition (2002)

# An Advanced Algorithm for Illumination-Based Synchronization of High-Speed Vision Sensors in Dynamic Scenes

Lei Hou, Shingo Kagami, and Koichi Hashimoto

Graduate School of Information Sciences, Tohoku University,  
6-6-01 Aramaki Aza Aoba, Aoba-ku, Sendai 980-8579, Japan

{lei,swk,koichi}@ic.is.tohoku.ac.jp

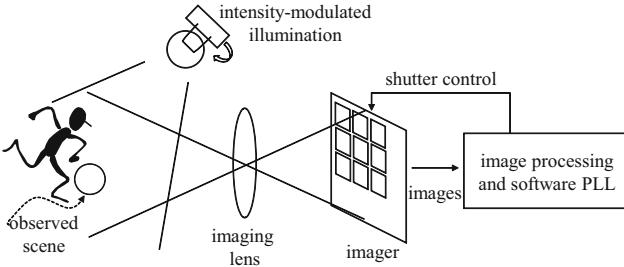
<http://www.ic.is.tohoku.ac.jp/en/>

**Abstract.** To acquire images of dynamic scenes from multiple points of view simultaneously, the acquisition time of vision sensors should be synchronized. In this paper, an advanced illumination-based synchronization method derived from the phase-locked loop (PLL) algorithm is proposed and evaluated. To remove the dependency of system behavior on the amplitude of the illumination, which can be affected by moving objects or the positional relation of the illumination and objects, the feedback amount of the PLL system is normalized per frame by the estimated amplitude of the reference signal to generate stable synchronization even in highly dynamic scenes. Both simulated results and real world experiments demonstrated successful synchronization that a 1,000-Hz frame rate vision sensor was successfully synchronized to both direct and indirect illumination with only 28- $\mu$ s peak-to-peak jitters.

**Keywords:** Robot Vision; Vision Chip; Camera Synchronization; Phase-locked Loop; Visible Light Communication; Visual Servo; Intensity Modulation; Signal Normalization; Quadrature Detection.

## 1 Introduction

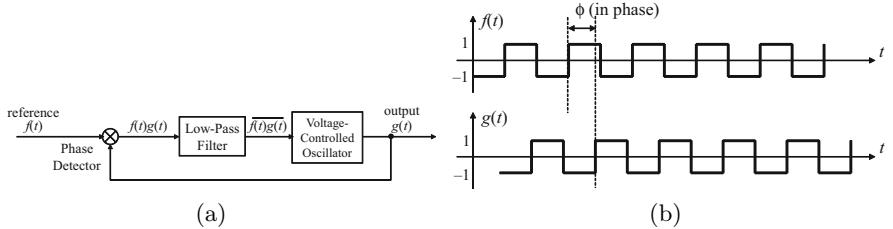
When multiple vision sensors are used to acquire images of a scene from multiple points of view to achieve, for example, binocular robot; cooperative tracking, wide area monitoring or 3D motion measurement, the image sequences given by the sensors should be synchronized. In some applications, virtual synchronization, eg. [1], in which interpolation and prediction between the frames of unsynchronized cameras are used, can be an alternative, but real synchronization is apparently advantageous when, for example, the motion of target objects



**Fig. 1.** Conceptual diagram of the proposed illumination-based synchronization

is fast and random, and/or highly precise position information is required. Instead of dedicated synchronization wires, some systems allow synchronization through standard electronic buses used for image transfer such as IEEE 1394 [2] and Ethernet [3,4]. These systems bring higher flexibility, but they still require wired connections and are unsuitable for wireless vision sensor networks. The principal difficulty in time synchronization of wireless network systems lies in nondeterminism in wireless media access time [5]. Due to this nondeterminism, it is difficult to make certain when a synchronization packet started to propagate from the sender. RBS [6] introduced a receiver-receiver synchronization scheme to remove the effect of the sender nondeterminism, but requires many message exchanges between receivers to achieve high precision. TPSN [7] and FTSP [8] suppress this nondeterminism by timestamping at the media access control (MAC) layer, but they inherently require special MAC implementations. On the other hand, illumination is always needed in visual sensing unless the observed target itself is a light emitter, and can be a common source of a synchronization reference for all the vision sensors. we proposed an algorithm [9] based on the phase-locked loop (PLL) technology [10,11]. Fig. 1 illustrates the conceptual diagram of the proposed system.

However, drastic amplitude fluctuation of reference signal caused by, for example, waving objects within the scene, or variation of relative distances between illumination and objects, can lead to unsuccessful synchronization. The dependency removal strategy lies in signal normalization of the amplitude of reference signal by estimating the real time amplitude of reference signal. This signal normalization algorithm has been originally proposed in [12], where negative feedback control with the estimation of reference amplitude was done every 4 frames. Shortly afterwards, an advanced algorithm with a moving window enabling per-frame feedback control was proposed to compensate for the large jitters and the unexpected instability when the real world experiment was carried out. In



**Fig. 2.** PLLs and its reference signal and output signal, (a) block diagram of a PLL, (b) reference signal  $f(t)$  and output signal  $g(t)$

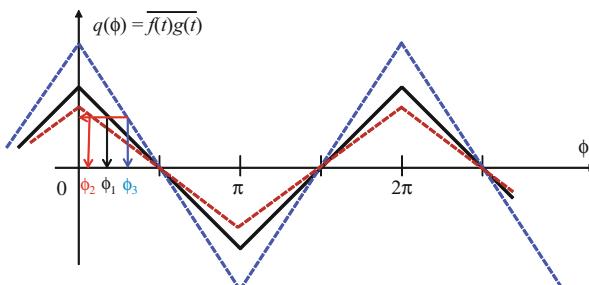
this paper, comprehensive performance evaluation and convincing experimental results will confirm the advantages and practicability of the proposed system.

## 2 Synchronization Algorithm

### 2.1 Imager-Based PLL Theory

For better understanding of the proposed method, we first describe an overview of the PLL theory. A PLL is a system synchronizing an output signal with a reference or input signal in frequency as well as in phase. In particular, our proposal is based on the binary-valued continuous-time PLL, where the reference and output signals are square waves. Successful synchronization has been achieved by the authors in static environment. In Fig. 2 (a), which illustrates a block diagram of a typical PLL. In Fig. 2 (b), both of  $f(t)$  and  $g(t)$  are square waves that alternately take two values, namely 1 and  $-1$ , with 50-% duty ratio. The product  $f(t)g(t)$  of the reference signal  $f(t)$  and the output signal  $g(t)$  is computed by the phase detector.

As can be seen in the black plot among Fig. 3, when  $q(\phi)$  is positive, the output signal  $g(t)$  is varied in the direction such that its phase will be lagged



**Fig. 3.** Fluctuation of the amplitude causes the confusion of phase error

relatively from that of the reference  $f(t)$ . When negative, of course,  $g(t)$  is varied so that its phase is advanced. As long as the feedback characteristic of this loop is properly designed, the system will be converged to the stable equilibrium point  $q(\phi) = \pi/2$ . This is called the locked state.

We formulate the time correlation of  $f(t)$  and  $g(t)$  as

$$\frac{1}{T} \int_{t-T}^t f(\tau)g(\tau) d\tau \quad (1)$$

where  $T$  denotes the period of the correlation time window, which is sufficiently longer than the vision frame time.

The main difference between this formulation and the standard theory described is that  $f(t)$  takes 1 and 0 instead of 1 and  $-1$ . This difference comes from the fact that light brightness cannot be negative. Nevertheless, the PLL system will behave in the same way as the standard one as long as  $g(t)$  has 50% duty ratio, because

$$\frac{1}{T} \int_{t-T}^t f(\tau)g(\tau)d\tau = \frac{1}{2T} \int_{t-T}^t f'(\tau)g(\tau)d\tau \quad (2)$$

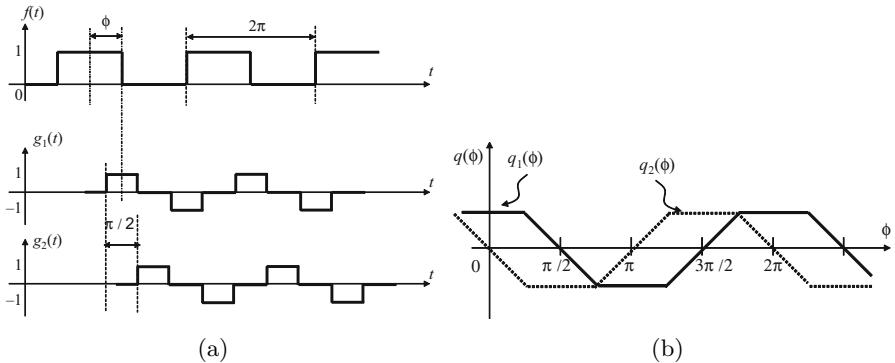
where  $f'(t)$  is the square wave with the values 1 and  $-1$ . This is a common technique used when correlation of optical signals is computed [13][4].

Unfortunately, we still cannot implement this time correlation computation as it is formulated. Most vision sensors output an image, which is the result of time integration of incident light over the frame time, just frame by frame, and therefore the reference value  $f(t)$  or the product  $f(t)g(t)$  at any arbitrary time instant is unavailable. However, by considering that  $g(t)$  is a constant during one frame period, we can obtain the time correlation as

$$\frac{1}{T} \int_{t-T}^t f(\tau)g(\tau)d\tau = \frac{1}{T} \sum_i (-1)^{i-1} F[i] \quad (3)$$

where  $i$  is the frame number index and  $F(i)$  is the sum of the pixel values obtained within the frame  $i$ .

Since the phase error is linearly proportional to the amplitude of illumination, the fluctuation of the illumination amplitude may confuse the phase error relation temporally, either advanced like  $q_2(\phi)$  or lagged like  $q_3(\phi)$ , shown in the red and the blue plot respectively in Fig. 3. Consequently, even if the correlation value stays the same, the shifts of phase error may be caused by the amplitude fluctuation. Therefore, for that very frame, synchronization will be provisionally lost due to the dysfunction of feedback gain, although the real time synchronization will recover once the fluctuation disappear. This imperfection greatly limits the practical application of the proposed technique.



**Fig. 4.** (a) Reference signal  $f(t)$ , output signal  $g_1(t)$ , and its quadrature counterpart  $g_2(t)$ , (b) The relation between time correlations and phase difference

## 2.2 Signal Normalization of the Amplitude of Reference Signal

Although a novel illumination-based synchronization method has been proposed and evaluated in [15]. To realize its practical application and to remove the dependency due to the amplitude fluctuation of reference signal, we propose an advanced algorithm in which the amplitude of the reference signal is normalized by the estimated amplitude of the reference signal every frame. This algorithm is designated as the *signal normalization algorithm* in the rest of this paper. In this signal normalization algorithm, the output signal  $g_1(t)$  has a quadrature counterpart  $g_2(t)$ , which has a 90-degree phase difference with  $g_1(t)$  and hence the designation of this method, as shown in Fig. 4(a). Given that most vision sensors arbitrarily output an image just frame by frame, the frequency of  $f(t)$  is doubled, therefore the phase quadrature can be spontaneously generated by the lag of a whole frame time.  $\overline{f(t)g_1(t)}$  and  $\overline{f(t)g_2(t)}$  are then computed respectively to remove the confusing effect of the amplitude fluctuation of  $f(t)$ .

The relations between  $q_1(\phi) \equiv \overline{f(t)g_1(t)}$ ,  $q_2(\phi) \equiv \overline{f(t)g_2(t)}$ , and  $\phi$  are shown in Fig. 4(b). The former has a  $\pi/2$  phase shift compared to the latter. Mathematically, for a certain frame, the maximum between  $|q_1(\phi)|$  and  $|q_2(\phi)|$  turns out to be a fixed value with respect to  $\phi$  and is proportional to the illumination amplitude. By updating this maximum correlation per frame, and dividing the feedback amount by this value, fluctuating amplitudes of the reference signal can be tolerated in real time.

## 2.3 Feedback Algorithm

The reference signal  $f(t)$  is injected into the PLL loops. In each output frame, after an image is acquired and  $F(i)$  is computed,  $F(i)$  is stored in one of these

four variables,  $E_1$ ,  $E_2$ ,  $H_1$ , and  $H_2$  so that:

$$E_1 = F[i], \text{ when } i = 0 \bmod 4 \quad (4)$$

$$E_2 = F[i], \text{ when } i = 1 \bmod 4 \quad (5)$$

$$H_1 = F[i], \text{ when } i = 2 \bmod 4 \quad (6)$$

$$H_2 = F[i], \text{ when } i = 3 \bmod 4 \quad (7)$$

Using these variables, in each frame the discrete-time low-pass filters to give the time correlations  $q_1[i]$  and  $q_2[i]$  at frame  $i$  are implemented as simple first-order recursive filters with a four-frame moving window

$$q_1[i] = a_1 \cdot q_1[i - 1] + a_2 \cdot (E_1 - H_1) \quad (8)$$

$$q_2[i] = a_1 \cdot q_2[i - 1] + a_2 \cdot (E_2 - H_2) \quad (9)$$

where  $q_1[i]$  and  $q_2[i]$  are achieved at the end of every frame. The optimal coefficients  $a_1$  and  $a_2$  can be chosen by analyzing the convergence time and the jitters, as well as the undershoot and overshoot of system before convergence, as discussed in the previous work.

Each frame of the simulated vision sensor consists of an integration period and a non-integration period. The integration period is fixed to 0.8 ms, and the length of the non-integration period  $\tau_{\text{nonint}}$  is changed every frames in accordance with the time correlation  $q_1$  and the normalizer  $\max(|q_1|, |q_2|)$

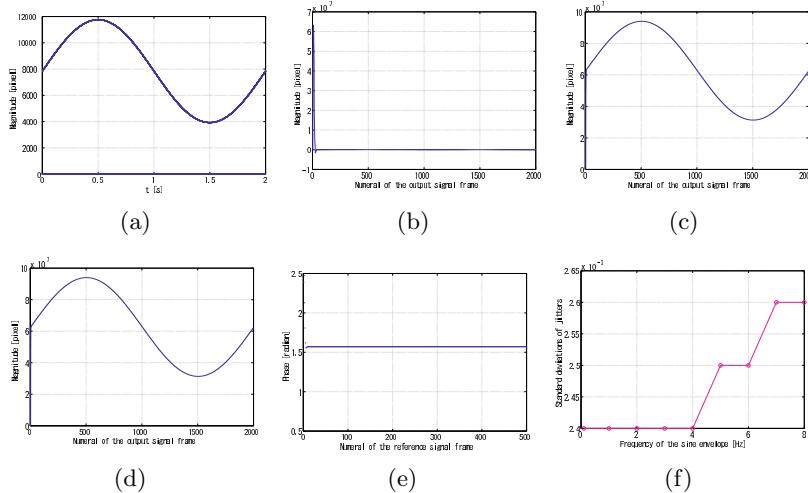
$$\tau_{\text{nonint}}[i] = \tau_0 + \frac{G \cdot q_1[i]}{\max(|q_1[i]|, |q_2[i]|)} \quad (10)$$

where  $\tau_0$  is a constant set to 0.2 ms according to the system property, and the actual value of gain  $G$  is explored in the simulation section. The resolution of adjustment of  $\tau_{\text{nonint}}$  is 100 ns, which is the instruction cycle of the system used in the experiment chapter.

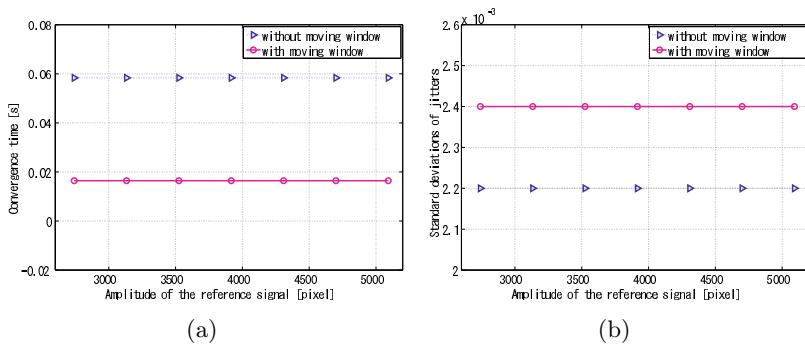
### 3 Simulation and Performance Evaluation

This chapter presents the simulated results to analyze the system behavior. The purpose of the simulations is twofold. Firstly, we aim at exploring feasible parameters for the system while evaluating the synchronization performance. We model visual measurement using a high-speed vision sensor with 1,000-Hz frame rate and  $64 \times 64$  pixels, which requires 250-Hz modulated illumination according to the central frequency of normalization algorithm. The frame rate and the number of pixels are decided so that they are equivalent to those of the vision sensor used in the real world experiments in the next chapter.

Fig. 5(a) shows the reference signal with a slow sinusoidal envelope curve, where the gain  $G$  was set to 550. Fig. 5(b) shows the correlation of the output



**Fig. 5.** Simulation results with a slow sinusoidal envelope curve, (a)  $f(t)$ , (b)  $q_1(t)$ , (c)  $q_2(t)$ , (d) Maximum, (e) Phase, (f) Jitters of different envelope frequencies



**Fig. 6.** Synchronization is independent of the amplitudes of reference signal, (a) convergence time of different amplitudes of the reference signal, (b) jitters of different amplitudes of the reference signal

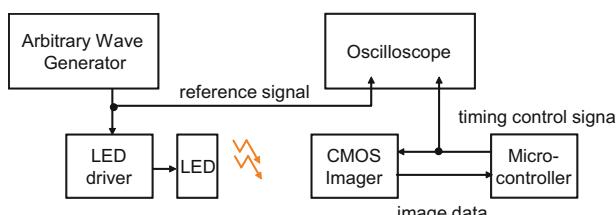
signal  $g_1(t)$ , and Fig. 5 (c) shows the correlation of its quadrature counterpart  $g_2(t)$ , while Fig. 5 (d) shows the maximum between  $q_1$  and  $q_2$ . By updating and normalizing  $q_1$  by this maximum, comparatively, the gain can be maintained as the same. The fluctuating reference signal is normalized in this way. Fig. 5 (e) shows the relative phase of the output signal to the reference signal. It can be seen that the system immediately converged to the  $\pi/2$  relative phase and became quite stable. Fig. 5 (f) shows the jitters of different reference frequencies. The faster the envelope frequency, the larger the jitters. Conceptually, the jitters were evaluated for the duration between the convergence time and 2 s from the beginning of the measurement. Leastwise, the relative phase between the reference and output is very stable and on the order of  $10^{-3}$  rad within this range of the gain.

Previously, the algorithm without signal normalization cannot tolerate too large variations of the amplitude of the reference signal. Actually, when 5-Hz sinusoidal signals with different amplitudes added to a constant direct-current (DC) synchronization was successful only as long as the sinusoidal amplitude is within 7.75 % of the DC component. The synchronization performance is analyzed in Fig. 6, in which the convergence time and jitters are completely independent of the amplitudes of the reference signal covering feasibly wide range. Furthermore, this advanced per-frame feedback control greatly reduces the convergence time on the order of  $10^{-2}$  s in Fig. 6 (a), while the jitters keep almost the same on the order of  $10^{-3}$  rad in Fig. 6 (b). The selection of coefficients has been discussed elaborately in [15].

## 4 Experiments

### 4.1 Experimental Setup

To demonstrate the proposed method, the normalization algorithm in the same way as the simulations was implemented on a real vision sensor. We employed a high-speed vision system called VCS-IV developed by the authors [16], which captures and processes images in real time at the frame rate up to around 1,000



**Fig. 7.** Block diagram of the experiment setup

Hz. The VCS-IV vision system is equipped with a  $64 \times 64$ -pixels CMOS imager called *Digital Vision Chip*, which has capability of pixel-parallel image processing programs on the focal-plane processing element array. This capability is not utilized in the presented experiment except for computation of summation of 6-bits digital pixel values over the array, which is used as  $F[i]$  in the same way as done in simulation. Fig. 7(a) shows the block diagram of the experimental setup. The illumination system consists of a Nissin Electronics LDR-90 LED array and an LPR-30W-D power supply system, which are driven by the reference square-wave signal from a Tektronics AFG3102 arbitrary wave generator.

The operation of the vision system was measured by observing the pixel reset signal of the imager, whose positive edge corresponds to the beginning of an integration period, by a Tektronics TDS3034 oscilloscope. If the operation of the vision system is locked to the illumination, synchronized waveforms of the pixel reset and the reference signal will be observed in the oscilloscope. The blue signal is the reference signal, while the yellow one locked to it is the output signal. The LED directly shed light on the vision system with an imaging optics in a normal laboratory environment as shown in Fig. 7(b). The reference signal can be reflection or indirect illumination. And the synchronization is unaffected by any real-world dynamic disturbance.

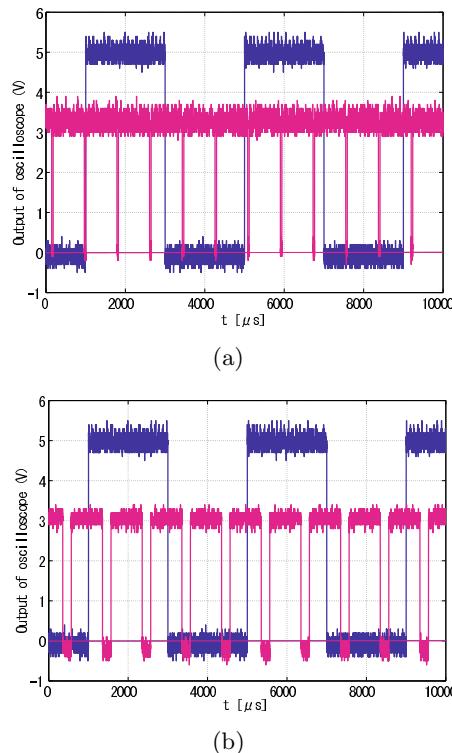
However, there is always a frequency mismatch between the oscillator that drives an illumination and the one that drives a vision sensor, and thus the central frequency of the vision sensor does not strictly agree with four times the frequency of the illumination. This mismatch causes steady-state residual phase error even in the locked state. In order to remove this error, a PI control is applied. Specifically, an integral term is added to Eq. (10) as

$$\begin{aligned} \tau_{\text{nonint}}[i] = & \tau_0 + \frac{G_p \cdot q_1[i]}{\max(|q_1[i]|, |q_2[i]|)} \\ & + G_i \sum_{j=0}^{\infty} \frac{q_1[i-j]}{\max(|q_1[i-j]|, |q_2[i-j]|)} \end{aligned} \quad (11)$$

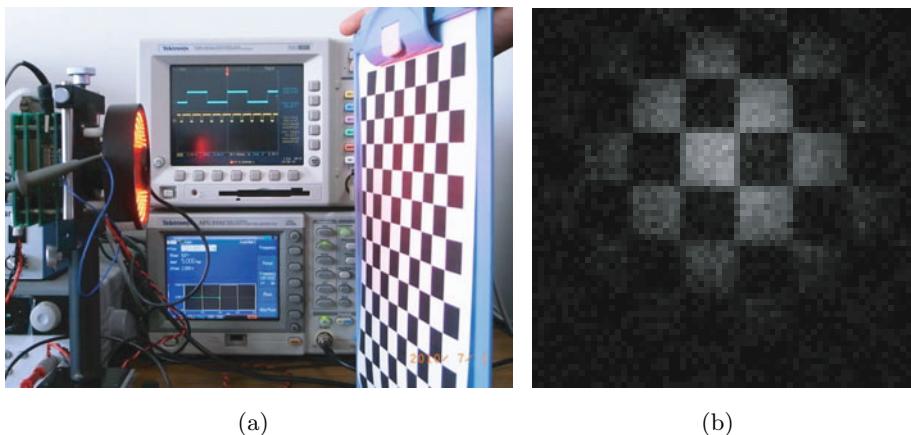
where  $G_p$  and  $G_i$  are set to 256 and 32, respectively. The coefficients of the low-pass filter in Eqs. (8) and (9) are also empirically optimized to 0.5 and 0.125.

## 4.2 Experimental Results

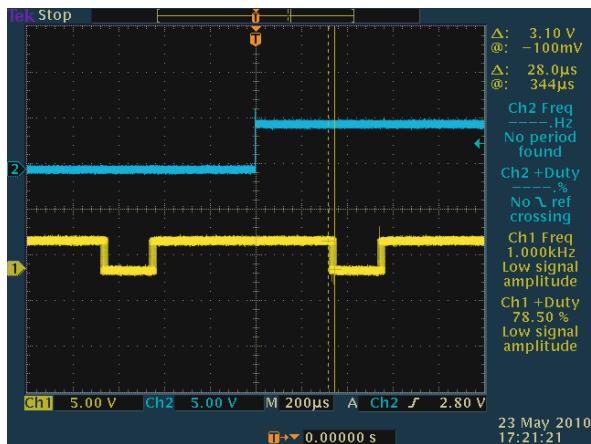
Fig. 8 (b) shows the plots of the 250-Hz square-wave signal to drive the LED and the output (pixel reset) signal that successfully locked to the reference signal compared to an unsynchronized case in Fig. 8 (a). The output signal got synchronized to the reference signal with  $\pi/2$  relative phase shift and twice the frequency. The indirect illumination employing the reflected LED light from a checkerboard is shown in Fig. 9 (a). The image taken by the vision chip is shown



**Fig. 8.** Output of the oscilloscope, (a) unsynchronized, (b) synchronized



**Fig. 9.** (a) Indirect illumination, (b) image of the checkerboard



**Fig. 10.** Peak-to-peak jitters of the output signal, 28- $\mu$ s

in Fig. 9 (b). The peak to peak jitter is around 28- $\mu$ s shown in Fig. 10 in either normal condition or dynamic scenes, which is only 1.0 % of the reference period and thus 0.1-rad phase error. This 0.1-rad phase error is satisfactory enough for practical utility.

## 5 Conclusion

An high performance illumination-based synchronization method employing signal normalization algorithm based on PLL for high-speed vision sensors has been described. The influence of illumination fluctuation can be effectively eliminated by quadrature detection combined with the novel intensity modulation method of the reference signal. Both the simulated and the experimental results showed that the electronic shutter of the vision sensors can be successfully locked to both the direct and the indirect LED illumination signal even under significant brightness fluctuation in the real-world environment.

## References

1. Fujiyoshi, H., Shimizu, S., Nishi, T.: Fast 3d position measurement with two unsynchronized cameras. In: Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation, Kobe, Japan, pp. 1239–1244 (July 2003)
2. Point Grey Research Inc.: Dragonfly Camera Synchronization, <http://www.ptgrey.com/products/multisync/index.asp> (as of 2009/02/16)

3. Rai, P.K., Tiwari, K., Guha, P., Mukerjee, A.: A cost-effective multiple camera vision system using firewire cameras and software synchronization. In: The 10th International Conference on High Performance Computing, Hyderabad, India (December 2003)
4. Litos, G., Zabulis, X., Triantafyllidis, G.: Synchronous image acquisition based on network synchronization. In: CVPRW 2006: Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop, Washington, DC, USA, pp. 167–167. IEEE Computer Society, Los Alamitos (June 2006)
5. Sivrikaya, F., Yener, B.: Time synchronization in sensor networks: a survey. *IEEE Netw.* 18, 45–55 (2004)
6. Elson, J., Girod, L., Estrin, D.: Fine-grained network time synchronization using reference broadcasts. In: Proceedings of the 5th symposium on Operating systems design and implementation, Number SI, pp. 147–163. ACM, New York (2002)
7. Ganeriwal, S., Kumar, R., Srivastava, M.B.: Timing-sync protocol for sensor networks. In: SenSys 2003: Proceedings of the 1st international conference on Embedded networked sensor systems, pp. 138–149. ACM, New York (2003)
8. Maróti, M., Kusy, B., Simon, G., Lédeczi, A.: The flooding time synchronization protocol. In: SenSys 2004: Proceedings of the 2nd international conference on Embedded networked sensor systems, pp. 39–49. ACM, New York (2004)
9. Hou, L., Kagami, S., Hashimoto, K.: Illumination-based real-time contactless synchronization of high-speed vision sensors. In: ROBIO 2009: Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics, Washington, DC, USA, pp. 1750–1755. IEEE Computer Society, Los Alamitos (2009)
10. Best, R.E.: Phase-locked loops: design, simulation, and applications. McGraw-Hill Professional, New York (2007)
11. Gardner, F.M.: Phaselock Techniques, 3rd edn. Wiley-Interscience, Hoboken (2005)
12. Hou, L., Kagami, S., Hashimoto, K.: Performance evaluation of illumination-based synchronization of high-speed vision sensors in dynamic scenes. In: International Conference on Advanced Mechatronics 2010, Osaka, Japan (2010) (to appear)
13. Ando, S., Kimachi, A.: Correlation image sensor: Two-dimensional matched detection of amplitude-modulated light. *IEEE Trans. Electron Devices* 50(10), 2059–2066 (2003)
14. Ohta, J., Yamamoto, K., Hirai, T., Kagawa, K., Nunoshita, M., Yamada, M., Yamasaki, Y., Sugishita, S., Watanabe, K.: An image sensor with an in-pixel demodulation function for detecting the intensity of a modulated light signal. *IEEE Trans. Electron Devices* 50(1), 166–172 (2003)
15. Hou, L., Kagami, S., Hashimoto, K.: Illumination-based synchronization of high-speed vision sensors. *Sensors* 10, 5530–5547 (2010)
16. Kagami, S., Komuro, T., Ishikawa, M.: A high-speed vision system with in-pixel programmable adcs and pes for real-time visual sensing. In: 8th IEEE International Workshop on Advanced Motion Control, Kawasaki, Japan, pp. 439–443 (March 2004)

# A Method to Control Ankle Exoskeleton with Surface Electromyography Signals

Zhen Zhang<sup>1</sup>, Jiaxin Jiang<sup>1</sup>, Liling Peng<sup>1</sup>, and Hongchao Fan<sup>2</sup>

<sup>1</sup> School of Mechatronics Engineering & Automation, Shanghai University,  
Shanghai, 200072, China

<sup>2</sup> School of Mechanical Engineering, Ningbo University of Technology,  
Zhejiang, 315000, China  
zhangzhen\_ta@shu.edu.cn

**Abstract.** This paper is concerned with a control method of an ankle exoskeleton with electromyographic(EMG) signals. The EMG signals of human ankle and the ankle exoskeleton are introduced first. Next a control method is proposed to control the ankle exoskeleton using EMG signals. The feed-forward neural network model applied here is composed of four layers and uses the back-propagation training algorithm. The output signals from neural network are processed by wavelet transform. At last, the control orders generated from the output signals are passed to the motor controller and drive the exoskeleton to move. Through experiment, the equality of neural network prediction of ankle movement is evaluated by correlation coefficient. It is shown from the experiment results that the proposed method can accurately control the movement of ankle joint.

**Keywords:** EMG, ankle exoskeleton, neural network, control method.

## 1 Introduction

Myoelectric signals are formed by physiological variations in the state of muscle fiber membranes. Electromyography (EMG) is an experimental technique concerned with the development, recording and analysis of myoelectric signals.[1]

For many decades surface EMG has been studied by many researchers in the medical and biomechanical fields to get a better understanding of how muscles work internally and how and when they are activated. In recent years more and more studies have explored the relationship between single muscles and the complex movements of the human body. Most of these studies are focused on analyzing disabilities, anomalies and how to track a progress in rehabilitation. In contrast to that, only a few publications focus on using EMG signals in real-time to control biomechanical robots. The main reason for this is the difficulty to map the EMG signal into the force a muscle is producing.

Beside the standard application of EMG signals to analyze disabilities or to track progress in rehabilitation, more focus has been put on controlling robot arms and exoskeletons with EMG signals (Lee[2], Fukuda[3], Mirota[4]) in recent years. In Lloyd[5] a promising but very complex musculoskeletal model was presented that takes into account 13 muscles crossing the knee to estimate the resulting knee torque.

In Fleischer's research[6], surface EMG signals were mapped to joint force of knee so that to control exoskeleton of knee. In Japan, Sankai[7] used surface EMG signals instead of goniometer to estimate the joint angle of knee.

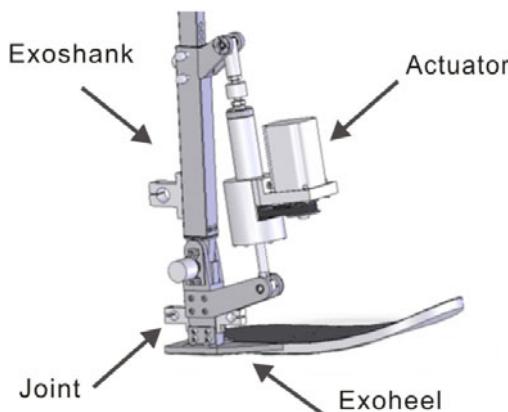
In recent years, lots of researchers focused on how to do pattern recognition or control exoskeleton with the EMG signal. A neuro-fuzzy algorithm was proposed and effectively recognized the upper action with the EMG signal[8]. Neural network was used to indentify four types of hand movement and the recognition rate was as high as 95 percent[9]. For lower limbs, there are many muscles which control human's movement. So it is difficult to control exoskeleton using surface EMG signals. In this paper, a method of controlling the ankle exoskeleton using surface EMG signals is introduced. A feed-forward neural network model is designed. The surface EMG signals of five muscles which control the ankle movement are inputted into the neural network. The ankle joint angles are estimated from the model. The control orders generated from the output signals are passed to the motor controller and can drive the exoskeleton to move.

We would like to draw your attention to the fact that it is not possible to modify a paper in any way, once it has been published. This applies to both the printed book and the online version of the publication. Every detail, including the order of the names of the authors, should be checked before the paper is sent to the Volume Editors.

## 2 Ankle Exoskeleton

### 2.1 Ankle Exoskeleton

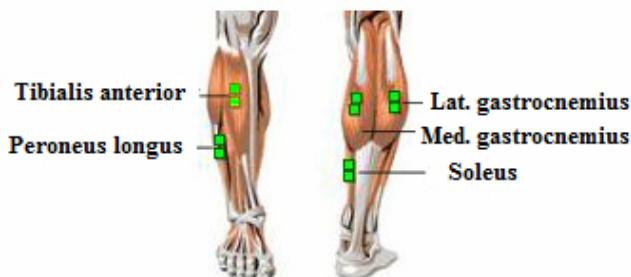
Ankle exoskeleton is a part of the powered gait orthosis. It can move legs of a patient in a physiological way on the moving treadmill. The ankle exoskeleton mechanism, shown in Fig.1, is a fairly straightforward design, utilizing an off-the-shelf ankle joint brace. Additional structural pieces are added to the ankle joint brace to extend it and to provide attachment points for the actuator. A linear series elastic actuator with a precision ball screw is connected between the shank brace and the foot brace so that it provides a torque about the ankle.



**Fig. 1.** The mechanism of ankle exoskeleton

## 2.2 Surface EMG of Ankle Joint

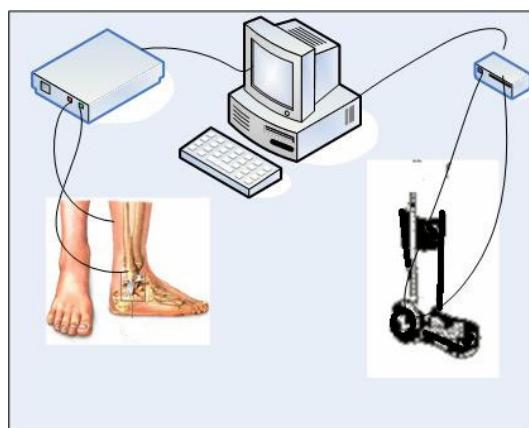
There are many muscles which control ankle joint including tibialis anterior, lateral gastrocnemius, medial gastrocnemius, peroneus longus, soleus, peroneus brevis, plantaris, flexor digitorum longus, flexor hallucis longus, tibialis posterior, extensor digitorum longus and extensor hallucis longus. Tibialis anterior(TA), lateral gastrocnemius(LG), medial gastrocnemius(MG), peroneus longus(PER) and soleus(SOL) are primary to ankle joint and they have strong surface EMG signals which could be analyzed.



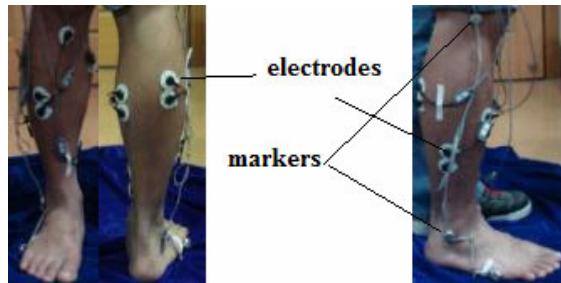
**Fig. 2.** Primary muscles of ankle joint control

## 2.3 Control System of Ankle Exoskeleton Using Surface EMG

The supposed control system is an intelligent device including a control method using surface EMG signals. As shown in Fig.3, the surface EMG signals of muscles which control human ankle movement are gathered into computer. The computer analyzes the surface EMG signals and implements the control method. The control signals from the computer are sent to the actuator controller of exoskeleton. The actuator controller realizes the movement of the ankle exoskeleton.



**Fig. 3.** The control system module



**Fig. 4.** The placement of electrodes and markers

### 3 Control Method Based on Neural Network

#### 3.1 Surface EMG Data Acquisition

The surface EMG of TA, MG, LG, PER and SOL are collected with an 8-channel surface EMG apparatus (TELEMYO2004T, Noraxon Company, USA) in the study. 2 cm silver electrodes are used as the surface EMG electrodes. The electrodes placement is shown in Fig.4. The sampling frequency is 1500Hz. The right ankle joint of subjects were asked to do regular dorsiflexion and plantar flexion movement in the sagittal plane.

For training neural network, the angle of ankle joint is recorded by means of the Optotrak system (Northern Digital, Canada). The sampling frequency of the Optotrak system is 100Hz.

#### 3.2 Feature Extraction of Surface EMG

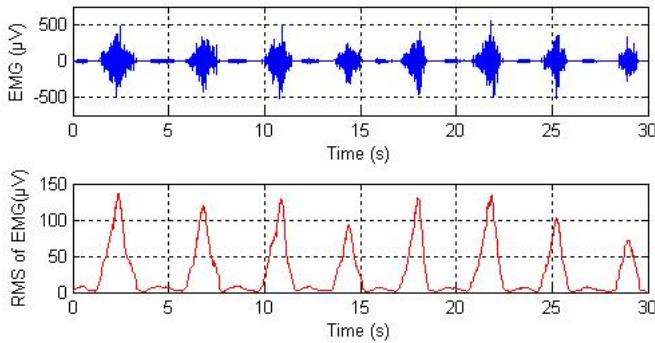
Untreated surface EMG signals are very complicated and can't be used for the control. So the rude surface EMG signals need to be treated. The surface EMG signals are rectified and filtered first. Then the root mean square(RMS) of the signals is calculated. The RMS of the surface EMG is to remove the noise. Its waveform is similar to the waveform of the linear envelope of the EMG. The RMS, which is linear relationship with muscle power, is a mathematical expression that is better than the linear envelope of the EMG. It is often used to control the threshold of myoelectric prostheses. Fig.5 is the surface EMG signals after RMS. The RMS can be computed as:

$$\sigma = \left[ \frac{1}{N} \sum_{k=1}^N s^2(k) \right]^{1/2} \quad (1)$$

Where  $s(k)$  is surface EMG signal,  $N$  is the total number of sample points.  $\sigma$  is the RMS of the surface EMG signal.

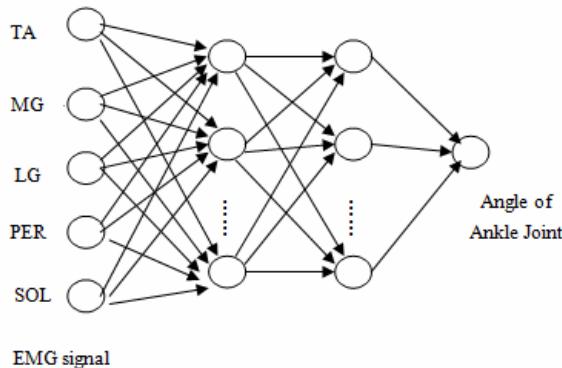
#### 3.3 Back-Propagation Neural Network

Since surface EMG signals have non-linear characteristics. It is difficult to establish a mapping model between surface EMG signals and joint angles of ankle with mathematical equation. But artificial neural network has some merits such as good adaptive



**Fig. 5.** Surface EMG signals after RMS

learning, fault-tolerance and non-linear mapping capability and so on[10][11]. Therefore, artificial neural network is used to establish a mapping relation between surface EMG signals and joint angles of ankle here. Four-layered feed forward neural network is used to estimate the joint angles from surface EMG signals in the study. And the network is trained by back-propagation algorithm. Surface EMG signals are the input signals. There are 5 neurons in input layer. Joint angle is the target output signal. There is only one neuron in output layer. Therefore, the mapping between the surface EMG and the joint angle is many-to-one. The neuron number of the first hided layer is 13. And the neuron number of the second hided layer is 23. The structure of neural network model is shown in Fig.6. Three transfer functions is used which respectively are *tansig* function, *tansig* function and *purelin* function.



**Fig. 6.** The feed-forward neural network model applied here

### 3.4 Data Post-processing

Neural network prediction output signal usually contains some noise signal. Noise has seriously impact on the forecast result. The wavelet transforms is chosen to reduce the noise. In this study, signal with noise is carried out by the six-layer wavelet decomposition. The original signal could be reconstructed with the low-frequency coefficient. This process can eliminate the high-frequency noise included the original signal. The

output signal of neural network is carried out by six layers wavelet decomposition with wavedec function in the study. And wrcoef function is used for reconstruction signal. Wavelet function is *sym8*.

### 3.5 Implement of Exoskeleton Control

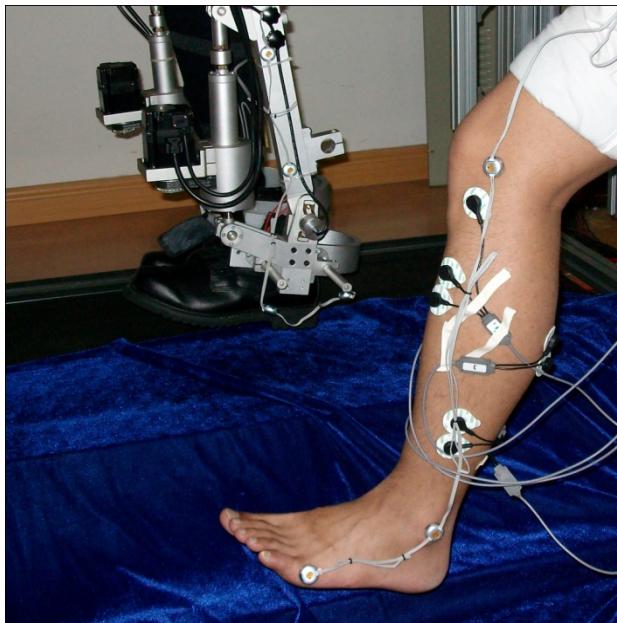
After the computer implements the control method, it sends control order to the actuator controller. The exoskeleton controller is responsible for converting the position data output from neural network to the motor controller. The NI PXI motion control card is used to realize the closed-loop control under LabVIEW platform. The 200w AC servo motor of Yaskawa Electricis is used to drive the actuator.

## 4 Experiment and Results

### 4.1 Experiment

Surface EMG signals of TA, MG, LG, PER and SOL were collected from nine male subjects, ages 24–32 (mean 26). None of the subjects had any previous ankle problems or had undergone any athletic training geared towards strengthening the ankle[12]. And they also had not undergone strenuous exercise in 24 hours before experiment. The right ankle joint of subjects were asked to do regular dorsiflexion and plantarflexion movement in the sagittal plane. The surface EMG was recorded for about 30s, which was repeated 3 times(Fig.7).

After surface EMG signal data and joint angle data are normalized, they were used for the neural network training. The training goal of neural network is 0.005. Learning rate is 0.1.

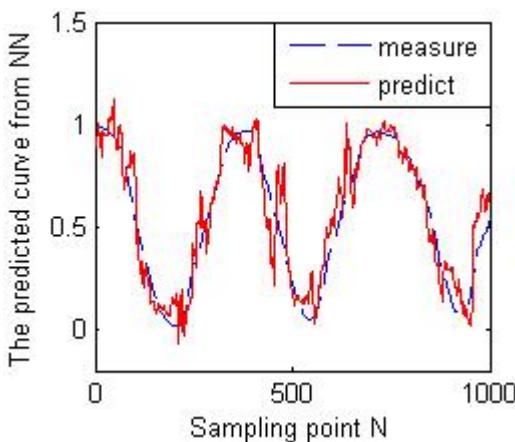


**Fig. 7.** The feed-forward neural network model applied here

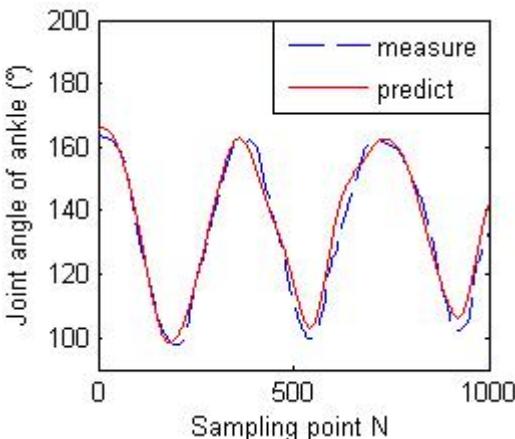
## 4.2 Results

The number of training is usually less than 60 when the neural network converges to the target value.

The predicted data are carried out by six layers wavelet decomposition and reconstructed signal (Fig.8). Quality of neural network prediction of ankle movement is evaluated by correlation coefficient. The correlation coefficient is the ratio of the predictive value and actual measured value of ankle angle. In the study, the correlation coefficients are shown in Table 1. The correlation coefficient is between the range of 0.9558-0.9915 and the average value of correlation coefficients is 0.9760. It means that the prediction result is close to the real measure value. The proposed method can accurately predict the movement of ankle joint (Fig.9).



**Fig. 8.** The contrast of the prediction result and the real measurement value before post processing (Subject9)



**Fig. 9.** The contrast of the prediction result and the real measurement value after post processing (Subject9)

## 5 Conclusion

This paper presents a control method of ankle exoskeleton based on neural network. Surface EMG signals of TA, MG, LG, PER and SOL are collected as the input of neural network. And joint angle is the target output signal. After training, the joint angle could be predicted by neural network. After post processing, the prediction result is very close to the real measurement value. Experiment proves that the proposed method can accurately predict the movement of ankle joint.

**Acknowledgments.** The research is jointly sponsored by National Natural Science Foundation(Grant No.50975165); Subject Leaders project in Shanghai(Grant No.10XD1401900); Shanghai Natural Science Fund(Grant No.01ZR1411500); Shanghai Education Research and Innovation Project(Grant No.10YZ17); Mechatronics Engineering Innovation Group project from Shanghai Education Commission; Scientific Research Fund of Zhejiang Provincial Education Department(Grant No.Y200906890).

## References

- Peter, K.: The ABC of EMG. Noraxon Inc., USA (April 2005)
- Sukhan, L., George, N.S.: The control of a prosthetic arm by EMG pattern recognition. *IEEE Transactions on Automatic Control*, AC 29, 290–302 (1984)
- Osamu, F., Toshio, T., Hiroki, S., et al.: An EMG controlled human supporting robot using neural network. In: Proceedings of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 1586–1591 (1999)
- Satoshi, M., Toshiyuki, K., Koji, I.: Estimation of forearm movement from EMG signal and application to prosthetic hand control. In: Proceedings of the IEEE Int. Conf. on Robotics & Automation, pp. 3692–3697 (2001)
- David, G.L., Thor, F.B.: An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo. *Journal of Biomechanics* 36, 765–776 (2003)
- Christian, F., Konstantin, K., Christian, R., et al.: Online calibration of the emg-toforce relationship. In: Proceedings of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (2004)
- Hiroaki, K., Suwoong, L., Shigehiro, K., et al.: Power assist method for HAL-3 using EMG-based feedback controller. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, vol. 2, pp. 1648–1653 (2003)
- Kiguchi, K.: A study on EMG-Based Human Motion Prediction for Power Assist Exoskeleton. In: Proceeding of the 2007 IEEE International Symposium On Computational Intelligence in Robotics and Automation, pp. 190–195 (2007)
- Korenberg, M.J., Morin, E.L.: Parallel cascade classification of myoelectric signals. In: Proceeding of IEEE 17th Annual Conference of Engineering in Medicine and Biology Society, vol. 2, pp. 1399–1400 (1995)
- Liu, M.M., Herzog, W., Savelberg, M.: Dynamic muscle force prediction from EMG: an artificial neural network approach. *Journal of Electromyography and Kinesiology* 9, 391–400 (1999)
- Sepulveda, F., Wells, D.M., Vaughan, C.L.: A neural network representation of electromyography and joint dynamics in human gait. *Journal of Biomechanics* 26, 101–109 (1993)
- Zhang, Z., Yao, S.L., Zhang, Y.N., Qian, J.W.: On the Surface Electromyography Sensor Network of Human Ankle Movement. In: IEEE International Conference on Robotics and Biomimetics, pp. 1688–1692 (2007)

# A Resistance Compensation Control Algorithm for a Cable-Driven Hand Exoskeleton for Motor Function Rehabilitation

Shuang Wang, Jiting Li, and Ruoyin Zheng

The State Key Laboratory of Virtual Reality Technology and Systems  
Robotics Institute, Beihang University, Beijing, China  
wangshuang\_110@yahoo.com.cn, lijiting@buaa.edu.cn,  
ryzheng860909@126.com

**Abstract.** The resistance compensation, especially the friction compensation in the Bowden cable transmission is a difficult issue to be handled. Aimed to the resistance reduction requirement in the active rehabilitative motion, a resistance compensation control method is proposed. Based on the simplified transmission model, the resistance, including the cable friction as well as the mechanical moment of inertial, is formulated. To realize the compensation, force sensors are used to measure the force exerted by the human fingertip. With the proposed algorithm, the maximum finger-exerted force is reduced to less than one third of before. The experimental result demonstrates the validity of the proposed method.

**Keywords:** Hand rehabilitation, active control mode, resistance compensation, cable/sheath transmission, hand exoskeleton.

## 1 Introduction

As we know, the motor capability of hand is crucial and important for human-being's activity of daily life. Hands, however, as the most vulnerable limb compared with arms and legs are much easier to be injured in accident and by some diseases. The impaired motor function needs to be recovered through rehabilitation. The traditional rehabilitation approach is costly for patients and laborious for therapists. Recent research showed that hand rehabilitative training using mechatronic devices and virtual reality is possible and effective [1] and is attracting much research interests [2-7]. Some robot-assisted rehabilitation systems are developed and the corresponding control algorithms are investigated [8-11].

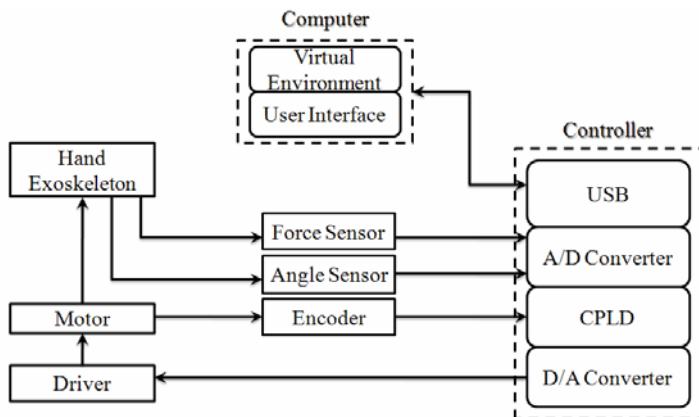
For some existing exoskeleton-type hand rehabilitation devices, the cable and sheath is utilized in mechanical systems to satisfy the need of remote and changeable distance transmission and reduce the weight exerted on the patient hand. Although there reaches a consensus that the friction between the cable and sheath accounts for the great part of the mechanical resistance, the problem to compensate the friction is still open and unresolved. In this paper we propose a force control algorithm which can provide resistance compensation for the active rehabilitative motion. With the

proposed approach, the measured maximum force which hand output is reduced to about one third.

The remainder of the paper is organized as follows. Section 2 introduces the system architecture. Section 3 presents force control algorithm with resistance compensation. Section 4 depicts experiment and the results. Section 5 gives conclusion and future work.

## 2 System Description

The system architecture is shown as Fig.1, which consists of the hand exoskeleton integrated with angle and force sensors, the controller and the virtual environment.

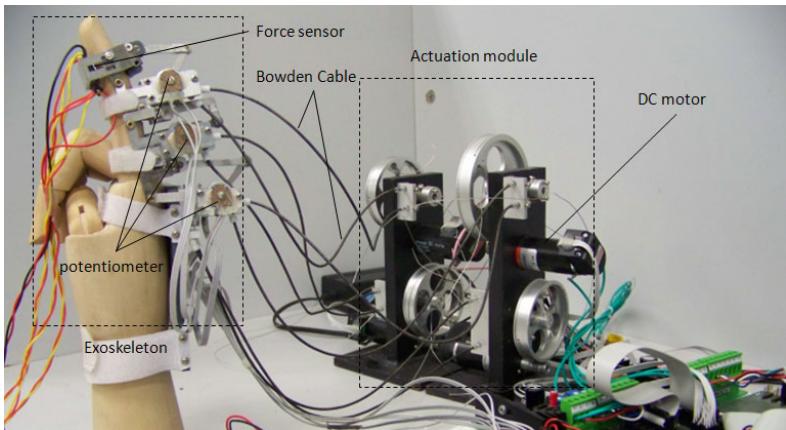


**Fig. 1.** The System Architecture

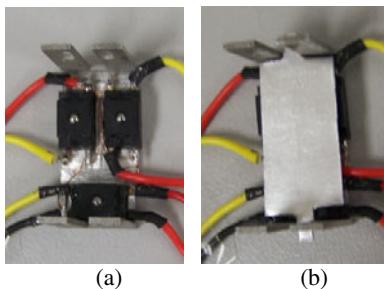
The exoskeleton device for index finger rehabilitation has 4 degrees of freedom and consists of three parts: the actuation module, the cable transmission and the exoskeleton, as shown in Fig 2 [12].

The exoskeleton is worn on the dorsal side of the hand. It is actuated by four actuators and can implement the motion of flexion/extension for the DIP, PIP and MCP1 joints, and the motion of adduction/abduction for the MCP2 joint. To reduce the weight which is imposed on the patient hand, the actuator is placed in a distance from the hand. Corresponding to the human hand anatomical structure the exoskeleton is comprised of three parts which are attached to the phalanges – distal, middle and proximal phalanges, respectively. For each finger joint, two cables, each of which is housed in a sheath, are used to transmit force and motion from the actuator to the exoskeleton. A potentiometer is installed on the each joint shaft of the exoskeleton to measure the rotational angle of the finger joint. Honeywell\_FSS force sensor is used to measure the fingertip force exerted by the human fingertips. Because the measure area (a small spherical surface) of the sensor is too small for soft fingertip to measure the reliable

force, three force sensors (Honeywell\_FSS shown in Fig. 3) are assembled together on the bottom of the distal module as shown in fig. 3(a) and are covered with a metal plate as shown in fig. 3(b) to ensure that the force exerted by the hand could be measured accurately. The forces on three contact points read from the force sensor are summed to obtain the resultant force which is exerted by the finger [13].



**Fig. 2.** Prototype of the hand exoskeleton



**Fig. 3.** (a) The Layout of Three force sensors (b) Covered by a Metal Plate

The real-time controller is developed by our research group. It could sample the angle data and the force data in real time. The controller links to the host computer by the USB, and the sampling frequency is 100HZ.

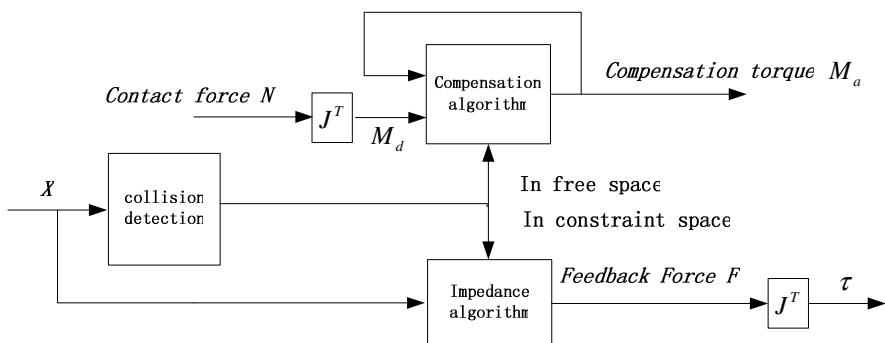
The motors are driven by the driver (EM-28 DC-MOTOR CONTROL UNIT). The driver is connected to the controller through analog output channels. And it runs under the torque control mode.

The host computer runs the virtual environment and gives the force feedback to the patient. The operation interface allows the therapist to set the training parameters for the rehabilitation therapy, for example, range of joint angle, training time, training speed and so on. In the graphic interface, the virtual hand, which is controlled by the human hand, is used to accomplish the different rehabilitation tasks.

### 3 Control Algorithm

During the active rehabilitation training, the virtual hand works in two different states. When it is in contact with the virtual object, we call it is in constraint space and the virtual contact force is feedback to the human finger. Otherwise it is called in free space and the human finger is supposed to move without resistance. Therefore the mechanical resistance is demanded to be compensated.

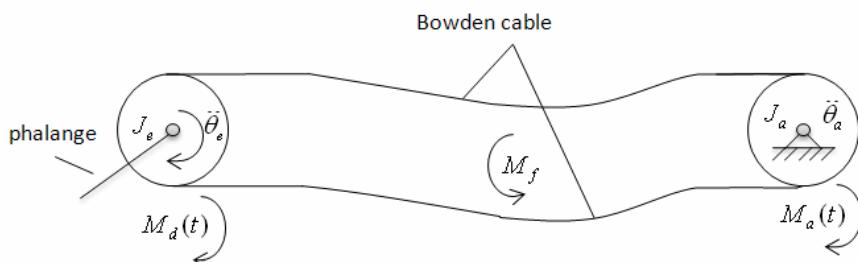
To satisfy the requirements in both spaces, the impedance control is used [14]. The block diagram of control structure is shown in Fig. 4, where,  $N$  is the contact forces exerted by human finger. The control method in constraint space will not be described in detail and we only focus on the resistance compensation control algorithm in the free space in the following part.



**Fig. 4.** The block diagram of control structure

As aforementioned, in free space the exoskeleton is supposed to follow the motion of the finger and be compliant with the finger's motion. However, the friction between the cable and sheath, as well as the moment of inertial of the system, cause great resistance to the finger flexion/extension. Thus we propose a method to compensate the resistance of the finger flexion.

Taking one joint for example, the actuator module, the cable/sheath transmission and the exoskeleton module are simplified as follows (Fig 5).



**Fig. 5.** The simplified model for one finger joint

According to the equilibrium equation of the moment at a certain time  $t$ , we have:

$$M_d(t) + M_a(t) = J_e * \ddot{\theta}_e(t) + J_a * \ddot{\theta}_a(t) + M_f(t) \quad (1)$$

where  $M_d(t)$  is the driving torque which is produced by the fingertip-exerted force;

$M_a(t)$  is the output torque of the actuator module which is used to compensate the system resistance;

$M_f(t)$  is the resistance torque due to the friction between the cables and the sheaths;

$J_e$  and  $J_a$  are the moment of inertia of the exoskeleton module and of the actuator module, respectively;

$\ddot{\theta}_e(t)$  and  $\ddot{\theta}_a(t)$  are the angular acceleration of the exoskeleton module and the actuator module, respectively.

If we want to compensate the mechanical transmission caused by the friction and the moment of inertia, the actuator module output  $M_a(t)$  should be:

$$M_a(t) = J_e * \ddot{\theta}_e(t) + J_a * \ddot{\theta}_a(t) + M_f(t) \quad (2)$$

Because  $M_f(t)$  at time  $t$  is unknown, so we estimate  $M_a(t)$  with the value at time  $t-1$ . Considering that the time interval is small so the estimation is reasonable. So we have:

$$M_a(t) = M_d(t-1) + M_a(t-1) \quad (3)$$

The active rehabilitative motion requires that it is the human hand that drives the exoskeleton. However, if the estimated  $M_a(t)$  is greater than the actual resistance, the exoskeleton is then driven by the actuator, and the active rehabilitative motion is changed to passive rehabilitative motion which is not desired. On the other hand, the human fingertip will not contact with the sensors which means the contact force can not be measured. Therefore we modify equation (3) as follows.

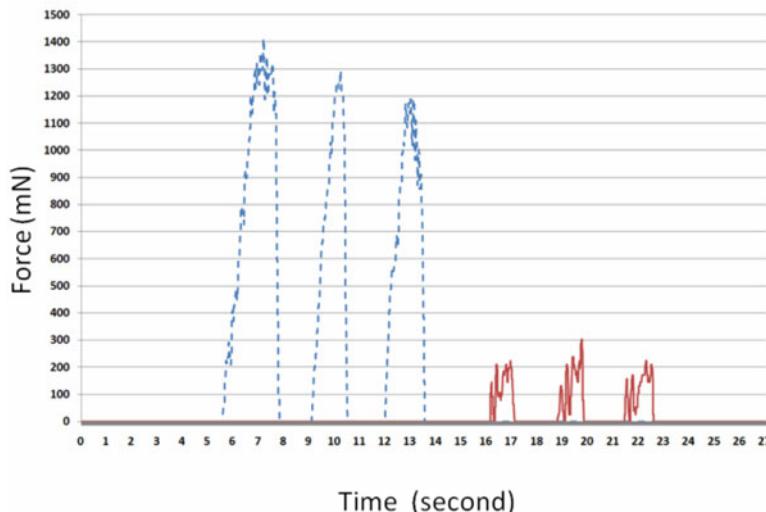
$$M_a(t) = \xi * [M_d(t-1) + M_a(t-1)] \quad (4)$$

where the coefficient  $\xi$  is set to be not more than 1.0 and its value could determined by experiment. The driving torque  $M_d(t-1)$  could be calculated from the jacobian matrix and the contact force measured from the force sensors.

## 4 Experiments and Result

In the experiment, the human hand wears the exoskeleton and index finger makes the motion of flexion and extension continuously for three times in two conditions: with

and without resistance compensation. In two conditions, the relative position of the exoskeleton module and the actuator module is approximately kept unchanged. The contact forces exerted by human finger in the two conditions are measured, respectively. The result is shown in Fig.6, where the coefficient  $\xi$  is set to be 0.9 and the accuracy of the force sensor is 10mN.



**Fig. 6.** The measure forces exerted by the human fingertip. Blue dashed line—without the resistance compensation, red solid line—with the resistance compensation

Fig 6 shows that with the compensation control algorithm, the maximum output force of the human fingertip is less than one third of the maximum force without. And the subject felt much easier with the compensation control. The experimental result demonstrates the validity of the proposed method.

## 5 Conclusion

Focused on the special requirements of the active rehabilitative motion for hand motor capability, a control structure integrated with resistance compensation is proposed. With the proposed controller the undesired resistance is greatly reduced. The advantage of the algorithm is that the complicated resistance model is avoided and the method is very simple.

Next, we are going to further test the performance of the proposed control algorithm, such as the force feedback in the constraint space. We will also expand the method to compensate the resistance during finger extension and the other two rehabilitation modes, i.e. assisted and resisted mode.

## References

1. Fischer, H.C., Stubblefield, K., Kline, T., Luo, X., Kenyon, R.V., Kamper, D.G.: Hand Rehabilitation Following Stroke: A Pilot Study of Assisted Finger Extension Training in a Virtual Environment. *Topics in Stroke Rehabilitation* (January- February 2007)
2. Ito, S., Kawasaki, H., Ishigre, Y.: A Design of Fine Motion Assist Equipment For Disabled Hand in Robotic Rehabilitation System. In: Proceeding of the First International Conference on Modeling, Simulation and Applied Optimization (February 2005)
3. Lelieveld, M.J., Maeno, T., Tomiyama, T.: Design and Development of Two Concepts for a 4 DOF Portable Haptic Interface With Active and Passive Multi-Point Force Feedback for The Index Finger. In: ASME International Design Engineering Technical Conference & Computers and Information in Engineering Conference (2006)
4. Fang, H., Xie, Z., Liu, H.: An Exoskeleton Master Hand for Controlling DLR/HIT Hand. In: The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, USA, October 11-15 (2009)
5. Fontana, M., Dettori, A., Salsedo, F., Bergamasco, M.: Mechanical design of a novel Hand Exoskeleton for accurate force displaying. In: 2009 IEEE International Conference on Robotics and Automation Kobe International Conference Center Kobe, Japan, May 12-17 (2009)
6. Wang, J., Li, J., Zhang, Y., Wang, S.: Design of an Exoskeleton for Index Finger Rehabilitation. In: 31st Annual International Conference of the IEEE EMBS Minneapolis, Minnesota, USA, September 2-6 (2009)
7. Fu, Y., Wang, P., Wang, S., Liu, H., Zhang, F.: Design and development of a portable exoskeleton based CPM machine for rehabilitation of hand injuries. In: IEEE International Conference on Robotics and Biomimetics, pp. 1476–1481 (December 2007)
8. Wege, A., Kondak, K., Hommel, G.: Mechanical Design and Motion Control of a Hand Exoskeleton for Rehabilitation. In: IEEE International Conference on Mechatronics and Automation, pp. 155–159 (2005)
9. Wege, A., Hommel, G.: Development and Control of a Hand Exoskeleton for Rehabilitation of Hand Injuries. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3046–3051 (August 2005)
10. Wang, S., Li, J., Zhang, Y., Wang, J.: Active and Passive Control of an Exoskeleton with Cable Transmission for Hand Rehabilitation. In: CISPR 2009-BMEI 2009 international Conference (October 2009)
11. Vallery, H., Venenman, J., Van Asseldonk, E., Ekkelenkamp, P., Buss, M., Van Der Kooij, H.: Compliant Actuation of Rehabilitation Robots. *IEEE Robotics & Automation Magazine* 15(3), 60–69 (2008)
12. Li, J., Wang, S., Wang, J.: Design and Analysis of an Exoskeleton Device for Finger Rehabilitation. In: 1st International Conference on Applied Bionics and Biomechanics, ICABB (2010) (submitted)
13. Jiang, L., Cutkosky, M.R., Ruutiainen, J., Raisamo, R.: Using Haptic Feedback to Improve Grasp Force Control in Multiple Sclerosis Patients. *IEEE Transactions on Robotics* 25(3) (June 2009)
14. Hogan, N.: Impedance control an approach to manipulatio I II III. *ASME, Transactions, Journal of Dynamic Systems, Measurement, and Control* 107, 1–24 (1985)

# Kinematic Modeling of Bevel Tip Flexible Needle

Zhang Yong-de and Zhao Yan-jiang

Intelligent Machine Institute, Harbin University of Science and Technology,  
Harbin, Heilongjiang Province, China  
[{zhangydy, zhaoyj}@hrbust.edu.cn](mailto:{zhangydy,zhaoyj}@hrbust.edu.cn)

**Abstract.** Aiming at the kinematic modeling of bevel tip flexible needle, this paper proposes the deflection unicycle model (DUM) for the nonholonomic system of the needle. Based on this, according to the actual path and the error analysis, a novel deflection unicycle model with rebound (DUMR) is proposed, which more accords with the facts. Kinematics is simplified and forward kinematics is calculated using screw theory and product of exponentials formula. The least square method is adopted to fit the experimental data and the parameters of the model are gained. Results show that the model with rebound is much superior to the model without rebound. What's more, the maximum error, the number of large errors and the root mean square relative to the model are little enough to fit the actual path well.

**Keywords:** Flexible needle; nonholonomic system; kinematic modeling; screw theory.

## 1 Introduction

Needle insertion is perhaps the most widely used technique in surgery, especially in minimally invasive surgery. It is a critical aspect of many medical diagnoses, treatments, and scientific studies. However there are distortions between tissue and traditional steel needle because of their interaction. The distortions cause the large errors that could not be compensated. This badly influences the effect of biopsy and therapy. What's more, as there are some sensitive organs (like nerve or vein) and obstacles (like bones) to be avoided, traditional steel can hardly meet this need. Some people proposed the bendable needle ideas. Okazawa et al. incorporated a pre-bent stylet inside a straight cannula, if needed, the curve path would be obtained when the stylet inserts outside the cannula [1]. Sears et al. coupled two and more curved concentric tubes to realize different curvatures of paths [2]. Glozman and Shoham controlled the needle base, utilizing the interaction between needle and tissue to make the needle bent [3]. Tang et al. adopted magnetic force to drive the symmetric magnetized head with a hinge [4]. However, the bendable needles mentioned above are still rigid to tissue. In this paper, we adopt a bevel tip flexible needle (“flexible needle” for short) proposed by Webster et al. [5]. The needle, which is flexible enough to tissue, can move along a curve path because of the lateral force between the bevel tip and tissue.

Therefor, the needle path can be precisely controlled, errors can be greatly minished and the obstacles can be avoided to reach the target.

For several years, there are rare researches on kinematic modeling. DiMaio and Salcudean do some researches on force modeling and path planning, but the studies are aiming at symmetric-tip steel needle [6]. There is only one research on flexible needle by Webster III et al. They proposed a bicycle model and a unicycle model, considering that the path of flexible needle is equivalent to that of a bicycle with constant angle of front wheel or equivalent to that of a unicycle. But these models neglect the elastic force of needle shaft, considering that it is absolutely flexible and the needle tip “stays in place” during reorientation. This definitely gives rise to a sort of errors between model and actual path. In this paper, based on the actual path, after analysis of errors causation, we presented a novel kinematic model that more accord with the facts, and obtained the model parameters with experiment data.

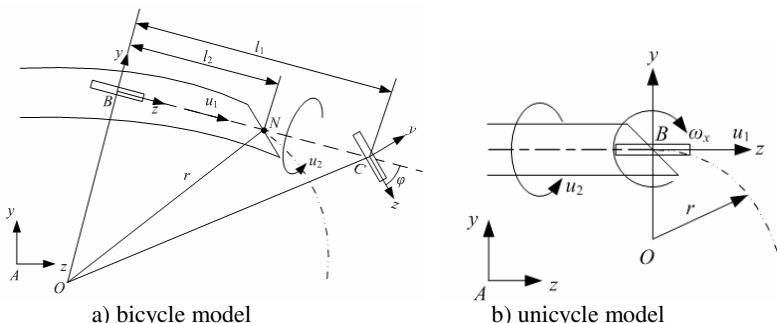
## 2 Kinematic Model of Flexible Needle

### 2.1 Deflection Unicycle Model (DUM)

Bevel tip flexible needle was driven by two inputs, namely the feed input  $u_1$  and the rotation input  $u_2$ , as shown in Fig.1. This system is a nonholonomic system. In order to simplify the question, reasonable assumptions are proposed as follows:

- (1) Organs are uniform and isotropic;
- (2) The needle movement entirely depends on the tip, that is to say the needle shaft is exactly follow the path of needle tip.

From the mechanics analysis: as the needle is inserted into tissue with feed input  $u_1$ , the tissue imposes a reaction force on the bevel that deflects the needle tip, causing it to follow an arc with a constant radius  $r$ , which is independent of feed speed [5]. Rotation motion  $u_2$  is to change the direction of the bevel tip, thus to change the moving direction of the needle tip. Therefor the movement of the needle tip is equivalent to that of front or rear wheel or a point at the frame of a bicycle, or equivalent to that of unicycle wheel, as shown in Fig.1. And the bicycle model or unicycle model system is also a nonholonomic system [5,7].

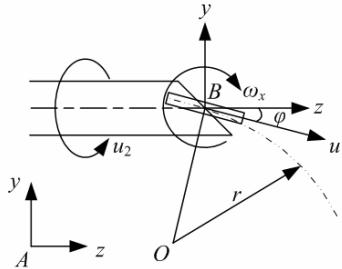


**Fig. 1.** Bicycle model and unicycle model

For bicycle model and unicycle model, the the needle path functions are as follow:

$$\begin{cases} \mathbf{p} = f(\varphi, l_1, l_2) & \text{bicycle model} \\ \mathbf{p} = f(r) & \text{unicycle model} \end{cases}. \quad (1)$$

Although the unicycle model has only one parameter, it is inferior to bicycle model, which more accords with the actual path. For a system model, the more parameters it has the more complex the calculation is. Accordingly, the purpose of this paper is to create a model with the least parameters and to fit the actual motion well, meanwhile the kinematics calculation is simplified. Here we present DUM as shown in Fig.2.



**Fig. 2.** Deflection unicycle model

Supposing that the feed speed and the rotation speed are  $u_1$  and  $u_2$  respectively; the arc radius is  $r$ ; the unicycle is attached to the needle tip with a deflection  $\varphi$ , leaning to the bevel direction; the frame  $B$  is attached to the unicycle and the bevel tip, the  $z$  axis points to the direction of movement along the centerline of needle shaft, and the direction of  $x$  axis is the same as that of the needle instantaneous rotation axis when moving along the arc of radius  $r$ . Thus, there is an angular drive  $\omega_x$  along the positive direction of  $x$  axis.

$$\omega_x = u_1 / r. \quad (2)$$

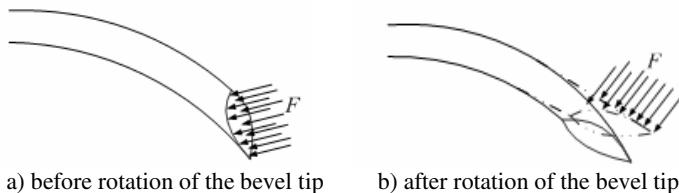
As the needle path radius  $r$  is independent of feed speed [5], so it is independent of the angular velocity  $\omega_x$  as well, and also the angular velocity  $\omega_x$  is proportional with the feed speed  $u_1$ . Thus the path of needle tip is only related to two parameters, deflection angle  $\varphi$  and radius  $r$ . Thereby the path equation is simplified into a function with two parameters:

$$p = f(\varphi, r). \quad (3)$$

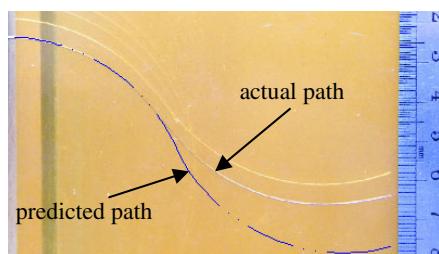
## 2.2 Deflection Unicycle Model with Rebound (DUMR)

From the results of reference [5], although the bicycle model is better than unicycle model, it remained considerable errors. The main cause of these errors is to excessively idealize needle and tissue when modeling, neglecting the tissue deformation and needle bending force. In this way, the needle tip will “stay in place” during reorientation of the bevel tip. But in fact, when the needle is inserted to some distance and then rotated with some angle, the needle tip will destroy the tissue, and what’s more, the effective interactive area between needle tip and tissue will be reduced as the direction of the bevel is changed, and the force imposed on the needle tip is

reduced, too. And therefore, there will be a certain degree of rebound of the needle shaft, as shown in Fig.3. So there will be large errors between the model predictive path and actual path in the next arc, as shown in Fig.4.

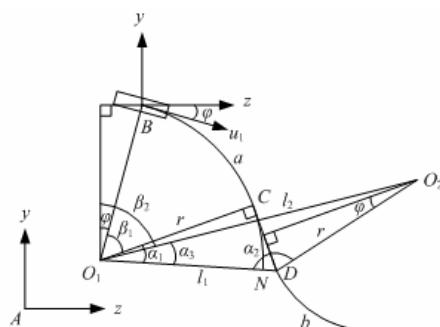


**Fig. 3.** Changes of position and force before and after rotation of the bevel tip



**Fig. 4.** Contrast between the prediction of DUM and actual path

So we must consider the rebound of the needle tip after rotation. In order to illustrate the problem easily and obviously, we are going to study the planar path with rebound. As shown in Fig.5, assuming that the needle tip is inserted to point  $N$  along the arc  $a$ , after rotation, a section of the original terminal arc  $CN$  (defined as  $reb$ ) rebounds into a section of line, and the original arc is held to point  $C$ , the needle tip rebounds to point  $D$ , so the length of  $CD$  equals the length of  $reb$ , and the line  $CD$  is tangent to the arc  $a$  at point  $C$ . And arc  $b$  from point  $D$  is the next section of path according to the model after rebound. Thereby the whole motion of the needle is equivalent to: firstly moving to  $C$  along arc  $a$ , then translation to  $D$  along the tangent direction or feed direction, rotating some angle (180 degrees), and then moving along arc  $b$ .



**Fig. 5.** Trajectory of DUMR

In this way, one more parameter is needed for the kinematic model, so the path function of DUMR is as follow:

$$p = f(\varphi, r, reb). \quad (4)$$

### 3 Calculation of Kinematics of Flexible Needle

Unlike traditional robots, the flexible needle has no joints and rods, so the kinematics cannot be calculated by D-H method. Here we resort to product of exponentials formula whose most charming character is that it only uses two frames, the world frame and the body frame, to realize the kinematics. Combining with geometric character of twist, it is the best substitute for D-H method [7].

#### 3.1 Definitions and Calculation

Fig.2 shows two frames, world frame  $A$  and body frame  $B$ . The homogeneous transformation from  $B$  to  $A$  is as follow:

$$\mathbf{g}_{ab} = \begin{bmatrix} \mathbf{R}_{ab} & \mathbf{P}_{ab} \\ 0 & 1 \end{bmatrix} \in SE(3). \quad (5)$$

We introduce the  $\vee$  (vee) and  $\wedge$  (wedge) operations and isomorphism between  $\mathbb{R}^3$  and  $so(3)$ :

$$\begin{cases} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}^\wedge = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \in so(3) \\ \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}^\vee = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \in \mathbb{R}^3 \end{cases}. \quad (6)$$

$$\begin{cases} \begin{bmatrix} \hat{\boldsymbol{\omega}} & \mathbf{v} \\ \mathbf{0} & 0 \end{bmatrix}^\vee = \begin{bmatrix} \mathbf{v} \\ \boldsymbol{\omega} \end{bmatrix} = \boldsymbol{\xi} \\ \hat{\boldsymbol{\xi}} = \begin{bmatrix} \mathbf{v} \\ \boldsymbol{\omega} \end{bmatrix}^\wedge = \begin{bmatrix} \hat{\boldsymbol{\omega}} & \mathbf{v} \\ \mathbf{0} & 0 \end{bmatrix} \end{cases}. \quad (7)$$

The transformation from  $B$  to  $A$  is defined as  $\mathbf{g}_{ab}$ , and the body velocity is as follow:

$$\mathbf{V}_{ab}^b = \begin{bmatrix} \mathbf{v}_{ab}^b \\ \boldsymbol{\omega}_{ab}^b \end{bmatrix} = (\mathbf{g}_{ab}^{-1} \dot{\mathbf{g}}_{ab})^\vee = \boldsymbol{\xi}. \quad (8)$$

Besides, the unit vectors  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$  are the standard basis.

#### 3.2 Kinematics Calculation

As is shown in Fig.2, body frame  $B$  attaches to the needle tip. In DUM, there are five constraints. Frame  $B$  can not move along  $x$  axis of frame  $B$ . The velocities of  $y$  and  $z$

axis are the projection of the feed speed  $u_1$  along the  $y$  and  $z$  axis of frame  $B$  respectively, and the deflection angle between the direction of feed speed  $u_1$  and  $z$  axis is a constant  $\varphi$ . The rotational speed around  $x$  axis is proportional with the feed speed, as is expressed in equation (2). Frame  $B$  can not rotate around  $y$  axis. The constraints are expressed as follow:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & r & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_{ab}^b \\ \boldsymbol{\omega}_{ab}^b \end{bmatrix} = \begin{bmatrix} 0 \\ -u_1 \sin \varphi \\ u_1 \cos \varphi \\ u_1 \\ u_1 \\ 0 \end{bmatrix}. \quad (9)$$

We can obtain:

$$\xi = \mathbf{V}_{ab}^b = [0 \ -u_1 \sin \varphi \ u_1 \cos \varphi \ u_1 / r \ 0 \ u_2]^T. \quad (10)$$

As the feed and rotation motion can be effect respectively, let  $\mathbf{u} = (u_1, u_2)$  denote the control inputs, then the object velocity can be described:

$$\xi = \mathbf{V}_{ab}^b = u_1 \mathbf{V}_1 + u_2 \mathbf{V}_2. \quad (11)$$

Where  $\mathbf{V}_1$  and  $\mathbf{V}_2$  are the of pure feed velocity vector and pure rotation velocity vector respectively.

$$\mathbf{V}_1 = \begin{bmatrix} \mathbf{v}_1 \\ \boldsymbol{\omega}_1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\sin \varphi \\ \cos \varphi \\ \mathbf{e}_1 / r \end{bmatrix}, \quad \mathbf{V}_2 = \begin{bmatrix} \mathbf{v}_2 \\ \boldsymbol{\omega}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{3 \times 1} \\ \mathbf{e}_3 \end{bmatrix}. \quad (12)$$

According to the equation (11), exponentials product form of the homogeneous transformation matrix of kinematics can be denoted as follow:

$$\mathbf{g}_{ab}(t) = \mathbf{g}_{ab}(0) \exp(\xi t) = \mathbf{g}_{ab}(0) \exp((u_1(t)\hat{\mathbf{V}}_1 + u_2(t)\hat{\mathbf{V}}_2)t). \quad (13)$$

The potion of the needle tip is:

$$\mathbf{p}(t) = \mathbf{p}_{ab}(t). \quad (14)$$

### 3.3 Discrete Model

Let  $\mathbf{g}_{ab}(k)$  denote the homogeneous transformation from body frame  $B$  to the world frame  $A$  in step  $k$ ,  $\mathbf{u}(k+1) = (u_1(k+1), u_2(k+1))$  denote the velocity in step  $k+1$ , the needle moves  $T$  s with no direction change of the needle tip, then the homogeneous transformation in  $k+1$  step can be denoted:

$$\mathbf{g}_{ab}(k+1) = \mathbf{g}_{ab}(k) \exp((u_1(k+1)\hat{\mathbf{V}}_1 + u_2(k+1)\hat{\mathbf{V}}_2)T), \quad k = 0, 1, 2, \dots. \quad (15)$$

If after  $k+1$  step, the direction of the needle tip changes, and then the needle moves  $T$  s, then the modified homogeneous transformation of step  $k+1$  and the homogeneous transformation of step  $k+2$  can be denoted:

$$\begin{cases} \tilde{\mathbf{g}}_{ab}(k+1) = \mathbf{g}_{ab}(k) \exp((u_1(k+1)\hat{\mathbf{V}}_1 + u_2(k+1)\hat{\mathbf{V}}_2)(T-t)) \\ \mathbf{g}_{ab}(k+2) = \tilde{\mathbf{g}}_{ab}(k+1) \mathbf{g}_{tran} \exp((u_1(k+2)\hat{\mathbf{V}}_1 + u_2(k+2)\hat{\mathbf{V}}_2)T) \end{cases} \quad k = 0, 1, 2, \dots. \quad (16)$$

Where  $t$  is modified time,  $t = reb / u_1$ ,  $reb$  is the rebound,  $\mathbf{g}_{tran}$  is the transition transformation,  $\mathbf{g}_{tran} = \mathbf{g}_L(reb)\mathbf{g}_z(\theta_{k+2})$ , where  $\mathbf{g}_L(reb)$  is the homogeneous transformation after moving a distance of  $reb$  along the direction of feed ( $z$  axis of the body frame).  $\mathbf{g}_z(\theta_{k+2})$  is the homogeneous transformation after the direction the needle tip changes that is the rotation around  $z$  axis.

Let  $\mathbf{g}_{tran} \exp((u_1(k)\hat{\mathbf{V}}_1 + u_2(k)\hat{\mathbf{V}}_2)t_k) = \mathbf{g}_k(t_k)$ , then the forward kinematics can be described:

$$\mathbf{g}_{ab}(t) = \mathbf{g}_{ab}(0)\mathbf{g}_1(t_1)\mathbf{g}_2(t_2)\dots\mathbf{g}_n(t_n). \quad (17)$$

Where  $t$  is the total time of the movement,  $t_1 \sim t_n$  is movement time at each step.

From the equation (14), we can obtain:

$$\mathbf{p}(k) = \mathbf{p}_{ab}(k). \quad (18)$$

## 4 Experiment and Model Parameters Fitting

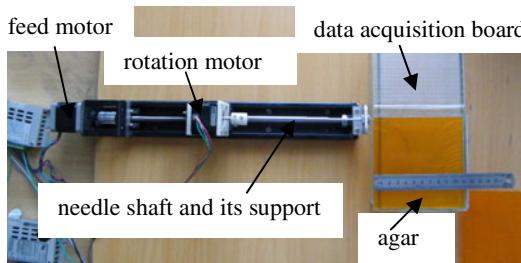
From the equations (4) and (16), we know that:

$$\mathbf{p} = f(u_1, u_2, \varphi, r, reb). \quad (19)$$

As  $u_1$  and  $u_2$  are independent of the path, we can choose the values in the experiment. But when they simultaneously effect on the system, their ratio will affect the path shapes [8]. As for  $\varphi$ ,  $r$  and  $reb$ , they can be obtained by fitting experimental data. However it is too complicated to fit the parameters using the kinematics mentioned above, so we resort to geometry.

### 4.1 Experiment Equipments

The whole needle insertion system consists of a control system, a needle insertion mechanism and an imitation of organ, as shown in Fig.6. Needle insertion mechanism is composed of two motors, screw-nut mechanism, needle shaft and its support. Feed motor drives screw-nut pairs to realize feed motion of needle shaft, and rotation motor drives needle spin. We adopt agar to imitate organs, use a grid with an interval of 1mm as the data acquisition board, so the measurement error is  $\pm 1\text{mm}$ . The needle used in the experiment is a cylinder nitinol wire with the diameter of 0.4mm, and with a bevel tip of 30 degrees. Though the value of the path radius  $r$  will be affected by the

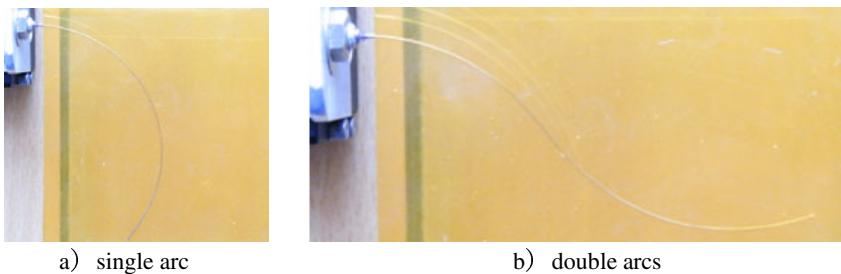


**Fig. 6.** Experiment equipments of flexible needle system

needle shaft diameter, bevel tip angle and simulated organs, according to previous theoretical analysis, we can get the conclusion that their paths are similar and kinematic model is the same. So we only choose this case to study, the relationship between model parameters and them will be studied in future work.

## 4.2 Experiment Result and Parameters Fitting

The experiment is composed of two parts, one is single arc:  $v=0.009\text{m/s}$ ,  $\omega=0$ , insertion into organ with 110mm, the other is double arcs: firstly,  $v=0.009\text{m/s}$ ,  $\omega=0$ , insertion into organ with 50mm; secondly,  $v=0$ ,  $\omega=3.5\text{rad/s}$ , rotation with 180 degrees, and lastly,  $v=0.009\text{m/s}$ ,  $\omega=0$ , insertion into organ with 60mm, as is shown in Fig.7. Each part of experiment performs for five times.



**Fig. 7.** Needle insertion experiment result

In order to prove the importance of the deflection and the rebound, we firstly compare DUM with bicycle model and unicycle model. Secondly, we compare the models with and without rebound. Here, we assume that the radius of the arc is  $r$ , the deflection of DUM is  $\varphi$ . So the central position of the arc is:

$$\mathbf{c} = f(r, \varphi). \quad (20)$$

Assuming that the insertion position is  $(0, y_0, z_0)$ , so for single arc, the central position of the arc is:

$$\mathbf{c} = [0 \quad y_0 \pm r \cos \varphi \quad z_0 - r \sin \varphi]^T. \quad (21)$$

Where  $\pm$  depends on the bending orientation of arcs, while bending to the positive of y axis, it is “+”, while bending to the negative of y axis, it is “-”.

For the double arcs or more, The central positon of first arc is same as that of the single arc.

$$\mathbf{c}_1 = [0 \quad y_{c_1} \quad z_{c_1}]^T = \mathbf{c}. \quad (22)$$

Then the central positions of the following arcs are:

For DUM without rebound:

$$\begin{aligned} \mathbf{c}_i &= [0 \quad y_{c_i} \quad z_{c_i}]^T = \\ &[0 \quad y_{c_{i-1}} + 2r \cos \varphi \cos(S_{i-1} / r) \quad z_{c_{i-1}} + 2r \sin \varphi \sin(S_{i-1} / r)]^T \quad (i = 2, 3, \dots) \end{aligned} \quad (23)$$

Where  $S_i$  is the length of the  $i$ th arc section.

For DUMR, as shown in Fig.5:

$$\begin{cases} \alpha_1 = \arctan(reb / r) \\ \alpha_2 = \pi - \alpha_1 - \varphi \end{cases} . \quad (24)$$

$$\begin{cases} l_1 = reb / \sin \alpha_1 \\ l_2 = \sqrt{l_1^2 + r^2 - 2l_1 r \cos \alpha_2} \end{cases} . \quad (25)$$

$$\alpha_3 = \arccos\left(\frac{l_1^2 + l_2^2 - r^2}{2l_1 l_2}\right). \quad (26)$$

$$\begin{cases} \beta_1 = (S_1 - reb) / r \\ \beta_2 = \varphi + \beta_1 + \alpha_1 - \alpha_3 \end{cases} . \quad (27)$$

So the central position of the following arc:

$$\mathbf{c}_i = \begin{bmatrix} 0 & y_{c_i} & z_{c_i} \end{bmatrix}^T = \begin{bmatrix} 0 & y_{c_{i-1}} + l_2 \cos \beta_2 & z_{c_{i-1}} + l_2 \sin \beta_2 \end{bmatrix}^T. \quad (28)$$

The central positions for unicycle model with and without rebound are easier to calculate according to their geometric characters.

In the models, the relationship between  $y$  and  $z$  is:

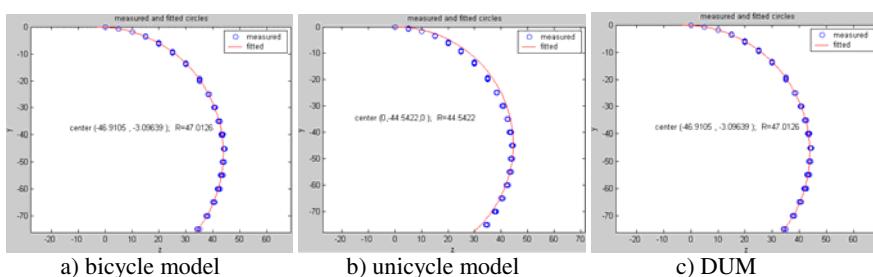
$$y = f_c(z) = \pm \sqrt{r^2 - (z - z_{c_i})^2} + y_{c_i}. \quad (29)$$

Where  $\pm$  depends on the position of sample point, while the coordinate value of  $y$  of the point is larger than that of the arc center, it is "+", on the contrary, it is "-".

We use the least square method to fit the model:

$$\min \sum_{i=1}^n (y_i - f_c(z_i))^2. \quad (30)$$

We believe that all the radii of the arcs are the same, so we fit the single arcs in advance. The results of bicycle model, unicycle model and DUM are respectively shown in Fig.8 and Table 1.



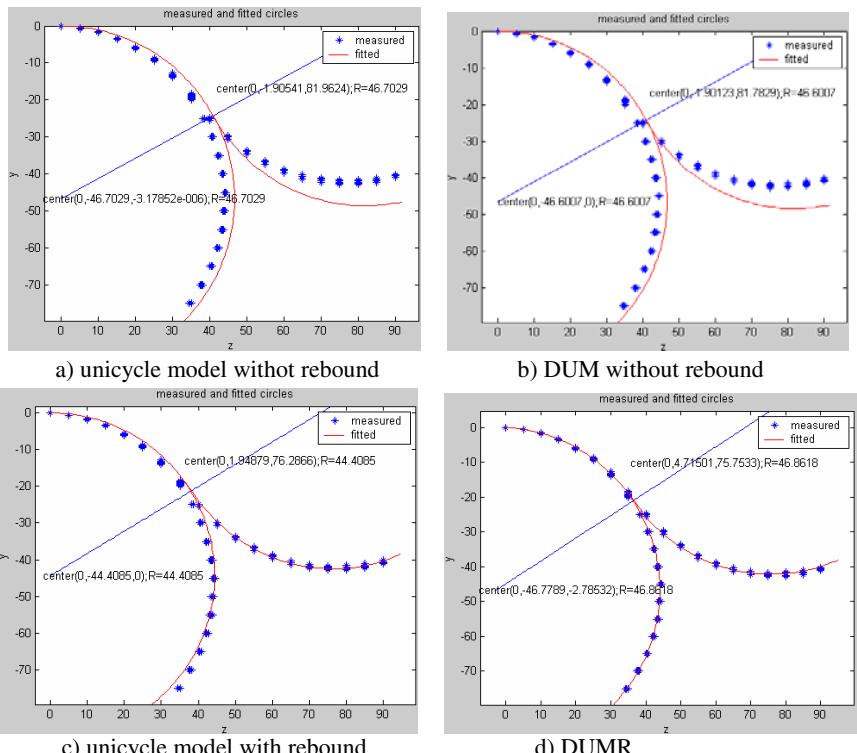
**Fig. 8.** Fitting results of models for single arc

**Table 1.** Fitting results of models for single arc

	$r$	$\varphi$	$l_1$	$l_2$	root mean square	maximum error
bicycle model	47.0126	9.9902°	8.2633	3.0964	0.2478	0.6226
unicycle model	44.5422	/	/	/	1.5848	2.8657
DUM	47.0126	3.7758°	/	/	0.2478	0.6226

From the data in figures and table, we can conclude that, the results of bicycle model and DUM are the same. That is to say the paths of them are the same. However, it is much better to choose DUM for the better parameter design and kinematics calculation (contrast with reference [5]). Although unicycle model is simple, with only one parameter  $r$ , it doesn't accord with the facts. The needle path accords with DUM. Although DUM is better than unicycle model (without deflection) for single arc, it is suspending for the double arcs. So we still have to compare the two models for double arcs. In order to verify the worth of rebound, we also contrast the models with and without rebound.

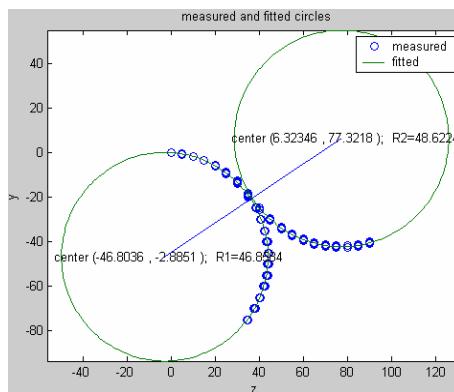
To substitute the 10 groups of data into the formula (32), to fit them with the four models mentioned above. The results are shown in Fig.9 and Table 2.

**Fig. 9.** Fitting results of models with and without rebound

**Table 2.** Fitting results of models with and without rebound

	$r$	$\varphi$	$reb$	root mean square	maximum error	number of large errors
unicycle model without rebound	46.6007	/	/	3.3347	7.7826	163
DUM without rebound	46.7029	$6.8 \times 10^{-8}$	/	3.3358	7.7115	163
unicycle model with rebound	44.4085	/	8.9543	1.2757	2.8088	85
DUMR	46.8618	0.0595	9.9517	0.3937	1.0713	1

From the data in figures and table, we can get the conclusions that the model with rebound is much superior to the model without rebound. And further, DUMR is superior to the unicycle model with rebound. What's more, the fitted parameters are close to the single arc parameters respectively. In order to prove this, we purely use multiple points to fit circles. The result is shown in figure 10. The radius of the circle fitted according to single arc and the first arc is 46.8584mm, while the radius of the circle fitted according to the second arc is 48.6224mm. So we can say that DUMR is much more accord with the fact.

**Fig. 10.** Circle fitting using multiple points

The most important issue is that the root mean square of the model is small enough, and the maximum error is small enough, too. Although a certain large error which a little exceeds the measurement error of  $\pm 1$ mm, there is only one such a datum among the 190 data. From the fitting results, the large error appears at the end of the second arc, it is due to the rotation angle error of the needle tip caused by torsional compliance. The path of the needle will offset the measuring plane no matter positive or negative error appears in rotation. According the projection theory, although the raduis of second arc remain constant in its insertion plane, the measurement value will appear larger, so the large error comes out, which accords with Fig.10. From the experiment, we can see that the path is very sensitive to the rotation angle errors. So the path will be more exactitude and the model will more accord with the facts if we can control the rotation angle precisely.

## 5 Conclusion

Making reference to the bicyclic model, we present DUM based on the path analysis of the flexible needle. Furthermore, according to the error analysis, DUMR is proposed. Forward kinematics has been calculated using screw theory and product of exponentials formula. The least square method is adopted to fit the experimental data, and the parameters of the model are gained. Through the contrast with other models, we can draw the conclusions as follows:

1. This model not only has the simple kinematics, but it much accords with the facts as well;
2. The model with rebound is much superior to the model without rebound, and DUMR is superior to the unicycle model with rebound. What's more, the root mean square of the model is small enough, and the maximum error is small enough, too. This model accord with the facts well;
3. Although a certain large error which a little exceeds the measurement error of  $\pm 1\text{mm}$ , there is only one datum of this kind among the 190 data. It is due to the rotation angle error of the needle tip.

So how to precisely control or correctly compensate the rotation angle of the tip is an important work in future. And the relationship between the model parameters and needle diameter, bevel angle and characters of organs is also the future work with great importance.

**Acknowledgments.** This work was supported in part by National Natural Science Foundation under grant 51075105 and Heilongjiang Province Education Bureau under grant 20080820.

## References

1. Okazawa, S., Ebrahimi, R., Chuang, J., et al.: Hand-held steerable needle device. *IEEE/ASME Transactions on Mechatronics* 10(3), 285–296 (2005)
2. Sears, P., Dupont, P.: A steerable needle technology using curved concentric tubes. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2850–2856. IEEE, Piscataway (2006)
3. Glozman, D., Shoham, M.: Image-guided robotic flexible needle steering. *IEEE Transactions on Robotics* 23(3), 459–467 (2007)
4. Tang, L.B., Chen, Y.H., He, X.H.: Magnetic force aided compliant needle navigation and needle performance analysis. In: *2007 IEEE International Conference on Robotics and Biomimetics*, Sanya, China, pp. 612–616 (2007)
5. Webster III, R.J., Kim, J.S., Cowan, N.J., et al.: Nonholonomic modeling of needle steering. *International Journal of Robotics Research* 25(5/6), 509–525 (2006)
6. DiMaio, S.P., Salcudean, E.S.: Needle steering and motion planning in soft tissues. *IEEE Transactions on Biomedical Engineering* 52(6), 965–974 (2005)
7. Murray, R.M., Li, Z., Sastry, S.S.: *A mathematical introduction to robotic manipulation*. CRC Press, USA (1994)
8. Engh, J.A., Podnar, G., Khoo, S.Y., et al.: Flexible needle steering system for percutaneous access to deep zones of the brain. In: *IEEE 32nd Annual Northeast Bioengineering Conference*, pp. 103–104. IEEE, Piscataway (2006)

# An Adaptive RBF Neural Network Control Strategy for Lower Limb Rehabilitation Robot

Feng Zhang, Pengfeng Li, Zengguang Hou, Xiaoliang Xie, Yixiong Chen,  
Qingling Li, and Min Tan

Laboratory of Complex Systems and Intelligence Science, Institute of Automation,  
the Chinese Academy of Sciences, Beijing 100190, China

[zfeng1203@yahoo.com.cn](mailto:zfeng1203@yahoo.com.cn), [pengfeng.li@ia.ac.cn](mailto:pengfeng.li@ia.ac.cn), [hou@compsys.ia.ac.cn](mailto:hou@compsys.ia.ac.cn),  
[xiaoliang.xie@ia.ac.cn](mailto:xiaoliang.xie@ia.ac.cn), [yixiong.chen@ia.ac.cn](mailto:yixiong.chen@ia.ac.cn), [doudouhit@163.com](mailto:doudouhit@163.com),  
[min.tan@ia.ac.cn](mailto:min.tan@ia.ac.cn)

**Abstract.** This paper proposed an adaptive control strategy based on RBF (radial basis function) neural network and PD Computed-Torque algorithm for precise tracking of a predefined trajectory. This control strategy can not only give a small tracking error, but also have a good robustness to the modeling errors of the robot dynamics equation and also to the system friction. With this control algorithm, the robot can work in assist-as-needed mode by detecting the human active joint torque. At last, a simulation result using matlab simulink is given to illustrate the effectiveness of our control strategy.

**Keywords:** Rehabilitation robot, RBF, active training, adaptive control.

## 1 Introduction

There has been more and more patients in the world who suffered from SCI or stroke and the treatment to these patients is a long way after surgery. Conventionally speaking, doctors help these patients do rehabilitation training by hand or by simple device such as automated vehicles. These traditional treatments are useful to the rehabilitation of patients to some extent. But more and more studies proved that the active participation of patients in training can do a better favor to the neuro-rehabilitation. Literature [1][2] also suggest that active training can improve the cortical reorganization. Clinic trials have also proved that repetitive movement exercise is helpful to neurologic rehabilitation [3][4], even though that the scientific basis for neuro-rehabilitation remains ill-defined. So most of the limb rehabilitation robots also focus on the various movement training. Exoskeleton robots are mostly often used for these purpose.

Rehabilitation robot is designed for active training and repetitive movement exercise of patients who suffered from paraplegia or hemiplegia, so that the basic function of a rehabilitation robot is that the robot can detect the active participation of patients and can automatically assist the patients as needed. There are many ways to detect the active participation of patients, for example, the

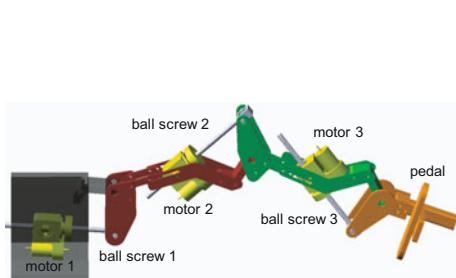
force sensor or torque sensor can be used to detect the patients' force that exerted onto the robot, the biological signal such as electromyography (EMG) or electroencephalogram (EEG) can be used to detect the intent of patients and also the status of muscles [5][6]. The MIT-MANUS which was designed by MIT in 1991 is an upper limb rehabilitation robot, it uses EEG to detect the patients' intent and hemiplegia patients can do some easy work with the help of the robot. The MotionMaker is an lower limb rehabilitation robot which was designed by SWORTEC, and EMG is used to detect the muscle activity. The Locomat is another lower limb rehabilitation robot which was designed by HOCOMA, it consists of a treadmill and a body weight support system. The Locomat robot can detect the torque the patients exerts onto the Locomat and then pull the patient's leg into the desired direction [7]. Most of the literatures about rehabilitation robots are mainly focus on the smart design or kinetic control [8][9][10], there is little literature about the dynamics control strategy to accomplish the precise trajectory tracking. In the following sections of this paper, we will give a introduction of our self-designed lower limb rehabilitation robot, and propose an adaptive control strategy based on RBF neural network and PD Computed-Torque algorithm for precise tracking of a predefined trajectory.

## 2 Simplified Dynamical Model of System

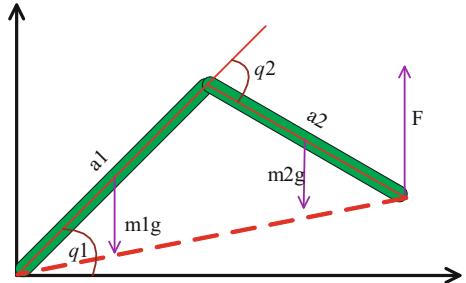
The self-designed 3-DOF exoskeleton lower limb rehabilitation robot model is showed in Fig.1. We only give an inner mechanical structure of the robot. The people's leg can be attached to the robot links by straps, and the foot can be fastened to the pedal. According to actual needs, the exoskeleton lower limb rehabilitation robot can help patients do various movement in a vertical plane, including single-joint movement and joint united movement. Three joints of the robot are all driven by ball screws. The ball screw is a line unit which changes the rotational motion of electrical motor into linear motion, and then the linear motion of the ball screw drives the link of the robot to rotate around the joint fulcrum. Compared with geared drive, this design can reduce the system friction, and increase the system damping, and reduce the influence of the external disturbance to the whole system. Three absolute encoders are installed in every joint to record the real joint angle, and the absolute encoders can maintain the current output even when the power is off. Also there is a force sensor installed at the end of each ball screw which is used to detect the joint torque. There is another more important force sensor placed on top of the foot treadle which is also the third link of the robot. This force sensor is used to detect the force the human foot exerts onto the foot treadle.

When using the platform for rehabilitation training, the length of each link should be adjusted firstly to set the link length equal to the human thigh and crus. Then the foot of the patient is fastened to the foot treadle. We can see the robot combined with the human leg as a whole, a three-link system. For simplification, the mass of human leg and the robot link are seen as evenly distributed, and the third link of the system is omitted because it just impacts

the foot posture and has no effect to the end trajectory. After simplification, the whole system can be seen as a well-proportioned two-link system showed in Fig. 2.



**Fig. 1.** 3-DOF rehabilitation robot for lower limbs



**Fig. 2.** Simplified model of system

## 2.1 Lagrange's Dynamics of Human Leg and Robot

By using Lagrange-Euler method, we can easily get the dynamics of the whole system, as showed in Fig. 2. Firstly, the Lagrange's equation of motion for the two-link system is defined as followed

$$\tau_i = \frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}_i} \right] - \frac{\partial L}{\partial q_i} \quad (1)$$

where  $\tau_i$  is joint torque,  $q_i$  is the joint angle,  $L$  is the Lagrangian which represents the difference between the kinetic and potential energies

$$L = K - P \quad (2)$$

For link 1, the kinetic and potential energies can be given by

$$\begin{cases} K_1 = \frac{1}{6} m_1 a_1^2 \dot{q}_1^2 \\ P_1 = \frac{1}{2} m_1 g a_1 \sin q_1 \end{cases} \quad (3)$$

where  $m_1$  is the total mass of robot link 1 and human thigh,  $a_1$  is the length of link 1 and human thigh,  $q_1$  is the joint angle of hip. For link 2, the kinetic and potential energies can be given by

$$\begin{cases} K_2 = \left( \frac{1}{2} m_2 a_1^2 + \frac{1}{6} m_2 a_2^2 + \frac{1}{2} m_2 a_1 a_2 \cos q_2 \right) \dot{q}_2^2 \\ \quad + \left( \frac{1}{3} m_2 a_2^2 + \frac{1}{2} m_2 a_1 a_2 \cos q_2 \right) \dot{q}_1 \dot{q}_2 + \frac{1}{6} m_2 a_2^2 \dot{q}_1^2 \\ P_2 = m_2 g a_1 \sin q_1 + \frac{1}{2} m_2 g a_2 \sin(q_1 + q_2) \end{cases} \quad (4)$$

where  $m_2$  is the total mass of robot link 2 and human crus,  $a_2$  is the length of link 2 and human thigh,  $q_2$  is the joint angle of knee.

By combining (3) and (4), and computing the formula (1), we can get the ideal dynamics of the whole system which can be written in the standard form

$$M_0(q)\ddot{q} + V_0(q, \dot{q})\dot{q} + G_0(q) = \tau \quad (5)$$

with  $M_0(q)$  the inertia matrix,  $V(q, \dot{q})$  the Coriolis/centripetal vector, and  $G_0(q)$  the gravity vector. The specific form of these matrices are given in the simulation section. Note that  $M_0(q)$  is symmetric and positive definite, and  $\dot{M}(q) - V_0(q, \dot{q})$  is a skew-symmetric matrix, and the gravity vector  $G_0(q)$  is bounded [11].

## 2.2 Detect the Joint Torque of Human Leg

Actually, in formula (5), the torque  $\tau$  is comprised of two parts, one is caused by the voluntary movement of patient and the other is caused by robot motor. To detect the torque of human leg that exerts to the joint is very important and it will decide the output torque of robot motor.

We suppose that the human foot and the third link is always horizontal in training, so that the force of foot is always perpendicular to the ground, and the drawing force and compression force transducer on top of the third link can detect the the human voluntary force. From Fig.2 we can see that if there is no voluntary force, the reading of the transducer should be produced by the leg gravity. The relationship between them is

$$F_s = \frac{(\frac{1}{2}m_t + m_c)ga_1 \cos q_1 + \frac{1}{2}m_c g a_2 \cos(q_1 + q_2)}{a_1 \cos q_1 + a_2 \cos(q_1 + q_2)} \quad (6)$$

With  $F_s$  the reading of transducer at static status,  $m_t$  and  $m_c$  the mass of thigh and crus. If the patient exerts voluntary force to the pedal, the reading of the transducer should be bigger or smaller than  $F_s$ . From the reading of transducer we can judge the system running mode. Specifically, if the reading is between 0 and  $F_s$ , then the robot arm will work in the assistance mode, else if the reading is bigger than  $F_s$  or less than 0, then the robot arm will work in the resistance mode. If the current reading of transducer is  $F$ , then the difference between  $F_s$  and  $F$  is the patient's voluntary force. Note that the leg gravity is not voluntary force. By transforming the force into joint torque, we can get

$$\tau_h = \begin{bmatrix} (F_s - F)a_1 \sin(q_1 + q_2) \sin q_2 \\ (F_s - F)a_2 \cos(q_1 + q_2) \end{bmatrix} \quad (7)$$

where  $\tau_h$  is patient's voluntary torque.

## 2.3 Control Strategy

In practice, the precise model of the plant is hard to get, and we can only build the ideal model. If the model (5) we have built is precise, then a Computed-PD

algorithm as followed is enough to make a good tracing of the predefined trajectory [11].

$$\tau = M_0(q)(\ddot{q}_d - K_v \dot{e} - K_p e) + V_0(q, \dot{q})\dot{s} + G_0(q) \quad (8)$$

where  $q_d$  is the desired trajectory, and  $e = q - q_d$ ,  $\dot{e} = \dot{q} - \dot{q}_d$ , the derivative gain matrix  $K_v$  and the proportional gain matrix  $K_p$  are selected positive definite.

But if the true model is

$$M(q)\ddot{q} + V(q, \dot{q})\dot{q} + G(q) + F(q, \dot{q}) = \tau \quad (9)$$

where  $F(q, \dot{q})$  is friction and disturbance item. Put the control law (8) into (9), and we get

$$\ddot{e} + K_v \dot{e} + K_p e = M_0^{-1}(\Delta M \ddot{q} + \Delta V \dot{q} + \Delta G - F(q, \dot{q})) \quad (10)$$

where  $\Delta M = M_0 - M$ ,  $\Delta V = V_0 - V$ , and  $\Delta G = G_0 - G$ . Easily we can see that the control law (8) is not satisfying. The unprecise part of the model is the right item of (10). We define the unprecise part of the model as

$$f(x) = M_0^{-1}(\Delta M \ddot{q} + \Delta V \dot{q} + \Delta G - F(q, \dot{q})) \quad (11)$$

If we have known the unprecise part, then we can design the controller as

$$\tau = M_0(q)(\ddot{q}_d - K_v \dot{e} - K_p e - f(x)) + V_0(q, \dot{q})\dot{s} + G_0(q) \quad (12)$$

Substitute (12) into (9), we can get the following error differential equation

$$\ddot{e} + K_v \dot{e} + K_p e = 0 \quad (13)$$

From this error differential equation, we can see that if the derivative gain matrix  $K_v$  and the proportional gain matrix  $K_p$  are selected positive definite, the system must converge within a certain time.

Unfortunately, the unprecise part of model is not available more often. The common method to resolve this problem is to compensate for the unprecise part of model. Neural network is a common method to approximate the unprecise item of a system [12]. The BP network and RBF network are two more often used neural networks. The BP network is a global approximation network, and the convergence is slow and be easy to fall into local minimum. On the contrary, the RBF network is a local approximation network, and has a fast learning speed, also it can avoid the local minimum [12]. Here we choose RBF neural network to approximate the unprecise part of the model. The structure of RBF neural network is showed in Fig.3. The gauss basis function is also chosen as followed

$$\phi_j = \exp\left(-\frac{\|X - c_j\|^2}{2b_j^2}\right), j = 1, 2, \dots, m \quad (14)$$

In the RBF network,  $X = [x_1, x_2, \dots, x_n]^T$  is the input vector,  $c_j$  is the central vector of node  $j$ ,  $c_j = [c_{j1}, c_{j2}, \dots, c_{jn}]$ , and  $b_j$  is the base width of node  $j$ .

$\varphi = [\phi_1, \dots, \phi_m]^T$  is the radial basis vector. If the weights vector of network is  $W = [w_1, \dots, w_m]^T$ , then the output of network is

$$y(t) = w_1\phi_1 + w_2\phi_2 + \dots + w_m\phi_m \quad (15)$$

In our control system, we choose  $X = [e \dot{e}]^T$  as input vector,  $\hat{W}$  as the weights vector which will be adjusted online,  $W^*$  is the best approximation weight vector, and the network output is  $\hat{f}(x, w) = \hat{w}^T \varphi(x)$ , and  $\hat{W}$  is the estimation value of  $W^*$ . At the basis of (8)-(10), we design the controller as

$$\tau = M_0(q)(\ddot{q}_d - K_v \dot{e} - K_p e - \hat{f}(x, w)) + V_0(q, \dot{q})\dot{s} + G_0(q) \quad (16)$$

If we substitute (16) into (9), then we will get

$$\dot{x} = Ax + B\{f(x) - \hat{f}(x, w)\} \quad (17)$$

where  $f(x) = M_0^{-1}(\Delta M \ddot{q} + \Delta V \dot{q} + \Delta G - F(q, \dot{q}))$ , and  $A = \begin{bmatrix} 0 & I \\ -k_p & -k_v \end{bmatrix}$ ,  $B = \begin{bmatrix} 0 \\ I \end{bmatrix}$ .

If we define the Lyapunov function as

$$V = \frac{1}{2}x^T Px + \frac{1}{2\gamma}\|\tilde{w}\|^2 \quad (18)$$

where  $P$  is a positive definite symmetric matrix, and it satisfies the Lyapunov function as followed

$$PA + A^T P = -Q \quad (19)$$

where  $Q \geq 0$ . If we choose the adaptive law as

$$\dot{\tilde{w}} = \gamma\varphi(x)x^T PB + k_1\gamma\|x\|\hat{w} \quad (20)$$

with  $\gamma > 0, k_1 > 0$ . The condition of convergence is

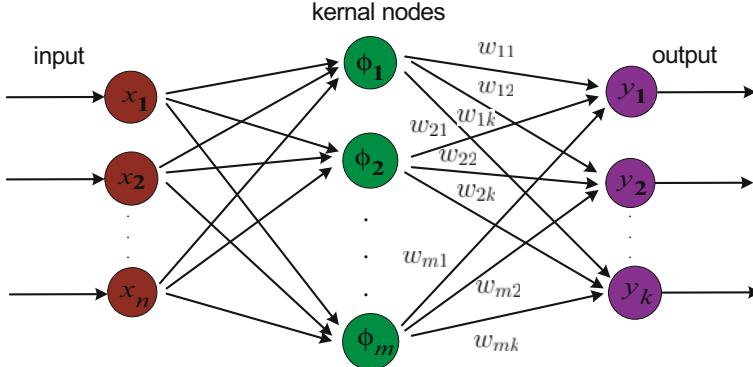
$$\|x\| \geq \frac{2}{\lambda_{min}(Q)} \left( \|\eta_0\| \lambda_{max}(P) + \frac{k_1}{4} w_{max}^2 \right) \quad (21)$$

with  $\eta_0$  the ideal error between  $f(x)$  and  $\hat{f}(x, w^*)$ .

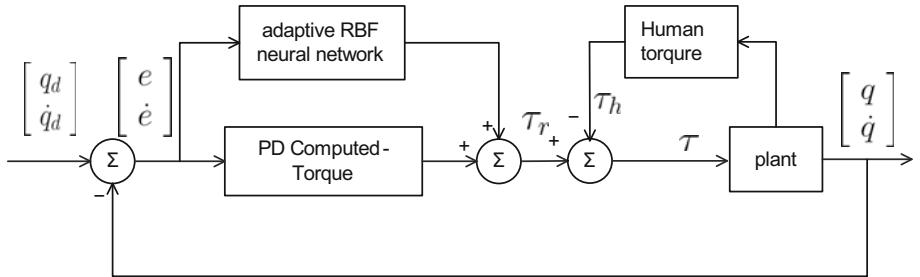
With above discussions, we can get the control system block diagram as Fig. 4.

### 3 Simulation

In order to verify the resulting performance and robustness of the control algorithm, the simulation is conducted using SIMULINK toolboxes of MATLAB. The namely model we use is the model we have built as (8), where the inertia matrix  $M_0(q)$  is given by  $M_0(q) = \begin{bmatrix} \alpha + 2\eta \cos q_2 & \beta + \eta \cos q_2 \\ \beta + \eta \cos q_2 & \beta \end{bmatrix}$ . The centrifugal/Coriolis torque  $V_0(q, \dot{q})$  is given by  $V_0(q, \dot{q}) = \begin{bmatrix} -\eta \sin q_2 \dot{q}_2 & -\eta \sin q_2 \dot{q}_1 - \eta \sin q_2 \dot{q}_2 \\ \eta \sin q_2 \dot{q}_1 & 0 \end{bmatrix}$ . The



**Fig. 3.** The radial-basis-function network



**Fig. 4.** Control system block diagram

gravity  $G_0(q)$  is given by  $\begin{bmatrix} (\frac{1}{2}m_1 + m_2)ga_1 \cos q_1 + \frac{1}{2}m_2ga_2 \cos(q_1 + q_2) \\ \frac{1}{2}m_2ga_2 \cos(q_1 + q_2) \end{bmatrix}$ . The real model we use is (9), with  $M(q), V(q, \dot{q}), G(q)$  are all twenty percent less than that of  $M_0(q), V_0(q, \dot{q}), G_0(q)$  for simulation, the disturbance and friction item is chosen as

$$F(q, \dot{q}) = k_r \dot{q} \quad (22)$$

The following parameters are used in the simulation:  $m_1 = 15.2\text{kg}$ ,  $m_2 = 12.51\text{kg}$ ,  $a_1 = 0.42\text{m}$ ,  $a_2 = 0.41\text{m}$ . The parameters  $\alpha$ ,  $\beta$  and  $\eta$  in  $M_0(q)$  and  $V_0(q, \dot{q})$  are representative of  $\alpha = (\frac{1}{3}m_1 + m_2)a_1^2 + \frac{1}{3}m_2a_2^2$ ,  $\beta = \frac{1}{3}m_2a_2^2$  and  $\eta = \frac{1}{2}m_2a_1a_2$ . The coefficient associated with the friction is  $k_r = 1.2$ .

Treadmill exercise is a mostly often used movement to paraplegia or hemiplegia patients, and it has been proven very effective to their recovery. In our simulation, we also choose this movement style which means that the trajectory of the rehabilitation robot end is a circle. Specifically speaking, we choose the trajectory as

$$\begin{cases} x = 0.63 + 0.1 \cos(0.5\pi t) \\ y = 0.63 + 0.1 \sin(0.5\pi t) \end{cases}$$

The center of the circle is selected as  $(0.63, 0)$ , and the radius is  $0.1m$ . According to [13], we know that the human walking speed ranges from  $67 \pm 3$  (very slow) to  $154 \pm 11$  steps/min (very fast), and a normal person walks  $100 \pm 1$  steps/min at the usual walking speed. We set the simulation period as  $4s$ , and it is suitable for a injured person. From the end trajectory of the robot, we can get the desired trajectory of every joint. It can be written as followed:

$$\begin{cases} q_2^d = -\arccos \frac{x^2 + y^2 - a_1^2 - a_2^2}{2a_1a_2} \\ q_1^d = \arcsin \frac{y}{\sqrt{x^2 + y^2}} - \arctan \frac{a_2 \sin(q_2^d)}{a_1 + a_2 \cos(q_2^d)} \end{cases}$$

The  $q_1^d$  is the desired trajectory of joint 1, and  $q_2^d$  is the desired trajectory of joint 2.

The torque that human exert onto the pedal is chosen as 20% of the system output torque for simulation. The gain matrices  $K_p$  and  $K_v$  are selected diagonal as  $K_p = \begin{bmatrix} 16 & 0 \\ 0 & 16 \end{bmatrix}$ ,  $K_v = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}$  respectively. The number neurons of the RBF

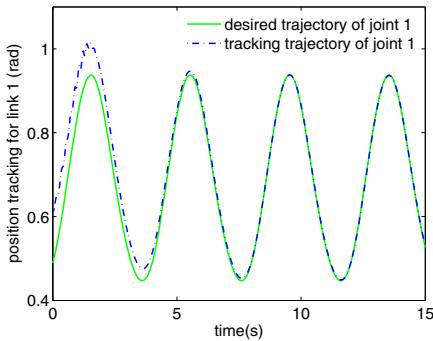
hidden layer we choose is four and the central vector is  $c = \begin{bmatrix} -4 & -2 & 0 & 2 \\ -4 & -2 & 0 & 2 \\ -4 & -2 & 0 & 2 \\ -4 & -2 & 0 & 2 \end{bmatrix}$ , the

base width of neurons is  $b_j = [5 \ 5 \ 5 \ 5]$ . The parameters in (20) are chosen as

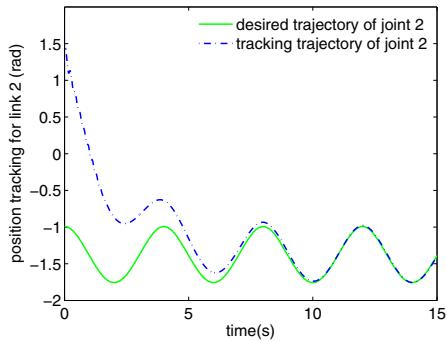
$\gamma = 100$ ,  $k_1 = 0.001$ , and the Lyapunov matrix as  $Q = \begin{bmatrix} 60 & 0 & 0 & 0 \\ 0 & 60 & 0 & 0 \\ 0 & 0 & 60 & 0 \\ 0 & 0 & 0 & 60 \end{bmatrix}$ . The

parameters above have different influences to the simulation. Specifically, the much bigger of the gain matrices  $K_p$  and  $K_v$  the faster convergence of trajectory. The  $\gamma$  and Lyapunov matrix  $Q$  is much bigger and the  $k_1$  is much smaller then the compensation for the unprecise part is more precise. Besides, the central vector and base width of neurons have small influence to the simulation. These conclusions are consistent with formula (13) and (21).

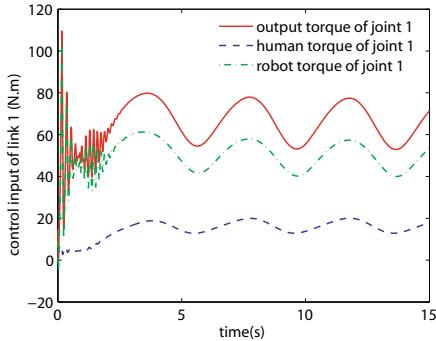
Figs 5 and 6 depict the position tracking results in 15 sec. The tracking error is almost less than 0.005 rad (0.287 deg) after 9 sec for both joints. Figs 7 and 8 show the output torque for tracking the desired trajectory. It can be seen that the output torque is the combination of human torque and robot torque. From figs 7 and 8 we also can see that the robot is working in assistance mode for the reason that the human torque and the robot torque are of the same direction. Once the human torque and the robot torque are of different directions, then the robot will work in resistive mode. Thus the assist-as-needed mode is accomplished. While in rehabilitation training for patients who have suffered from different degrees of disability, once the doctors have set the trajectory and the speed, the robot will decide to help or resist the patient according to the patients' degree of voluntary. If the patient can not control his lower limbs to follow the trajectory, then the robot will give him a hand. On the contrary, if the patient can not only



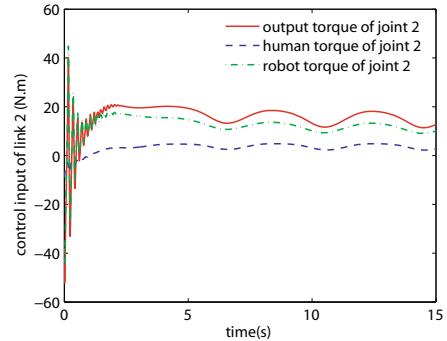
**Fig. 5.** Desired and actual trajectory of joint 1



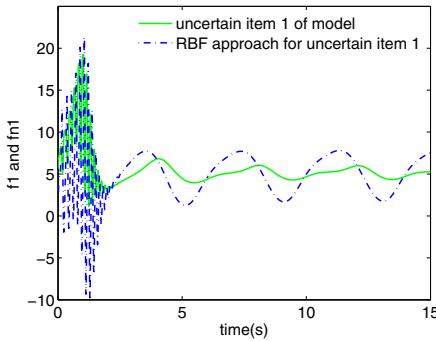
**Fig. 6.** Desired and actual trajectory of joint 2



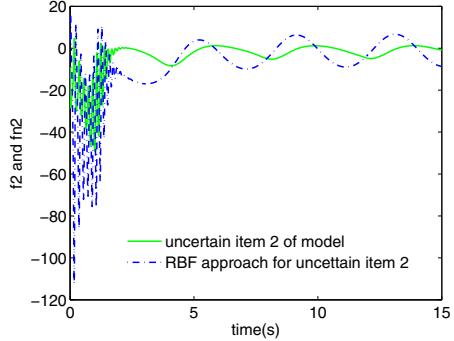
**Fig. 7.** Output torque of joint 1



**Fig. 8.** Output torque of joint 2



**Fig. 9.** RBF neural network approximation for uncertain item of joint 1



**Fig. 10.** RBF neural network approximation for uncertain item of joint 2

follow the trajectory but also want to move faster, then the robot will give him an resistance. So the whole system will always work in the predefined form the doctors have set. Figs 9 and 10 show the approximation to the unprecise part of the dynamics model as showed in formula (10).

## 4 Conclusion

In this paper, an adaptive RBF neural network combined with PD Computed-torque control algorithm is proposed for trajectory tracking of lower limbs rehabilitation robot. In this control algorithm, the RBF neural network is used to approximate the uncertain part of the dynamics model, including model error, friction and external disturbance, and the PD Computed-torque is used to control the known part of dynamics. From the simulation result we can see that the RBF approximation has a good performance and the control algorithm has a good robustness to the load changes and system friction. As is known to all that the rehabilitation robot load is different for different patients, and the RBF compensation can resolve the influence of physical difference. Furthermore, using this control algorithm, the robot can work in the assist-as-needed mode by detecting the joint torque. The control algorithm is just for simulation and has not been used in practice to date.

## Acknowledgements

This research is supported in part by the National Natural Science Foundation of China (Grant #60775043), the National Hi-Tech R & D Program (863) of China (Grant #2009AA04Z201), and the Sci. & Tech. for the Disabled Program of the Chinese Academy of Sciences (Grant #KGCX2-YW-618).

## References

- [1] Lotze, M., Braun, C., Birbaumer, N., Anders, S., Cohen, L.G.: Motor learning elicited by voluntary. *Brain* 126, 866–872 (2003)
- [2] Schoone, M., van Os, P., Campagne, A.: Robot-mediated Active Rehabilitation (ACRE) A user trial. In: IEEE 10th international conference on rehabilitation robotics, pp. 477–481 (2007)
- [3] Pohl, M., Werner, C., Holzgraefe, M., Kroczeck, G., Mehrholz, J., Wingendorf, I., Holig, G., Koch, R., Hesse, S.: Repetitive locomotor training and physiotherapy improve walking and basic activities of daily living after stroke: a single-blind, randomized multicentretrial (DEutsche GAngtrainerStudie, DEGAS). *Clinical Rehabilitation* 21, 17–27 (2007)
- [4] Hasan, M.K., Park, S.-H., Seo, S.-J., Sohn, D.-H., Hwang, S.-H., Khang, G.: A Gait Rehabilitation and Training System based on Task Specific Repetitive Approach. In: IEEE 3rd international conference on bioinformatics and biomedical engineering, pp. 1–4 (2009)

- [5] Yang, Q., Siemionow, V., Yao, W., Sahgal, V., Yue, G.H.: Single-trial EEG-EMG coherence analysis reveals muscle fatigue-related progressive alterations in corticomuscular coupling. *IEEE transactions on neural systems and rehabilitation engineering* 18(2), 97–106 (2010)
- [6] Hincapie, J.G., Kirsch, R.F.: Feasibility of EMG based neural network controller for an upper extremity neuroprosthesis. *IEEE transactions on neural systems and rehabilitation engineering* 17(1), 80–90 (2009)
- [7] Lunenburger, L., Colombo, G., Riener, R., Dietz, V.: Biofeedback in gait training with the robotic orthosis locomat. In: The 26th Annual International conference of the IEEE EMBS, vol. 7, pp. 4888–4891 (2004)
- [8] Den, A., Moughamir, S., Afilal, L., Zaytoon, J.: Control system design of a 3-dof upper limbs rehabilitation. *Computer methods and programs in biomedicine* 89, 202–214 (2007)
- [9] Jezernik, S., Scharer, R., Colombo, G., Morari, M.: Adaptive robotic rehabilitation of locomotion: a clinical study in spinally injured individuals. *Spinal cord* 41, 657–666 (2003)
- [10] Morita, Y., Hirose, A., Uno, T., Uchida, M., Ukai, H., Matsui, N.: Development of rehabilitation training support system using 3D force display robot. Springer, Heidelberg (2007)
- [11] Lewis, F.L., Dawson, D.M., Abdallah, C.T.: Robot manipulator control theory and practice. Marcel Dekker, Inc., New York (2004)
- [12] Park, J., Sandberg, I.W.: Universal approximation using radial-basis-function networks. *Neural Computation* 3, 246–257 (1991)
- [13] Mark, D.L., Hylton, B.M., Victor, S.F., Stephen, R.L.: Walking speed, cadence and step length are selected to optimize the stability of head and pelvis accelerations. *Experimental Brain Research* 184(2), 201–209 (2008)

# Model Complexity Reduction of the Hand Musculoskeletal System in Tremulous Motion

Ruwadi<sup>1</sup>, Shee Cheng Yap<sup>1</sup>, Philippe Poignet<sup>2</sup>, and Ang Wei Tech<sup>1</sup>

<sup>1</sup> Nanyang Technological University, Singapore

<sup>2</sup> Université Montpellier 2, France

{ruwa0002,cyshee,wtang}@ntu.edu.sg, philippe.poignet@lirmm.fr

**Abstract.** Musculoskeletal modeling is an important step in understanding the behavior of a body part for postural and motion control. A simple but reliable model is preferred over more complex models. Most of the musculoskeletal models that have been developed so far involved a number of parameters that sometimes, some of the parameters are not easily identified or require difficult or expensive procedures. In this paper, a simplification strategy of the musculoskeletal model in 1 degree of freedom (DOF) hand tremulous motion is presented. The key idea of the complexity reduction mainly on the combination of the two inputs into single input using a sequence of signal processing. 3 models (first order, second order and second order with one zero) plus time delay are considered to represent the simplified musculoskeletal model. The best fit was represented by a second order plus time delay and one zero.

**Keywords:** Musculoskeletal Modeling, Electromyography, Electromechanical Delay, EMG, EMD, Tremor, Tremulous Motion, Tremor Modeling.

## 1 Introduction

In general, mathematical modeling is necessary to provide understanding of the system behavior under certain condition. In control engineering, modeling is usually required to obtain suitable control input that will produce preferred plant output according to the control purpose. It is a great challenge to formulate a general mathematical model for a complicated system such us musculoskeletal system with nonlinearities characteristics. A customized model is usually developed to work well for a particular objective or application. One model may work well for one application but may not be suitable for another purpose. In order to obtain a good model that will serve the modeling purpose well, important characteristics of the system for that particular purpose have to be preserved.

In sport and rehabilitation engineering, musculoskeletal modeling is required for control the postural or motion of body parts either using motor – controlled or functional electrical stimulation (FES) – driven devices. It is usually developed to relate the muscle activity captured using electromyography sensor (EMG) to its corresponding measurable output (can be force, torque or motion parameters such as position, velocity or

acceleration). To date, musculoskeletal modeling remains one of the most challenging tasks. Not only that a musculoskeletal system is known to be highly nonlinear, its mechanism is also complicated to be fully-represented mathematically since it involves numbers of non-observable intermediate processes with non-measurable parameters.

In developing a good musculoskeletal model, numerous approaches have been proposed in the literature. The most commonly used method in musculoskeletal modeling is to model the muscle contraction using the well-known Hill-type muscle model [1-3], and Huxley sliding theory of muscle contraction [4, 5], and the variation of these two [6] and other approaches [7, 8]. Most of the models presented in the literature use EMG signals as the model input and force or torque as the model output. A few literatures considered the motion parameters as the model output, for example in [8]. There are no standard procedures in representing EMG signals as an input to the model. Some authors represented the EMG signals to muscle activations and then relate them to the muscle contractions mechanism [9, 10]. A force profile will then be predicted from the muscle contraction mechanism considering (most of the cases) the physiological characteristics of the muscle fiber.

From the modeling strategy of most available models, they are able to relate only one EMG signal as input and one output. In most of the tremulous motion cases, the motion is driven by two antagonistic muscle groups. This will yield a two-input single-input system. The current available approaches may also be problematic since an interaction effects when individual or both inputs present could not be evaluated easily. Hence, in this paper, we presented a model complexity reduction of the hand musculoskeletal model in tremulous motion that will reduce the number of EMG signals from the two antagonist muscle group as the musculoskeletal system input into single input.

## 2 Methods

The key feature in the model complexity reduction presented in this paper is on the strategy to merge the two EMG input signal as one input variable. The strategy lies in the signal processing sequence that will be described in the following sub-section. The two EMG input signals are represented into single parameters that have a close correlation with the motor output – acceleration data (ACC).

### 2.1 Data Acquisitions

In the current work, no tremor patient was recruited to participate in the experiment. One may comment the validity of the model reduction technique presented in this paper will not quite of a use when implemented to tremor patients. Burkhard, in his paper presented that voluntary simulated tremor by healthy subject shares common characteristics with pathological tremor in Parkinson's Diseases (PD) and Essential Tremor (ET) [11]. He examined that voluntary simulated tremor has indistinguishable frequency and regularity, but has relatively higher amplitude than tremor in PD and ET.

Tremulous motion data was recorded from healthy subjects performing hand tremor with predetermined frequency. The forearm was relaxed in a fixed support, while the hand was rested in a horizontally-free rotatable support (frictionless). The wrist was positioned just above the hinge of the rotatable support so as to make sure that the tremor imitated was 1 DOF and was free from gravitational acceleration effect. The gravitational effect, in this work, has to be minimized so as to maximize the relationship of the electrical activity of the muscles data recorded corresponds only to the motion observed. A beeping sound with a preset frequency was generated to guide the subject move his hand in a constant frequency set prior to the data collection.

The myoelectric signal of the muscle activity was recorded using EMG sensor and the motion parameter of the hand was captured using accelerometer sensor. Surface electrode EMG was used since we avoided invasive procedure during the data acquisition processes. Two differential electrodes (1 cm apart) were placed in the skin near the *Flexor Carpi Radialis* and *Extensor Carpi Ulnaris* muscles respectively with one electrode placed near the bone on the elbow as a common ground for both electrodes groups. An accelerometer was fixed on the back hand to record the corresponding movement – muscle contraction output.

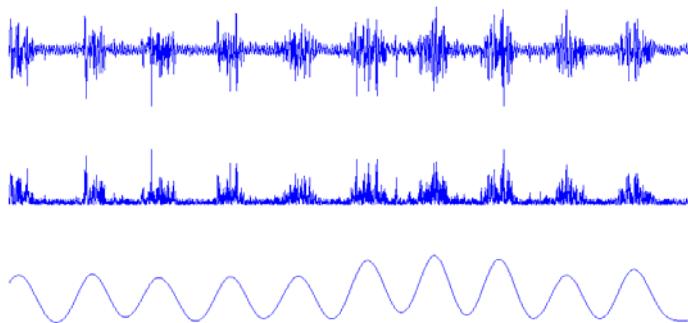
The EMG signals were sensed using Ag/AgCl circular surface electrodes F-301 from Skintact and amplified using EMG100C from Biopac Systems, Inc., USA, that was preset to pass 10-500Hz signals with 1000 gain. The acceleration data was captured using DE-ACCM3D accelerometer from Dimension Engineering, having sensitivity of 360mV/g ( $g =$  gravitational acceleration) and a bandwidth of 500Hz. Analog EMG signals and ACC data were captured with sampling frequency of 1000Hz via data acquisition NI DAQ CARD – 6036E from National Instrument and were interfaced using LabView 8.5 software also from National Instrument operated in windows PC. All data was processed offline using Matlab 7.1 from The MathWorks, Inc.

A slice of data from a subject simulation tremulous motion with preset frequency of 4 Hz will be used throughout the analysis in this paper. The subject was a male healthy subject with no prior diagnostic of motor or neurological disorder including tremor. The subject was 24 years old, 165 cm in height and weight of 65 kg when the data was collected.

## 2.2 Model Complexity Reduction via Signal Processing

EMG signal is generated from polarization-depolarization mechanism in the muscle fibers [1]. A polarization-depolarization process refers to the polarity changes in the muscle fiber for opening voltage-gated ion channel during muscle contraction. The signal captured by the recorder devices in this work is not the absolute value of the electrical potential, but the difference between the two recorded voltages. Hence, the raw EMG signal recorded will have positive-negative profiles. A temporal summation effect of the polarization-depolarization processes in muscle contraction makes the myoelectric signal generated could not be directly related to the corresponding muscle contraction output (force, torque or motion parameters). Hence, in this work, we propose a EMG signal processing sequence so as to be able to relate to the muscle contraction output (in this case acceleration) easily.

One of the key ideas in this work is lying in the signal processing strategy to combine the two EMG signals into single input variable. As described in the previous subsection, 2 inputs (EMG signals captured from *Flexor Carpi Radialis* and *Extensor Carpi Ulnaris* muscles) and 1 output (hand motion – acceleration data) were considered to model the 1-DOF tremulous motion of the hand. To reduce the number of input variables from two-input single-output system into single-input single-output model in this work, the two inputs were preprocessed and compared its relative value to obtain single input variable to the model, while the model output remained the same. The signal processing sequence of the two EMG signals will be described as follows:



**Fig. 1.** Signal processing sequence of the EMG signals. The raw EMG signal recorded by the electrodes was normalized to its maximum voluntary contraction (MVC) value obtained from maximum voluntary isometric muscle contraction (top). The normalized signal was then fully rectified by taking the absolute value of the signal (middle). The EMG signal envelope was obtained by passing the rectified signal through a low-pass filter (bottom).

### 1. Signal Normalization

EMG signals captured from one muscle cannot be directly compared to the signal captured from different muscle or different recording location. One of the reasons is due to the variation in volume conduction across different measurement spot from the muscle fiber to the skin. In addition, since surface electrodes were used, the crosstalk from the adjacent muscle fibers could not be quantified easily from one recording location to another. One method to increase comparability of EMG signals is by normalizing them to their respective maximum voluntary contraction (MVC). The MVC signals were recorded from isometric muscle contraction by forcing the subjects to flex (or extend) their hand to a static body so as to keep the deflection of the hand in a neutral position – hence isometric contraction.

### 2. Signal Rectification

Since EMG signal has positive – negative profiles, it is difficult to use it directly for further process due to the complex mechanism in the muscle contraction. Common practice is to rectify the signal and then to obtain the envelope of the rectified signal before apply the signal for further analysis. The principle of the rectification is elimination of the negative value of the signal. It can be done by zeroing the negative

value, taking absolute value of the signal, or squaring the signal. For this work, we obtained a full rectification by taking the absolute value of the signal. The rectified signal was further processed to get the envelope so it will have a better representation and correlation to the muscle contraction output (force or motion parameters).

### 3. Signal Filtering

There are no standard formulations in obtaining envelope from a signal. The idea of taking the envelopes is to get the outer profiles covering the signals. In this work, the envelopes of the rectified signals were obtained by passing the signals through zero phase low-pass filters. The filters are 3rd order Butterworth filter with cut-off frequency 2.5 times of the dominant tremulous motion frequencies (that have been set prior to the data collections).

### 4. Signal Combination

This is one of the key ideas of the model variables reduction in this work. Both EMG signals that have been processed were represented into single signal. To obtain the combined version of the signal, we took the relative difference of one preprocessed signal envelope to the other. This can simply be done by subtracting one signal from the other. Since the positive increment in the acceleration data refers to the extension direction (*Extensor Carpi Ulnaris* muscle contraction), the combined signal was obtained by subtracting the envelope of the EMG signal from extensor to the envelope of the EMG signal from flexor.

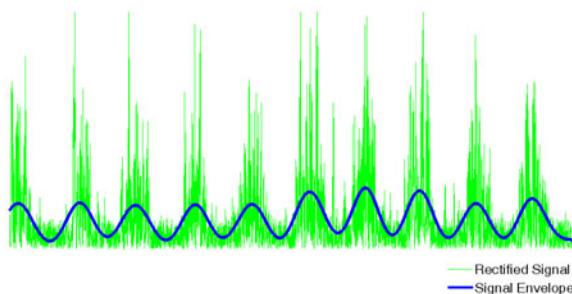
One may argue that direct signals subtraction may not be able to be justified due to the differences in the degree of contribution of each signal to the corresponding muscle response - output motion. It is to be noted that the signals has been normalized to their respective MVC, hence the normalized version of the signals represent their relative values to the maximum contraction they can produce. In addition, the model is to represent tremulous motion, sinusoidal-like motion with respect to its neutral motion (not to a neutral position of the hand about the wrist).

### 5. Signal Mean Removal

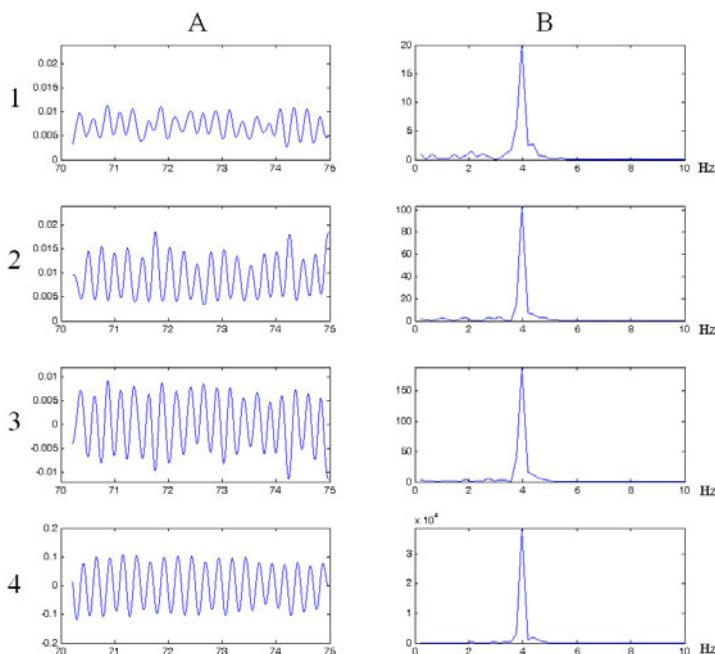
Tremulous motion was the main interest in this work. Tremulous motion has a roughly sinusoidal motion profile with periodic characteristics. It moves with certain frequency and observable amplitude. In this periodic tremulous motion, the neutral oscillation of the motion was set to be zero-value (not necessarily to be the neutral position of the hand, but the neutral position of the oscillation). Hence, the bias of the signal has to be removed.

This was performed by subtracting the signal by its mean. The most common practice to use EMG signal in musculoskeletal modeling is to apply a certain threshold value for the muscle to active (contract). By removing the signal mean, we will not consider a threshold value. In addition, since our motion of interest is a tremulous motion, we assume that the muscle contracts in a periodic tone. The resulting signal will be the input of the simplified model of the musculoskeletal model developed in this paper. In addition, signals mean removal will yield relative signal to a neutral value of the system output data.

The ACC signal processing was done to get rid of the noise outside the frequency band of interest. A 3<sup>rd</sup> order zero-phase band-pass filter was used to reject the noise signal contents in the ACC data. The band-pass filter consisted of a low-pass butterworth filter with cut-off frequency of 10 Hz and a high-pass butterworth filter with cut-off frequency of 1 Hz.



**Fig. 2.** Rectified EMG signal and its envelope obtained from low-pass filtering. The raw EMG signal was normalized to its maximum voluntary contraction (MVC) value, and then was full rectified by taking the absolute value of the signal. The rectified signal, due to the nature of the EMG signal generation from polarization-depolarization processes that acts as temporal summation in effect to the motion (force) produced, cannot be relate directly as a system input. The most commonly used method to make use of EMG signal is taking its signal envelope using low-pass filtering.



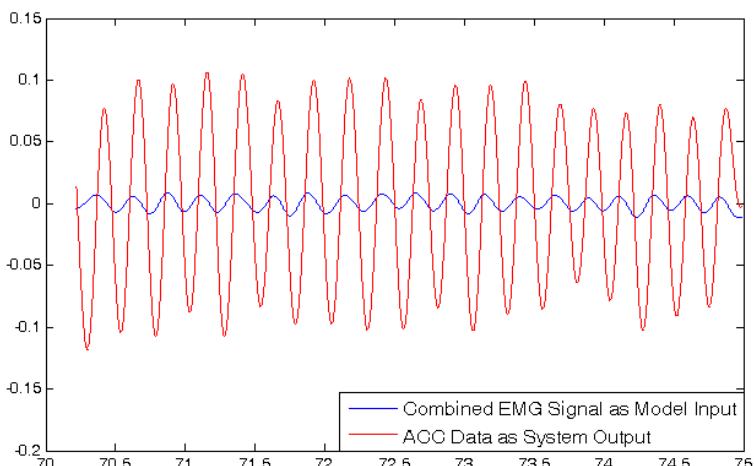
**Fig. 3.** Preprocessed EMG Signals and ACC signal in time and frequency domains. Comparisons of different signals representation in time domain are presented in A and the frequency contents of the corresponding signals are presented in B. 1 – EMG signal envelope from extensor muscle. 2 – EMG signal envelope from flexor muscle. 3 – The combined version of EMG signal obtained from subtracting the EMG signal envelope of the extensor muscle with EMG signal envelope of the flexor muscle. 4 – ACC signals of the tremulous motion.

### 3 Analysis

As described, the number of input variables reduction was achieved by the signal processing strategy. One may argue how reliable is the signal processing sequence presented in the previous section to combine the two EMG signals as one input to the musculoskeletal model. Before we apply the processed EMG signals to the model the musculoskeletal system, it is worth mentioning to look at the correlation of the input signals (the two EMG signals) and the output data (ACC signals).

Figure 3 shows a comparison of EMG signal envelopes from flexor, extensor, the combined result (extensor-flexor) as the input to the musculoskeletal model and the ACC signal – captured from the back hand as the output of the musculoskeletal model. From the figure, it is clear that the spectrum of the frequency contents in all signals have similar dominant frequency. Even though the frequency spectrum of the individual EMG signal 3(1) and 3(2) have slight larger frequency distribution than the ACC output 3(4), the frequency contents of the combined EMG signal 3(3) shows a closer profile relation to the ACC output. In the figure 3, the dominant frequencies of all the signals are at about 4 Hz.

The similarity in frequency contents between flexor and extensor shows that in this experiment, the tremulous motion was generated by the contraction of both muscles in an alternating sequence. The subtracted signal shown in figure 3 (3) shows a strong correlation with the ACC signals as shown in figure 3 (4). By observing the subtracted signal as the model input and the ACC signal as the model output as presented in figure 4, we could clearly see that the two signals have similar profile of sinusoidal, with roughly constant phase difference. From the figure, we may guess that the input-output relationship may have linear system characteristics of a first or second order system.



**Fig. 4.** Combined EMG signals obtained from the signal processing sequence and its corresponding motor output (ACC) signals

## 4 Model Transfer Functions Identification

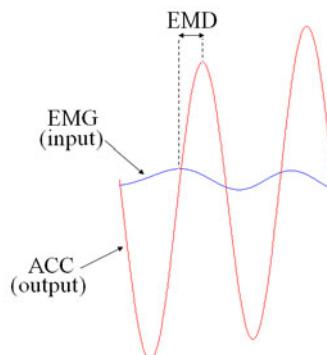
To validate the proposed strategy of model complexity reduction in the musculoskeletal modeling of 1-DOF hand tremulous motion presented above, system identification of the musculoskeletal model will be presented to estimate transfer function parameters in the first order, second order and second order system with one zero. Considering a strong profile correlation between the input (processed EMG signal) and the output (ACC signal), these three system transfer function were chosen to represent the input-output relationship of the musculoskeletal model. The frequency spectrum of the input (processed EMG signals) is similar to the output (ACC signal) as presented in figure 3. In addition, from observation, the similarity of the input-output profile becomes the basis for us to propose a linear system model.

Prediction error method system identification will be used to estimate the model transfer function parameters, but the time delay will be identified by observation from the input-output data.

### 4.1 Electromechanical Delay (EMD)

Electromechanical delay or known as EMD is defined as the time lag between the muscular electrical activity (EMG signals) and the motor output (usually force). In this work, since our parameter of interest is not a force, but the motion itself, the EMD will be re-defined to be the time delay between the electrical activity of the muscle and the motion generated.

The EMD, in this model parameters identification is an important measure to provide the value of the time delay between the input and the output in the transfer function of the musculoskeletal model. There number of EMD calculation methods available in the literature such as envelope estimation and thresholding, adaptive envelope estimation and thresholding and double threshold detector [12].



**Fig. 5.** Electromechanical delay between EMG signal (input) and the ACC signal (output) as the time delay in the model transfer function

In this work, the EMD was calculated by measuring the time differences between the peak of the input and the peak of the output. Figure 5 illustrates how the EMD was obtained. The EMD obtained will be used as a the time delay in the transfer function of the models.

From a slice of 4 Hz frequency tremulous data presented in figure 4, there were a distribution profiles in the EMD. The value of EMD from one peak is not the same to the EMD of the next peak. This is probably because of the frequency spectrum of the motion is not exactly at 4Hz but has a small distribution at about 4Hz as shown in figure 3. In addition, the EMD is also affected by mechanical factors such as muscle loading and muscle length that will change the stiffness of the musculoskeletal system, muscle fiber composition, firing rate of the muscle contraction, fatigue, etc [12].

From the same slice of data, the mean of the EMD was 38.63ms with standard deviation of 8.04ms and range of 31.00ms. The mean value of EMD will be used as the time delay in the system identification step and it was assumed to be constants.

## 4.2 Simulation Results

In all the three models that will be identified, the time delay will be fixed and assumed to be constant. The value of the time delay is the mean value of the EMD obtained from previous section, i.e. 38.63ms. The poles and zeros were estimated using the prediction error method and applied in Matlab. Validation of the transfer function was done via simulation using a data of a 4Hz tremulous motion. The simulation results of the three transfer function are presented in figure 6. The transfer functions estimated are as follows:

First order plus time delay:

$$G1 = \frac{11.37}{0.001s + 1} e^{-0.03863s}$$

Second order plus time delay:

$$G2 = \frac{11.73}{0.000001s^2 + 0.002s + 1} e^{-0.03863s}$$

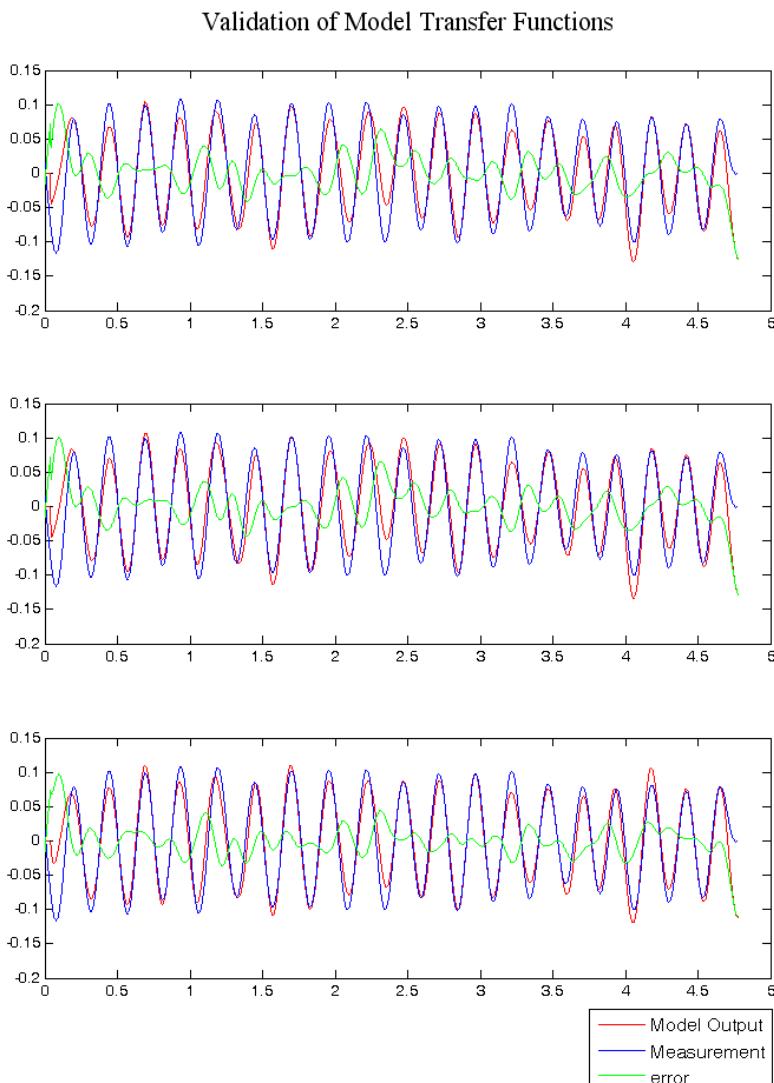
Second order plus time delay and one zero:

$$G1 = \frac{0.7545s + 4.146}{0.0009069s^2 + 0.06033s + 1} e^{-0.03863s}$$

## 5 Discussion

Although EMG signal has been widely used as an input in the musculoskeletal model and as a control input to a system, there are no standard procedures to process the signal to extract certain information required. Signal envelope is the most common form of EMG signal to use for further analysis. Even though it is the most commonly used, there are different approaches to get the envelope of the EMG signal.

The novelty proposed in this paper is to combine two EMG signals from agonist-antagonist muscle groups into single parameter. This new approach will bring significant model complexity reduction in the musculoskeletal modeling in tremulous motion especially when a control system has to be designed for postural or motion control related to tremulous motion or roughly periodic motion.



**Fig. 6.** Simulation results of the musculoskeletal SISO models represented in first order system plus time delay (top), second order system plus time delay (middle) and second order system plus time delay and a zero (bottom)

Even though the signal analyzed in this paper was taken from a tremulous motion that was driven by alternating contraction of two agonist-antagonist muscle group, the same signal processing sequence may also be implemented in a tremulous motion by which the motor is one muscle group only. In this regard, signal mean removal played an important role since the absolute position was not considered. Instead, a relative position of the tremulous motion was set as the neutral position of the oscillation. With this method, the non-alternating pace of the agonist-antagonist muscle group contractions problem could be tackled.

From the simulation results of model transfer function representing the musculoskeletal system as shown in figure 6, the best of the three systems is second order plus time delay and one zero (bottom). First order plus time delay model performs similar to the second order plus time delay model. From Matlab, the best fit of the two are about 59%, while additional zero on the second order plus time delay model increases the best fit to 70%.

Related to EMD, since the value obtained from a slice of data used in the analysis has quite large distribution (having a mean of 38.63ms with standard deviation of 8.04ms and range of 31.00ms), assumptions of constant EMD value may produce mismatch between model output and measurement output. An adaptive numerical method to calculate EMD may be required if a more robust model is preferred.

By judging from the good fit of the simulation results of the modeled musculoskeletal systems in a SISO model, it opens prospective uses of the method in developing system for pathological tremor control in the upper limb. But physiologically, the model parameter has no physical relationship except for the EMD. But this should be enough to provide understanding in the output behavior (hand motion) prediction from a known input (captured EMG signals) for control system development in tremulous motion.

## 6 Conclusion

An EMG signals processing sequence has been presented to combine 2 EMG signals input into single input variable in a tremulous motion driven by two antagonistic muscle groups. Signal characteristics of the individual signal envelope with the combined version are similar in the sense that frequency spectrum and the amplitude profiles. A validation of the model variables reduction has also been presented to see the feasibility of the signal processing sequence to simplify the musculoskeletal system modeling. A good profile fit was observed in the combined EMG input signal and the ACC output and can be well modeled in a second order system plus time delay and a zero with best fit measure of about 70%.

In the future, a validation to the proposed methods will be carried out using real tremor data collected from tremor patients. In addition, a rigorous system identification procedure will be conducted to obtain the best representation of the simplified musculoskeletal system model. From the proposed strategy, an implementation of musculoskeletal system model simplification presented in this paper will be expected to be able to predict the tremor motion and use the information to control functional electrical stimulation (FES) to reduce pathological tremor.

## References

- [1] Hof, A.L.: EMG and muscle force: An introduction. *Human Movement Science* 3, 119–153 (1984)
- [2] Langenderfer, J., LaScalza, S., Mell, A., Carpenter, J.E., Kuhn, J.E., Hughes, R.E.: An EMG-driven model of the upper extremity and estimation of long head biceps force. *Computers in Biology and Medicine* 35, 25–39 (2005)
- [3] Hof, A.L., Van den Berg, J.: EMG to force processing I: An electrical analogue of the hill muscle model. *Journal of Biomechanics* 14, 747–753, 755–758 (1981)
- [4] Bestel, J., Clément, F., Sorine, M.: A Biomechanical Model of Muscle Contraction. In: Niessen, W.J., Viergever, M.A. (eds.) *MICCAI 2001. LNCS*, vol. 2208, pp. 1159–1161. Springer, Heidelberg (2001)
- [5] Dijkstra, S., Denier van der Gon, J.J., Blangé, T., Karemaker, J.M., Kramer, A.E.J.L.: A simplified sliding-filament muscle model for simulation purposes. *Biological Cybernetics* 12, 94–101 (1973)
- [6] Hayashibe, M., Guiraud, D., Poignet, P.: EMG-based neuromuscular modeling with full physiological dynamics and its comparison with modified hill model. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2009, pp. 6530–6533 (2009)
- [7] Bean, J.C., Chaffin, D.B., Schultz, A.B.: Biomechanical model calculation of muscle contraction forces: A double linear programming method. *Journal of Biomechanics* 21, 59–66 (1988)
- [8] Wexler, A.S., Jun, D., Binder-Macleod, S.A.: A mathematical model that predicts skeletal muscle force. *IEEE Transactions on Biomedical Engineering* 44, 337–348 (1997)
- [9] Terry, K.K.K., Arthur, F.T.M.: Feasibility of using EMG driven neuromusculoskeletal model for prediction of dynamic movement of the elbow. *Journal of electromyography and kinesiology: official journal of the International Society of Electrophysiological Kinesiology* 15, 12–26 (2005)
- [10] David, G.L., Thor, F.B.: An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo. *Journal of Biomechanics* 36, 765–776 (2003)
- [11] Burkhard, P.R., Langston, J.W., Tetrud, J.W.: Voluntarily simulated tremor in normal subjects. *Neurophysiologie Clinique/Clinical Neurophysiology* 32, 119–126 (2002)
- [12] Conforto, S., Mathieu, P., Schmid, M., Bibbo, D., Florestal, J.R., D'Alessio, T.: How much can we trust the electromechanical delay estimated by using electromyography? In: 28th Annual International Conference of the IEEE in Engineering in Medicine and Biology Society, EMBS 2006, pp. 1256–1259 (2006)

# A Wearable Rehabilitation Robotic Hand Driven by PM-TS Actuators

Jun Wu, Jian Huang, Yongji Wang, and Kexin Xing

Key Laboratory of Image Processing and Intelligent Control,

Department of Control Science and Engineering,

Huazhong University of Science and Technology, 430074, Wuhan, China

[huang\\_jan@mail.hust.edu.cn](mailto:huang_jan@mail.hust.edu.cn)

**Abstract.** Robotic-assisted therapy is of great benefit to the recovery of motor function for the patients survived from stroke. However there have been few emphases on the patients' hand training/exercise during the rehabilitation process. The goal of this research is to develop a novel wearable device for robotic assisted hand therapy. Unlike the traditional agonist/antagonist PM actuator, we propose a new PM-TS actuator comprising a Pneumatic Muscle (PM) and a Torsion Spring (TS) for joint drive. Based on the proposed PM-TS actuator, we design a robotic hand which is wearable and provides assistive forces required for finger training. The robotic hand has two distinct degrees of freedom at the metacarpophalangeal (MP) and proximal interphalangeal (PIP) joints. The variable integral PID (VIPID) controller was designed to make the joint angle of robotic hand can accurately track a given trajectory. The results show that the VIPID controller has better performance than the conventional PID controller. The proposed rehabilitation robotic hand is potentially of providing supplemental at-home therapy in addition to the clinic treatment.

**Keywords:** Rehabilitation Robotic Hand, Pneumatic Muscle (PM), Torsion Spring (TS), PM-TS Actuator.

## 1 Introduction

Adult disability is a common result of stroke in many countries, and usually involves deficits of motor functions. In the case of stroke victims, it is widely accepted that spontaneous recovery accounts for the motor and functional restoration taking place within the first months after the stroke incident. The findings have shown that intense practice of repetitive movements can help to improve the strength and functional recovery of the affected arm/hand. However, traditional approach of rehabilitation training is very labor intensive and lack of consistency and objective assessment.

Many kinds of assist devices for hand therapy have been developed to offer many patients the intensive training, such as the Rutgers Hand Master II-ND Force-Feedback Glove [1], the HWARD robotic hand-therapy device [2], the finger therapy robot at the Rehabilitation Institute of Chicago [3], the GENTLE/G hand robotic therapy system developed at the University of Reading [4], and the cabledriven finger therapy robots [5] which have been successful in providing grasp assistance and hand

opening. Unfortunately, it should be noted that many robotic systems are developed for research purpose, these robotic hands/devices are too heavy, complex, bulky and unwearable, or too expensive for home use by individuals. Therefore, there is a need for affordable, safe and practical devices to assist therapy for practical use.

In robotic control systems, most researchers select motors as driving sources for rehabilitation training [6][7]. Recently, the compliant pneumatic muscle actuators (PMA) have found many applications in the robot-based rehabilitation therapy. A seven degree-of-freedom upper arm training/rehabilitation (exoskeleton) system driven by PMA was designed [8]. Compared with the electric motor actuator, which is lack of the necessary compliance between the actuator and the limb being moved in exoskeleton applications, PMA is soft and exhibits many properties of human muscle. It is more important that PMA is at a very low level of risk of injury. This is not the case for hydraulic and/or electric motor actuators, which produce a far greater risk of injury to the user adjacent to the device when failing. PMA possesses many unique advantages (e.g. cheapness, light weight, compliance). At the same time it has very high power/weight and power/volume ratios [9][10]. Therefore, the hand therapy devices with PMAs can interact with the patients in a safer and more natural way.

A challenge for the application of PMA is that it can be only operated in the contractile direction. Hence, PMs have to be used in antagonistic pairs to achieve bi-directional motion for a single joint. Antagonistic pairs of actuators are usually an acceptable design solution for a simple flexion and extension joint, like the arm or ankle [8][11]. In many cases, we hope to use a compact structure to realize the bi-directional motion of joint driven by the PM. If the PMs are used in antagonistic pairs, it will not only increase the complexity of mechanical design and control, but also make the device/system bulk, weight and even the cost greatly increased. An available way to meet the requirement is to minimize the number of PMs in the actuators. For the above reasons, we proposed a new actuator structure comprising a PM and a torsion spring (TS), called PM-TS actuator, and designed a novel wearable robotic hand for assisted repetitive driven by the new PM-TS actuator.

In the conventional PID control, the fixed PID parameters cannot guarantee a satisfactory performance in the entire control process. Some studies demonstrate that better control performance may be attained when using a PID control strategy with variable coefficient other than a conventional PID method. Thus, in this study we designed a variable integral PID (VPID) controller to control the proposed PM-TS actuators. Experimental results found that the VPID performs better in our application.

In summary, the contributions of this study include

- 1): proposing a new PM-TS actuator for joint drive,
- 2): designing a wearable robotic hand for rehabilitation training,
- 3): implementing the control of the rehabilitation robotic hand by using VPID,
- 4): integrating the robotic hand with the sensor system with virtual reality (VR).

This paper is organized as follows. Section 2 proposes a new PM-TS actuator and describes the mechanics of the wearable rehabilitation robotic hand. Section 3 analyzes the control structure of the entire rehabilitation system. In section 4, the experimental results of the trajectory tracking control of the proposed PM-TS driven rehabilitation robotic hand are presented. In the last section, the conclusion of the current work and the possible future improvements are described.

## 2 Mechanics of Wearable Rehabilitation Robotic Hand

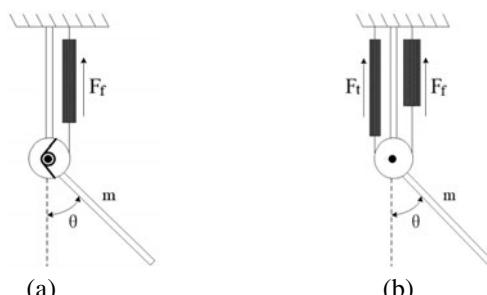
The design goal is to develop a therapeutic robotic hand that provides finger training. The robotic hand needs to be compact for portability, relatively easy to use for don and doff, capable of interacting with a personal computer based visual biofeedback system to capture the interest of the user, objective to evaluate the functional performance and has the potential to be inexpensive and amenable for home use.

### 2.1 Mechanic Structure

The “Mckibben” pneumatic actuators used in the joint drive provide compliant actuation and thus augment the safety compared to the traditional rigid drivers. PMA is made from inexpensive materials, such as natural latex rubber for the inner bladder covered by a polymer based braid (Fig. 1). The maximum contraction ratios of PMAs which are made in our Lab are about 25%. In order to transform the translation into the rotation, a pulley mechanism is designed and arranged above the joints. In traditional way, PMs have to be used in antagonistic pairs to achieve bi-directional movement for a single joint as shown in Fig. 2(b). This arrangement requires the control of at least two actuators simultaneously, and it will increase the complexity of mechanical design and control, the device’s bulk, weight and cost. An available way to meet the requirement is to minimize the number of PMs in the actuators. For these reasons, we proposed a new actuator structure comprising a PM and a torsion spring, called PM-TS actuator. The PM is arranged in the suitable place to provide the pulling moment and the torsion spring provides opposing torque, it is shown in Fig. 2(a).



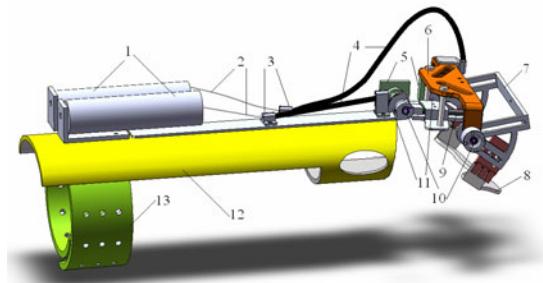
**Fig. 1.** Materials of pneumatic muscle actuator and pneumatic muscle actuator made in lab



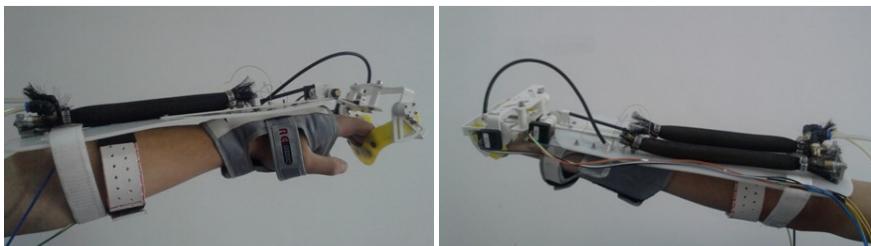
**Fig. 2.** (a) 1-DOF joint driven by PM-TS actuator. (b) 1-DOF joint driven by a pair of PMs.

Based on the proposed PM-TS actuator, we designed a rehabilitation robotic hand for hand therapy. Fig. 3 shows the 3D model of the wearable robotic hand, which is designed in Solidworks. The major design goals are summarized as follows:

- 1) There are 2 DOFs (Degrees of Freedom) for the fingers, which makes the movement of fingers more suitable and flexible.
- 2) The rotation scope of joint angle is limited by the mechanism for safety.
- 3) There is a bi-directional movement for each DOF.
- 4) The robotic hand can be worn comfortably and easily.



**Fig. 3.** The 3D model of the wearable robotic hand. 1-pneumatic muscle, 2-steel wire, 3- wire rack, 4-wire conduits, 5-angle sensor, 6-press plate, 7-finger splint link, 8-finger splint, 9-glide mechanism, 10-torsion spring, 11-pulley, 12-forearm attachment, 13-forearm hoop.



**Fig. 4.** The prototype of the wearable rehabilitation robotic hand

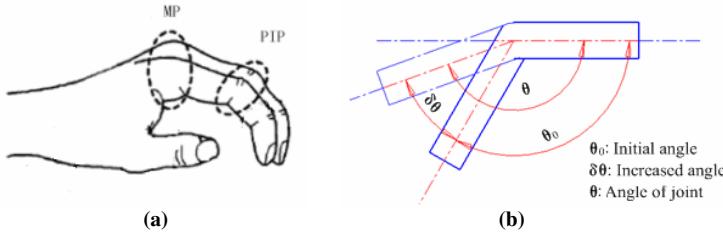
Fig. 4 shows the prototype of the wearable rehabilitation robotic hand worn on the user's left hand. The PMs are arranged on the top of the forearm attachment and the torsion springs are mounted on the joints. The steel wires are used to connect the PMs and the pulleys. The steel wires driven by PM-TS actuators lead to the rotation of the pulleys, which results in the rotation of finger joints. The wire conduit can avoid the movement coupling of the DOFs. A press plate is placed above the proximal phalanges to ensure that the proximal phalanges are parallel to the corresponding mechanical link. A glide mechanism is assembled on the press plate to adapt the relative movement between exoskeleton and human fingers. There are two Hall-effect encoders mounted to the axes of each joint, the Hall-effect encoders are used to measure the angles of the joints.

## 2.2 Fundamental Motion

The extensions of the metacarpophalangeal (MP) and proximal interphalangeal (PIP) joints of fingers in able-bodied volunteers with the wearable robotic hand are done as the fundamental experiment. Table 1 shows the joint angles are measured when the robotic hand works, where  $\theta_0$  means the initial angle,  $\delta\theta$  represents the increased angle and  $\theta$  is the current angle.  $\theta_0$ ,  $\delta\theta$  and  $\theta$  are illustrated respectively in Fig. 5.

**Table 1.** The ranges of motion for the robotic hand

	PIP[rad./deg.]	MP[rad./deg.]
$\theta_0$	1.5709/90	1.919/110
$\Delta\theta$	1.5709/90	1.2211/70
$\theta$	3.1416/180	3.1416/180



**Fig. 5.** (a) Disposition of joints. (b) Initial angle and increased angle for the fingers.  $\theta$  is the joint angle.  $\theta_0$  means the initial angle, and  $\delta\theta$  represents the increased angle.

## 2.3 Safety

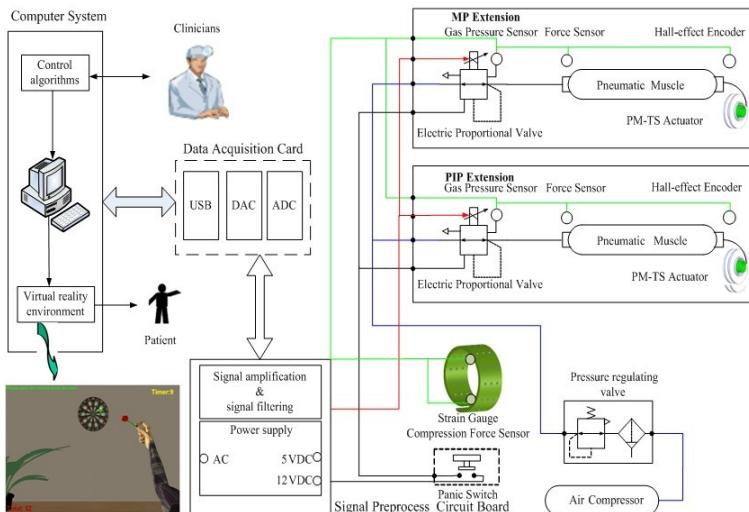
For a rehabilitation robotic hand, the safety is dominant within all performances. The affected hand needs to be absolutely safe when interacting with the robotic hand. Firstly, the risk to patients is minimized by utilizing the compliant PMAs. All assisted movements will be made slow to protect against significant increases in muscle tone. Secondly, the mechanical design allows the device physically to limit range of motion of individual joints. Moreover, the safety envelope in the control algorithm will further guarantee the safety. If the calculated control pressure for the PM exceeds the preset limit, the real control pressure will be set to the limit value. In addition, a panic button is provided for easy access by the unaffected hand to exhaust the PMAs and shut the device down if the patients are feeling any discomfort, fatigue or anxiety.

## 3 Control System

Task-oriented repetitive movements can improve muscular strength and movement coordination in patients with impairments due to neurological problems. A typical repetitive hand functional movement consists of grasping (flexion) and release (extension). Fig. 6 shows the block diagram of control system for the robotic hand.

The final goal of control scheme is to provide controllable, quantifiable assistance specific to some particular patients by adapting the level of the assistance provided.

The sensors can feedback angle, pressure and force information for adaptation of the assistive force and quantitative evaluation of task performance. The angle sensors are used to measure the joint flexion and extension. The robotic hand with the strain gauge compression force sensors involves the patients in their hand rehabilitation by encouraging self-powered motion by the fingers, and assisting movement when necessary. In addition, there is also a pressure sensor integrated in the electric proportional valve to provide the local feedback. The real-time data acquisition card is responsible for sampling all sensory data. The data is exchanged via a USB connection between the data acquisition card and the host computer. The analog output of the card is connected to the electric proportional valve to control the PMAs. The air transmission unit is an air-conditioning circuit which contains a filter valve and an electric proportional valve. The host computer runs the control algorithms and a virtual reality game. It can build a task-oriented interactive game environment to deliver an engaging and interesting training process. The evaluation of task performance is fed back to patients by intuitive cues integrated with the VR games.



**Fig. 6.** Block diagram of control system for the rehabilitation robotic hand

To realize the angle tracking, an incremental PID controller is firstly implemented in the host computer. The incremental PID control algorithm is described by

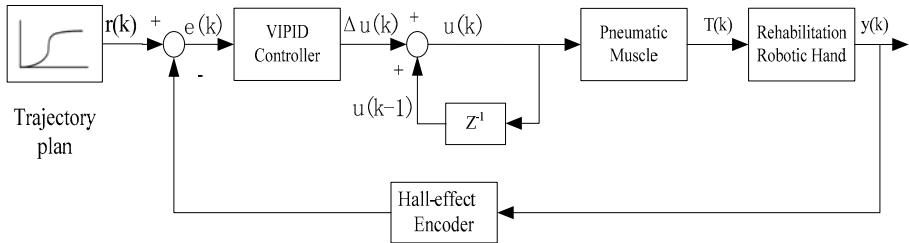
$$\Delta u(k) = k_p [e(k) - e(k-1)] + k_i e(k) + k_d [e(k) - 2e(k-1) + e(k-2)] \quad (1)$$

where  $\Delta u(k)$  is the control increment at the  $k$ -th sampling time.  $e(k)$  is the deviation of the system at the  $k$ -th sampling time.  $k_p$  is the proportional gain.  $k_i = k_p * T/T_i$  is the integral coefficient, and  $k_d = k_p * T_d/T$  is the derivative coefficient.

In the incremental PID control algorithm, the integral coefficient  $k_i$  is a constant. If the integral coefficient  $k_i$  is too small while the deviation is large, the incremental control variable may not be enough at some inflection points. On the other hand, if  $k_i$  is too large, the control process of the system may be unstable. Thus, it is important to

improve the system quality by changing the integral coefficient  $k_i$  according to the value of deviation. To attain better control performance, variable integral PID (VIPID) controller in an incremental way was applied. The control structure of the proposed rehabilitation robotic hand is depicted by Fig. 7. The relationship of integral coefficient  $f[e(k)]$  and error  $e(k)$  for VIPID can be designed as follows:

$$f[e(k)] = \begin{cases} k_i - \text{Min} & |e(k)| \leq B \\ k_i - \text{Min} + \frac{(k_i - \text{Max} - k_i - \text{Min})}{A} (|e(k)| - B) & B < |e(k)| \leq A + B \\ k_i - \text{Max} & |e(k)| > A + B \end{cases} \quad (2)$$



**Fig. 7.** The control structure of rehabilitation robotic hand with trajectory plan

Then the VIPID control algorithm is given by

$$\Delta u(k) = k_p [e(k) - e(k-1)] + f[e(k)]e(k) + k_d [e(k) - 2e(k-1) + e(k-2)] \quad (3)$$

## 4 Experiment

### 4.1 Experimental Condition

Fig. 4 and Fig. 6 show the experimental apparatus and schematic diagram for the control of the wearable rehabilitation robotic hand. The experimental apparatus includes pneumatic muscles, the electromagnetic proportional valve, pressure regulating valve, Hall-effect encoder, and the acquisition cards. Table 2 and 3 show the equipments and physical parameters of pneumatic muscles.

The parameters of conventional PID controller are:  $k_p=0.02$ ,  $k_i=0.005$  and  $k_d=0.02$ . The parameters of VIPID controller are:  $k_p=0.02$ ,  $k_d=0.02$ ,  $k_i-\text{Min}=0.005$ ,  $k_i-\text{Max}=0.02$ ,  $A=0.05$  and  $B=0.025$ . The two control algorithms were respectively implemented in the control system. Two control loops run independently for the two PM-TS actuators which drive the MP and the PIP joints to rotate. The controlled variable for each control loop is the measured angle at each joint. The software running on the PC was developed using Visual C++ 6.0.

### 4.2 Experimental Results

In order to further assess the performance of the system and the realization of the ultimate clinical application, we conducted preliminary clinical trials, as shown in

**Table 2.** The main performance parameters of the equipments

Name	Model	Performance	Manufacturer
Mute air compressor	FB-0.017/7	Rating exhaust pressure: 0.7Mpa	Taiwan JAGUAR Co., Ltd
Electromagnetic proportion valve	ITV1030-211BS	Input: 0-5V Output: 0.005-0.5Mpa	SMC Co., Ltd
Pressure regulating valve	AW20-02BCG	Range: 0.05-0.85Mpa	SMC Co., Ltd
Data acquisition card	USB-4716	8 channel DI, 8 channel DO 16 channel AI, 2 channel AO Sampling rate: 200 kS/s	Advantech Co., Ltd
Switching power supply	Q-120DE	Input: AC 220V±10% 50Hz Output: ±5V, ±12V, ±24V	Taiwan MEAN WELL
Hall-effect encoder (Angle sensor)	ED-18-SB-0050-V-P	Degree range: 0-360° Output: DC 0-5V Bushing mounting torque: 10 in-lb max	Measurement Specialties, Inc.
Strain Gauge Force Sensor	FSR400	Sensing area: 0.2" circle Force sensitive rang: 0 to 10kg	Interlink Electronics
Others		Joints connector, air pipe, etc	SMC Co., Ltd

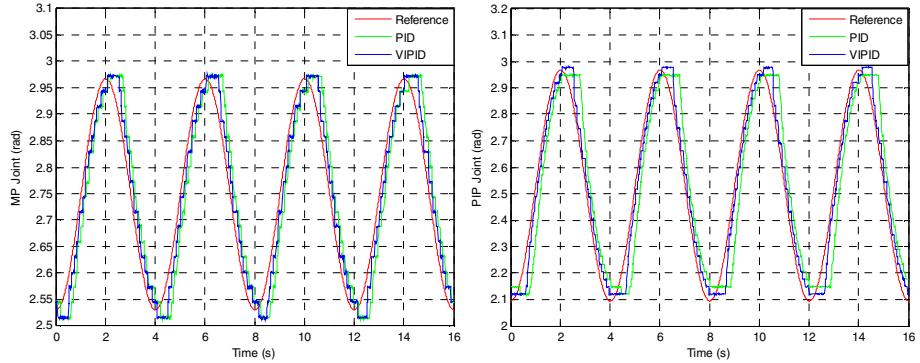
**Table 3.** The parameters of pneumatic muscles for MP and PIP

Joint	MP	PIP
Initial length $L_0$	211mm	232mm
Initial diameter $D_0$	12.26mm	12.26mm
Initial angle of mesh grid $\theta_0$	22°	22°
Thickness of rubber sleeve $t_k$	1.64mm	1.64mm

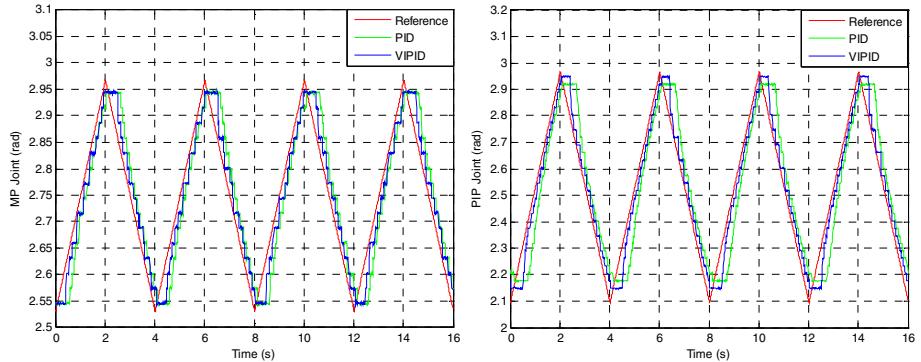
Fig. 8. Some experiments were performed on a healthy subject, the performance of PID and VIPID controllers were evaluated using sine and ramp waves as the desired joint angle trajectory that should be tracked for the extension/flexion of the MP joint and PIP joint. The user was passive and not working with or against the motion.

**Fig. 8.** The sequence of a clinical rehabilitation trial

The sine trajectory tracking results of MP and PIP using the conventional PID and VIPID control strategies are showed in Fig. 9. The ramp trajectory tracking results of PIP using the two control strategies are showed in Fig. 10. The self-tuning of VIPID parameters can be realized by using the error, it prevails over the PID in many performances, such as response time, flexibility, adaptability and control precision. But the jitter of MP is bigger than that of PIP during the extending. The main reason is that the friction force of MP has bigger influence than that of PIP, and the influence cannot be eliminated completely in the control system.



**Fig. 9.** The sine trajectory tracking result of MP and PIP using the PID and VIPID



**Fig. 10.** The ramp trajectory tracking result of MP and PIP using the PID and VIPID

## 5 Discussion and Conclusion

In this paper we proposed a new PM-TS actuator configuration that uses a PM and a torsion spring for bi-directional movement of joint. A wearable rehabilitation robotic hand actuated by the proposed PM-TS actuator is also developed for the stroke patients. It is designed with the intention of becoming a low cost and safe take-home supplemental therapy device for repetitive hand therapy. It is obviously more

easily-realized and compact than traditional actuators composed of PMs. In this way, the number of components which are often used in air circuit such as the expensive electric proportional valve and the fussy windpipe, can be greatly decreased. This meets the requirement of most rehabilitation devices – building lightweight, more compact, compliant and economical system. Further, this arrangement requires fewer actuators to be controlled. Finally, to realize the accurate tracking control of the designed rehabilitation robotic hand, the PID and VIPID controllers were designed to track some desired trajectories. The results show that the VIPID controller has better performance than the PID controller in trajectory tracking.

The future work is the development of an intelligent adaptive controller to encourage the patients' voluntary efforts during task-oriented rehabilitation training. More experiments will be implemented to improve the accuracy and stability of the angle/position control. Other control modes considering the measured force will be developed for more flexible training ways. The current robotic hand presented in the preliminary experiment can be further improved in mechanism and intelligence in the future. It is expected that this system will eventually: 1) empower patients' voluntary for intention control; 2) support therapists in a labor intensive task; 3) reduce dependency on hospital resources; and 4) assist in re-integrating patients with hand impairment back into the community.

## Acknowledgments

This work is partially supported by Hi-tech Research and Development Program of China under Grant 2007AA04Z204 and Grant 2008AA04Z207, and in part by the Natural Science Foundation of China under Grant 60674105 and 60975058.

## References

1. Bouzit, M., Burdea, G., Popescu, G., Boian, R.: The Rutgers Master II—New design force-feedback glove. *IEEE ASME Trans. Mechatron.* 7, 256–263 (2002)
2. Takahashi, C.D., Der-Yeghaia, L., Le, V.H., Cramer, S.C.: A robotic device for hand motor therapy after stroke. In: 9th IEEE International Conference on Rehabilitation Robotics Conference, pp. 17–20. IEEE Press, Chicago (2005)
3. Worsnopp, T.T., Peshkin, M.A., Colgate, J.E., Kamper, D.G.: An actuated finger exoskeleton for hand rehabilitation following stroke. In: 10th IEEE International Conference on Rehabilitation Robotics, pp. 896–901. IEEE Press, Noordwijk (2007)
4. Loureiro, R.C., Harwin, W.S.: Reach & grasp therapy: Design and control of a 9-DOF robotic neuro-rehabilitation system. In: 10th IEEE International Conference on Rehabilitation Robotics, pp. 757–763. IEEE Press, Noordwijk (2007)
5. Dovat, L., Lamberg, O., Johnson, V., Salman, B., Wong, S., Gassert, R., Burdet, E., Leong, T.C., Milner, T.: A cable driven robotic system to train finger function after stroke. In: 10th IEEE International Conference on Rehabilitation Robotics, pp. 222–227. IEEE Press, Noordwijk (2007)
6. Lamberg, O., Dovat, L., Gassert, R., Burdet, E., Chee, L.T., Milner, T.: A Haptic Knob for Rehabilitation of Hand Function. *IEEE Trans. Neural Syst. Rehabil. Eng.* 15, 356–366 (2007)

7. Mulas, M., Folgheraiter, M., Gini, G.: An EMG-controlled Exoskeleton for Hand Rehabilitation. In: 9th IEEE International Conference on Rehabilitation Robotics, pp. 371–374. IEEE Press, Chicago (2005)
8. Tsagarakis, N.G., Caldwell, D.G.: Development and Control of a ‘Soft-Actuated’ Exoskeleton for Use in Physiotherapy and Training. *Autonomous Robots* 15, 21–33 (2003)
9. Caldwell, D.G., Medrano-Cerda, G.A., Goodwin, M.: Control of pneumatic muscle actuators. *IEEE Control Syst. Mag.* 15, 40–48 (1995)
10. Chou, C.P., Hannaford, B.: Static and dynamic characteristics of McKibben pneumatic artificial muscles. In: IEEE Robotics and Automation Conf., pp. 281–286. IEEE Press, San Diego (1994)
11. Ferris, D.P., Czerniecki, J.M., Hannaford, B.: An Ankle-Foot Orthosis Powered by Artificial Pneumatic Muscles. *J. Appl. Biomech.* 21, 189–197 (2005)

# A Cooperated-Robot Arm Used for Rehabilitation Treatment with Hybrid Impedance Control Method

Jingguo Wang and Yangmin Li\*

Department of Electromechanical Engineering, University of Macau,  
Macao SAR, China  
YMLi@umac.mo

**Abstract.** Robot-assisted therapy instruments can help patients to recover their upper limbs or lower limbs function. Nowadays, most two-DOF(degree of freedom) robotic arms are applied for upper-limb planar motion in the rehabilitation treatment, while the wrist joint is usually neglected. In this paper, a three-DOF planar robotic arm is used as a rehabilitation treatment tool for the survivor after suffering a stroke or spinal cord injury, aiming to assist them to recover upper-limb function. In order to strengthen the patient's arm, force control in one direction will be desired while the position trajectories in the orthogonal direction should be tracked simultaneously, therefore hybrid impedance control (HIC) is exploited. Redundancy resolution is combined into the control scheme and the constraints of joint limits avoidance(JLA) are considered. The simulations are made to study the HIC strategies and the results are discussed accordingly and the tracking effectiveness of the proposed method is confirmed.

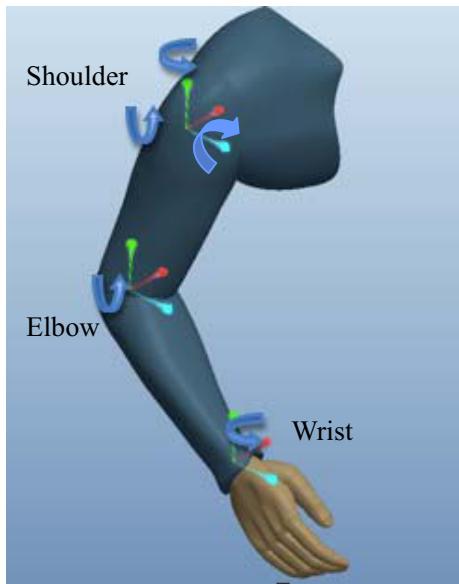
## 1 Introduction

Patients suffering from stroke, traumatic brain and spinal cord injuries always need assistance to rehabilitate the affected limbs [1]-[6]. By traditional methods with motor rehabilitation therapy in daily clinical routine, the frequent and intensive training guided manually by the physiotherapists is time consuming and requires much labor. Robotic training devices are widely used in rehabilitation therapy for the improvement of the upper-limb of patients .

Since human arm is very flexible and the upper limb has several DOFs as shown in Fig. 1, it is very hard to practice the rehabilitation as it really is. Many two-DOF robotic arms [1], like MIT-MANUS [3], are applied for upper-limb planar motion in the rehabilitation treatment, but the wrist joint is usually neglected in some works. The patient's wrist should be considered and trained during the basic movement therapy and it will help to recover the whole upper-limb function [4]. In this paper, the shoulder joint is simplified as a single DOF joint and a 3-DOF planar manipulator is formulated to do the initial-step rehabilitation, which will provide the basic analysis for the future design with spatial-DOF robotic arm. The movements of three patient's joints(shoulder, elbow and wrist) are driven by mechanical active-assisted robotic arm and then the joints data can be recorded and followed up in the process of relearning motor control later.

---

\* Corresponding author.



**Fig. 1.** The structure of human arm

Most of the novel-design rehabilitation robot is configured for safe, stable and compliant operation in close physical contact with humans, which is achieved using impedance control [2]- [12]. Its computer control system modulates the way the robot reacts to mechanical perturbation from a patient or clinician and ensures a gentle compliant behavior [4] [7]. The aim of control is to let the hand track a predefined path in one direction and to impose an assistive or resistant force in the other direction. The hybrid impedance control scheme [2] is a combination of these two fundamental force control strategies, both positions are commanded and a command force trajectory is followed. An augmented hybrid impedance control (AHIC) scheme [12] is proposed for kinematically redundant manipulators combining redundancy resolution to achieve the regulation of the end-effector interaction with the environment and satisfy user defined additional constraints.

The goal of some robot systems designed is to assist in the rehabilitation of patients with neuromuscular disorders by performing various facilitation movements. In [4], a planar robot was designed with impedance control for guiding patients to make movements along the specified trajectories, which had different stimulatory effects on elbow and shoulder joints. A hybrid position/force controller incorporating fuzzy logic was developed to constrain the movement in the desired direction and to maintain a constant force along the moving direction and the planned movements of linear or circular trajectories were considered [1]. A new semi-exoskeleton arm rehabilitation robot having four-DOF called ARMin was developed and tested with numerous stroke and SCI patients, and the patient-cooperative control strategy had been presented and evaluated that ARMin can be used to play ball games or to train ADL-related tasks in combination with an audiovisual display [5].

In this paper, we utilize a three-DOF robotic arm for rehabilitation therapy mainly aiming to assist the affected arms of survivors after the stroke and to train their upper-limb function. The robotic devices are connected with the haptic virtual environment. The virtual constraint is seen as a compliant environment and the desired impedance is assumed to be a mass-spring-damper system. The hybrid impedance control is applied not only to track the desired position trajectory but also to follow the force trajectory simultaneously. The other parts of the paper are organized as follows: In section 2 the manipulator's dynamics is introduced. In section 3, hybrid impedance control method for redundant manipulator is carefully discussed. Simulations are performed and their results are demonstrated in section 4 and the conclusion is given out in the last section.

## 2 Kinematics of the Robotic Arm

### 2.1 Inverse Kinematics

Assume that a task space trajectory is given  $(x(t), \dot{x}(t))$ , and the goal is to find a feasible joint space trajectory  $(q(t), \dot{q}(t))$  that reproduces the given trajectory. The differential kinematics equation establishes a linear mapping between joint space velocities and task space velocities and it has the following form:

$$\dot{x} = J(q)\dot{q} \quad (1)$$

Due to the non-square Jacobian matrix for redundant manipulator, the basic inverse solution to (1) is obtained by using the pseudoinverse  $J^\dagger$  of the matrix  $J$  and the inverse solution can then be written as

$$\dot{q} = J^\dagger(q)\dot{x} \quad (2)$$

where the pseudoinverse  $J^\dagger$  can be computed as  $J^\dagger = J^T(JJ^T)^{-1}$ .

For a kinematically redundant manipulator, a nonempty null space exists due to the excess of input space relative to the manipulable space ( $n > m$ ), which is available to set up systematic procedures for an effective handling of redundant DOFs. Redundancy will make the robot achieve many advantages such as the avoidance of obstacles, joint limits and the singular configurations, or minimization of impacting forces between the end-effector of a manipulator and an environment [13] [11].

### 2.2 Desired Acceleration Trajectories in Task Space

The joint space dynamics can be transformed into the task space and a set of minimal parameters in the null space by defining the weighted inner product in joint space [8], and a decoupled impedance controller can be applied to control the motion of the end-effector as well as the internal motion.

Based on the HIC idea, the spaces are split into the main task  $X$  and additional task  $Z$  into position and force controlled subspaces and their equations of motions have similar forms. The accelerations of main tasks ( $\ddot{X}$ ) and additional tasks ( $\ddot{Z}$ ) have the following forms:

$$\ddot{X} = S_x \ddot{X}_d - M_x^{-1}[B_x(\dot{X} - S_x \dot{X}_d) + K_x S_x(X - X_d) - (I - S_x)F_{xd} + F_{xe}] \quad (3)$$

$$\ddot{Z} = S_z \ddot{Z}_d - M_z^{-1} [B_z(\dot{Z} - S_z \dot{Z}_d) + K_z S_z (Z - Z_d) - (I - S_z) F_{zd} + F_{ze}] \quad (4)$$

where  $S_x$  and  $S_z$  are diagonal elements of the selection matrix  $S$ ,  $S = \text{diag}(S_x, S_z)$  and similarly,  $M_d = \text{diag}(M_x, M_z)$ ,  $K_d = \text{diag}(K_x, K_z)$ ,  $F_d = \text{diag}(F_{xd}, F_{zd})$  and  $F_e = \text{diag}(F_{xe}, F_{ze})$ .

### 2.3 Redundancy Resolution at the Acceleration Level

For compliant motion, dynamic control of redundant manipulators in task space requires the computation of joint accelerations. Hence, redundancy resolution should be performed at the acceleration level. The differential kinematics equations are as follows:

$$\begin{aligned}\dot{X} &= J_e(q)\dot{q} \\ \ddot{X} &= J_e(q)\ddot{q} + J_e(q)\dot{q}\end{aligned}\quad (5)$$

where  $J_e(q)$  is the Jacobian of the end-effector.

Then the joint acceleration can be solved by the pseudo inverse of the Jacobian matrix  $J_e^\dagger$ :

$$\ddot{q} = J_e^\dagger(\ddot{X} - J_e\dot{q}) \quad (6)$$

Similar to the pseudo-inverse solution, the following weighted damped least-squares solution [14] can be obtained:

$$\ddot{q} = [J_e^T W_e J_e + J_c^T W_c J_c + W_v]^{-1} [J_e^T W_e (\ddot{X} - J_e\dot{q}) + J_c^T W_c (\ddot{Z} - J_c\dot{q})] \quad (7)$$

which minimizes the following cost function:

$$L = \ddot{E}_e^T W_e \ddot{E}_e + \ddot{E}_c^T W_c \ddot{E}_c + \dot{q}^T W_v \dot{q} \quad (8)$$

where  $W_e$ ,  $W_c$  and  $W_v$  are diagonal positive-definite weighting matrices that assign priority between the main, additional and singularity robustness tasks. The velocity errors of the main and additional tasks are  $\ddot{E}_e = \ddot{X}_d - (\ddot{X} - J_e\dot{q})$  and  $\ddot{E}_c = \ddot{Z}_d - (\ddot{Z} - J_c\dot{q})$ .

## 3 Hybrid Impedance Control for Redundant Robotic Arm

### 3.1 Dynamics of Robotic Arm in Task Space

The dynamics of the open-chain robotic arm can be obtained using energy-based Lagrange method, which is based on the principle of virtual work function.  $L = T - V$ , where the *Lagrangian*,  $L$ , is defined as the difference between the kinetic energy  $T$  and the potential energy  $V$  of the whole system. For manipulator's compliant motion, control tasks are always described in task space. Then in the absence of friction, the dynamics of the robotic arm can be written as:

$$M(X)\ddot{X} + C(X, \dot{X})\dot{X} + G(X) = F - F_e \quad (9)$$

where  $M(X)$  is the nonsingular symmetric inertia matrix,  $C(X, \dot{X})$  is the vector that implicitly includes centrifugal, coriolis and viscous friction,  $G(X)$  is the gravity terms,  $F$  is the input control force and  $F_e$  is the interaction force between the robotic arm and the environment.

The vector of generalized input force  $F$  is related to the vector of input joint torque  $\tau$  by the Jacobian matrix  $J$ ,

$$\tau = J^T(q)F \quad (10)$$

### 3.2 Hybrid Impedance Control

The impedance control proposed in [8] [9] has the following form:

$$M_d(\ddot{x} - \ddot{x}_d) + B_d(\dot{x} - \dot{x}_d) + K_d(x - x_d) = -F_e \quad (11)$$

where  $M_d$ ,  $B_d$  and  $K_d \in R^{m \times m}$  are positive-definite matrices representing inertia, damping and stiffness respectively, which are usually chosen to be diagonal, rendering the system decoupled, and  $F_e$  is the environmental reaction forces.

Impedance control makes the robot behave as a mass-spring-dashpot system, so the force of interaction due to uncertainty on the location of the contact point and environmental structural properties is regulated [2]. However the impedance control is only a position control scheme, with adjustments made to control the contact forces, and it has not any attempt to follow a command force trajectory. Therefore, the hybrid impedance control proposed is defined as follows:

$$M_d(\ddot{x} - S\ddot{x}_d) + B_d(\dot{x} - S\dot{x}_d) + K_dS(x - x_d) - (I - S)F_d = -F_e \quad (12)$$

Here the task space is divided into orthogonal position and force controlled subspaces using the selection matrix  $S$ .  $S$  is diagonal with ones and zeros on the diagonal while ones on the diagonal relate to position-controlled subspaces and zeros relate to the force-controlled subspaces.  $F_d$  is the desired force. The desired equation of motion in the position-controlled subspaces is identical to (13).

$$M_d\ddot{e}_p + B_d\dot{e}_p + K_d e_p = -F_e \quad (13)$$

where  $e_p$  is the position tracking error and  $e_p = X - X_d$ .

In the force-controlled subspace, the desired impedance is defined by

$$M_d\ddot{x} + B_d\dot{x} - F_d = -F_e. \quad (14)$$

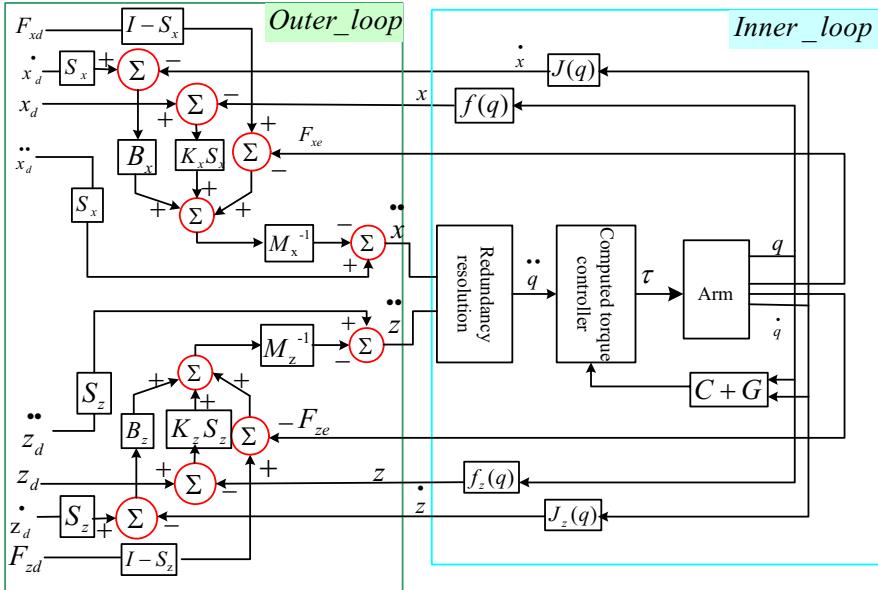
### 3.3 Control Law

In the absence of friction and disturbances, the joint-space dynamic equation of rigid manipulator's motion is described by:

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \tau + J_e^T F_e \quad (15)$$

The proposed control scheme is shown in Fig. 2. The inner and outer loop control strategy can be combined to achieve the closed-loop dynamics [12]. The outer loop outputs the acceleration trajectories reflecting the desired impedance in the position-controlled subspaces as pointed in (3) and (4), and the desired force in the force-controlled subspaces of the main and additional tasks as pointed in (14). The inner-loop control is to select an input to the actuators which makes the end-effector track the desired trajectories generated by the outer loop. The input torque of the manipulator dynamics is defined as:

$$\tau = M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) - J_e^T F_e \quad (16)$$



**Fig. 2.** The hybrid impedance control scheme

### 3.4 Null-Space Optimization

Many techniques of redundancy resolution are used to search the null space, therefore the search direction and the solution are determined by evaluation of chosen performance criteria. The work [16] presented a detailed discussion about performance criteria from two aspects of constraint-based criteria and operational goal-based criteria, while the former includes joint limit avoidance, velocity limit avoidance, peak torque avoidance, obstacle avoidance and mathematical singularity avoidance and the latter considers dexterity, speed of operation, load carrying capacity, manipulator precision, energy minimizing and other criteria.

The most popular method is the gradient projection method (GPM) first introduced in [15] to utilize the redundancy to avoid mechanical joint limits. The particular solution is found using either the pseudoinverse or some other forms of the particular inverse, and the homogeneous solution is then projected onto the null space of the Jacobian matrix. However, there are several problems associated with GPM, including the selection of the appropriate scalar coefficients that determine the magnitude of the self-motion and oscillations in the joint trajectory, and in many cases, inability to avoid the joint limits in time to avoid disruption.

Note that just the joint limit avoidance [11] is considered as local optimization of a performance criterion in this work, and the objective function considered is the distance from mechanical joint limits defined as

$$\Phi(q) = \frac{1}{n} \sum_{i=1}^n \left( \frac{q_i - \bar{q}_i}{q_{iM} - q_{im}} \right)^2 \quad (17)$$

where  $q_{iM}(q_{im})$  denotes the maximum(minimum) limit for  $q_i$  and  $\bar{q}_i$  the middle value of the joint range, and thus redundancy may be exploited to keep the manipulator off joint limits.

## 4 Case Study for Robot-Cooperated Therapy Based on HIC Method

In order to attempt to provide comprehensive solutions for disabled arm recovery, we utilize a simple planar manipulator to hold one patient's arm closely, which assist the patient just as much as needed to perform a particular movement task while some mechanical quantities such as the position, velocity, and forces will be recorded during the motion. A monitor connected with the whole robotic system provides the motion display for the patients and some games are designed to evaluate the patients performance of upper-limb motion. Here the simulated planar motion aims to train the stroke patients recovering their moving ability as the first step of the rehabilitation work, and a simplified model is to show its application purpose in Fig. 3. A constraint object was placed in the virtual reality (or custom games) and it had been simulated as an elastically compliant system, which can be sensed by haptic-feedback devices.

The robotic arm assists the patient to perform some particular movement tasks, which needs whole-joint coordination of shoulder, elbow and wrist. The rehabilitation work requires patients cooperate and interact with the manipulator as a safe, stable and compliant operation, which is achieved using impedance control. Here the hybrid impedance control is considered not only to track the desired position trajectory in one direction (like  $X$ ) but also to track the desired force trajectory in another direction (like  $Y$ ). In order to show the effectiveness of the proposed control method, the related simulation is carried out in the MATLAB/Simulink.

### 4.1 Initial Parameters of Simulation

In the design of rehabilitation robot, it is necessary to consider that the robotic arm is adaptable to the human limb in terms of segment lengths. In this simulation, the length parameters are defined as  $l_1 = 0.25m$ ,  $l_2 = 0.25m$ ,  $l_3 = 0.1m$  and the masses of links are  $m_1 = 3kg$ ,  $m_2 = 2kg$ ,  $m_3 = 1kg$ . Their inertias about the rotating axis have the following values:  $I_{z1} = 0.0625kg.m^2$ ,  $I_{z2} = 0.0417kg.m^2$ ,  $I_{z3} = 0.0033kg.m^2$ .

Another factor to take into account is the range of motion (ROM) of human limb, and then three joints limits of the planar manipulator designed are set as:  $0deg \leq q_1 \leq 150deg$  for shoulder joint,  $-100deg \leq q_2 \leq 0deg$  for elbow joint and  $-80deg \leq q_3 \leq 30deg$  for wrist joint.

The Jacobian matrix of the manipulator is given out as follows:

$$J_e(1, 1) = -l_1 s_1 - l_2 s_{12} - l_3 s_{123};$$

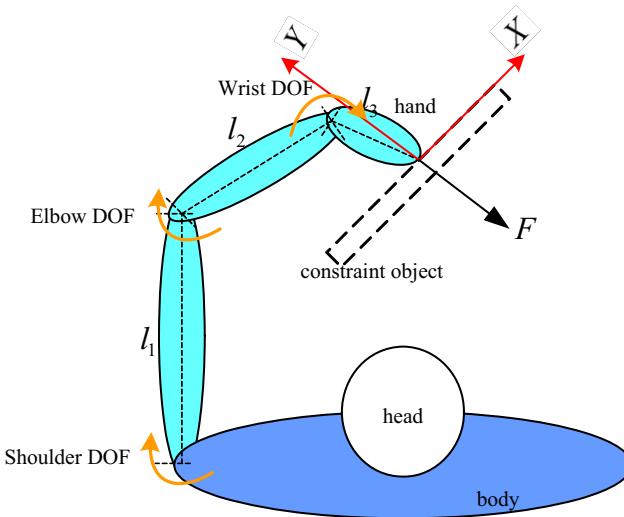
$$J_e(2, 1) = l_1 c_1 + l_2 c_{12} + l_3 c_{123};$$

$$J_e(1, 2) = -l_2 s_{12} - l_3 s_{123};$$

$$J_e(2, 2) = l_2 c_{12} + l_3 c_{123};$$

$$J_e(1, 3) = -l_3 s_{123};$$

$$J_e(2, 3) = l_3 c_{123};$$



**Fig. 3.** 3-DOF planar manipulator imitating human arm

The symmetric positive definite inertia matrix  $M(q)$  can be expressed by:

$$\begin{aligned} M(1, 1) &= 0.4381 + 0.025(10c_2 + c_3 + c_{23}); \\ M(1, 2) &= 0.1413 + 0.0125(10c_2 + 2c_3 + c_{23}); \\ M(1, 3) &= 0.0058 + 0.0125(c_3 + c_{23}); \\ M(2, 2) &= 0.1415 + 0.025c_3; \\ M(2, 3) &= 0.0058 + 0.0125c_3; \\ M(3, 3) &= 0.0058. \end{aligned}$$

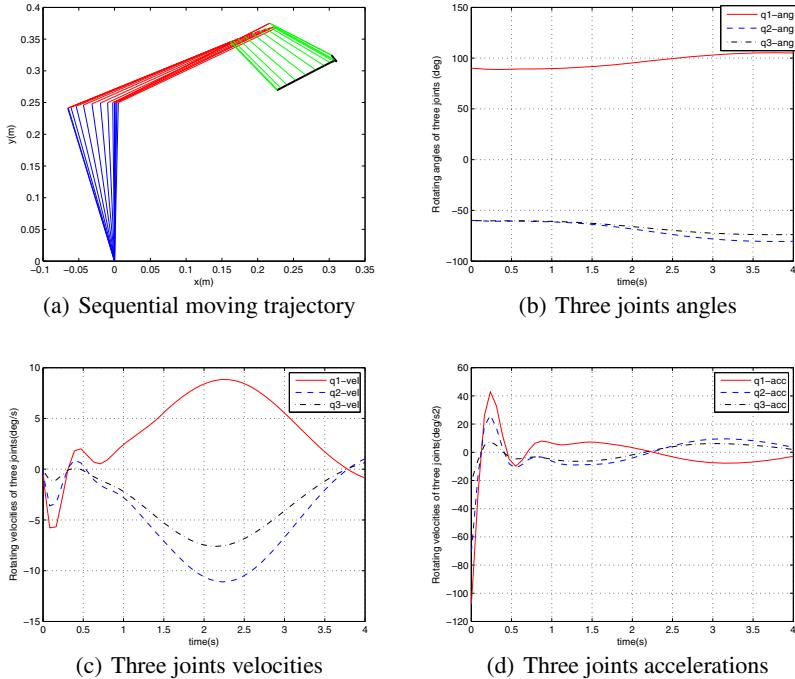
The matrix of Centrifugal and Coriolis forces  $C(q)$  will have the following form:

$$\begin{aligned} C(1, 1) &= -0.0125(10d_2 s_2 + d_3 s_3 + (d_2 + d_3)s_{23}); \\ C(1, 2) &= 0.0125(10d_1 s_2 - d_3 s_3 + d_1 s_{23}); \\ C(1, 3) &= 0.0125d_2 s_3; \\ C(2, 1) &= -0.0125(10(d_1 + d_2)s_2 + d_3 s_3 + (d_1 + d_2 + d_3)s_{23}); \\ C(2, 2) &= -0.0125d_3 s_3; \\ C(2, 3) &= 0.0125(d_1 + d_2)s_3; \\ C(3, 1) &= -0.0125(d_1 + d_2 + d_3)(s_3 + s_{23}); \\ C(3, 2) &= -0.0125(d_1 + d_2 + d_3)s_3; \\ C(3, 3) &= 0. \end{aligned}$$

where  $s_1 = \sin(q_1)$ ;  $c_1 = \cos(q_1)$ ;  $s_{12} = \sin(q_1 + q_2)$ ;  $d_1 = \dot{q}_1$ , and so on.

The weighting diagonal elements in the weighting matrices of the redundancy resolution have the values as  $we = 10$ ;  $wc = 100$ ;  $wv = 20$ .

Other parameters for simulation are selected as:  $M_d = diag(10)$ ;  $B_d = diag(150)$ ;  $K_d = diag(3000)$ . The stiffness of virtual constraint environment has the value  $K_e = 10000 N/m$ .



**Fig. 4.** Simulation results of hybrid impedance control considering the JLA constraint

## 4.2 Simulations Results

From the equations (3) and (4), when the selection matrix is set to  $S_x = \text{diag}(0, 1)$ , it becomes the hybrid impedance control, which is not only to track the position trajectory in  $X$  direction but also to track the force trajectory in  $Y$  direction.

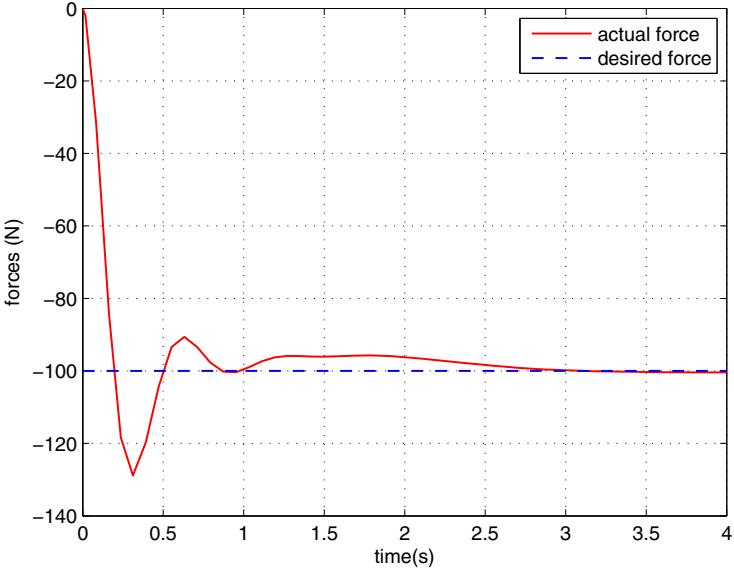
The robot end-effector is moved in the  $X$  direction of constraint frame from one point to another point with constant velocity in 4 seconds and the trajectory is calculated by linear interpolation. The desired force trajectory in the  $Y$  direction is defined as  $F_d = -100N$  as shown in Fig. 2.

An virtual constraint object has been simulated as an elastically compliant system and the interaction force between the robot and the constraint  $F_e$  can be calculated by the following force equation:

$$F_e = K_e \Delta X \quad (18)$$

where  $\Delta X$  denotes the position error between the constraint and the robot end-effector in the constraint frame.

The simulation is performed to track both the position trajectory in  $X$  direction and the force trajectory  $F_d$  in  $Y$  direction and the constraint of JLA is combined into the redundancy resolution.



**Fig. 5.** Desired force and actual force in the simulation

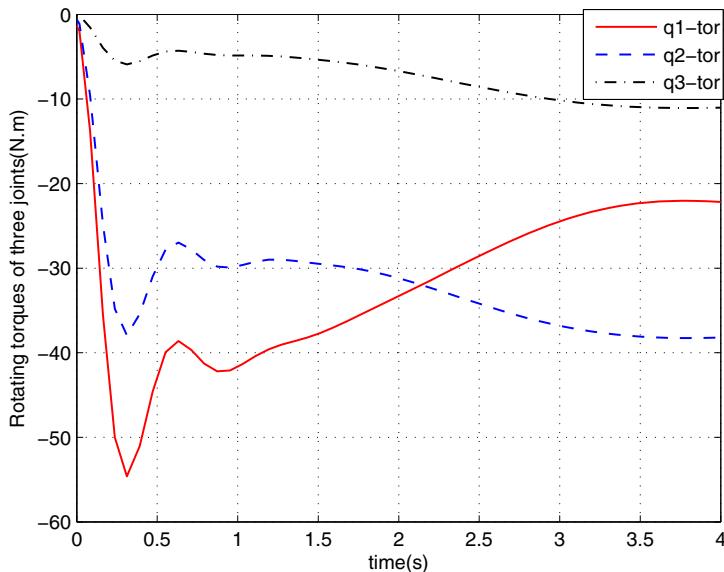
The sequential trajectories of manipulator motion are shown in Fig. 4(a), where the desired position trajectory is tracked well. The rotating angles of all three joints is shown in Fig. 4(b) and it is easily found that all the joints have not exceeded their mechanical limits. The velocities and accelerations of three joints are shown in Fig. 4(c) and Fig. 4(d).

It can be observed that the desired force is followed well from Fig. 5 while the error between the actual and desired force become smaller and smaller with time going on. The rotating torques of all three joints are given out in Fig. 6.

### 4.3 Discussions

A robotic arm with higher DOFs will allow a wide variety of movements, which may be more easier to assist the patients arm in space in order to do various rehabilitation training. However, a robotic arm designed with too many DOFs always comes with a lot of problems, such as heavy weight, large size, big inertia and so on. In this work, we only analyze the three-DOF planar manipulation, which will provide the basic analysis for our further design.

The robotic devices are connected with the haptic virtual environment. The motion designed in the game(or virtual reality) should guide and assist the patient to perform some defined movement tasks, which needs whole-joint coordination of shoulder, elbow



**Fig. 6.** Rotating torques of three joints in the simulation

and wrist. At the same time, some correlated sensory feedback, including visual, audio, and proprioceptive sensory stimuli, should be matched for better rehabilitation.

Similar as the structure of human being, the shoulder joint of the rehabilitation robot should be spherical joint, so one or two DOFs will be added to the shoulder joint for more flexible manipulation.

## 5 Conclusion and Future Work

In order to recover the upper-limb of a patient after the stroke, a three-DOF planar robotic arm is used to do the rehabilitation treatment, which imitates the actual human arm with shoulder, elbow and wrist joints while the wrist joint is always neglected in previous research works. Then hybrid impedance control method is applied to control both the position and the force trajectories in the constraint frame, where the task space is split into position-controlled subspace and force-controlled subspace by modifying the selection matrix  $S$ . The constraints of joint limits avoidance are combined into the redundancy resolution. Corresponding simulations are made with the method of hybrid impedance control, and their results show the proposed control method can track both the position and force trajectory well.

In the next step, a real robotic arm will be fabricated and haptic devices will be installed to provide feedback sense. Some games in virtual reality are displayed to interact with the robotic system and then the patient's movement is guided and assisted.

## Acknowledgment

This work is supported by Macao Science and Technology Development Fund under Grant no. 016/2008/A1 and the Research Committee of University of Macau under grant no. UL016/08-Y2/EME/LYM01/FST.

## References

- Ju, M.S., Lin, C.C.K., Lin, D.H., Hwang, I.S., Chen, S.M.: A rehabilitation robot with force-position hybrid fuzzy controller: hybrid fuzzy control of rehabilitation robot. *IEEE Trans. on Neural Systems and Rehabilitation Engineering* 13(3), 349–358 (2005)
- Liu, G.J., Goldenberg, A.A.: Robust hybrid impedance control of robot manipulators. In: *IEEE Int. Conf. on Robotics and Automation*, California, USA, pp. 287–292 (1991)
- Formica, D., Zollo, L., Guglielmelli, E.: Torque-dependent compliance control in the joint space of an operational robotic machine for motor therapy. In: *IEEE Int. Conf. on Rehabilitation Robotics*, Chicago, IL, USA, pp. 341–344 (2005)
- Krebs, H.I., Hogan, N., Aisen, M.L., Volpe, B.T.: Robot-aided neurorehabilitation. *IEEE Trans. Rehab. Eng.* 6(1), 75–87 (1998)
- Nef, T., Mihelj, M., Riener, R.: Armin: a robot for patient-cooperative arm therapy. *Medical and Biological Engineering and Computing* 45(9), 887–900 (2007)
- Jamwal, P.K., Xie, S., Aw, K.C.: Kinematic design optimization of a parallel ankle rehabilitation robot using modified genetic algorithm. *Robotics and Autonomous Systems* 57(10), 1018–1027 (2009)
- Hogan, N.: The mechanics of multi-joint posture and movement control. *Biological Cybernetics* 81, 39–60 (1999)
- Oh, Y., Chung, W.K., Youm, Y., Suh, I.H.: Motion/force decomposition of redundant manipulators and its application to hybrid impedance control. In: *IEEE Int. Conf. on Robotics and Automation*, Leuven, Belgium, pp. 1441–1446 (1998)
- Wang, J.-G., Li, Y.M.: Massaging Human Feet by a Redundant Manipulator Equipped with Tactile Sensor. In: *IEEE Int. Conf. on Advanced Intelligent Mechatronics*, Montreal, Canada, pp. 7–12 (2010)
- Newman, W.S., Dohring, M.E.: Augmented impedance control: An approach to compliant control of kinematically redundant manipulators. In: *IEEE Int. Conf. Robotics and Automation*, California, USA, pp. 30–35 (1991)
- Wang, J.-G., Li, Y.M.: Impedance control of a spatial redundant manipulator used for relaxing muscle fatigue. In: *IEEE Int. Conf. on Mechatronics and Automation*, Changchun, Jilin, China, pp. 2799–2804 (2009)
- Shadpey, F., Patel, R.V., Balaoutis, C., Tessier, C.: Compliant motion control and redundancy resolution for kinematically redundant manipulators. In: *American Control Conference*, Seattle, WA, pp. 392–396 (1995)
- de Wit, C.C., Siciliano, B., Bastin, G.: *Theory of Robot Control*. Springer, London (1996)
- Seraji, H., Colbaugh, R.: Singularity-robustness and task prioritization in configuration control of redundant robots. In: *29th IEEE Conf. on Decision and Control*, Honolulu, HI, USA, pp. 3089–3095 (1990)
- Liegeois, A.: Automatic supervisory control of the configuration and behavior of multibody mechanisms. *IEEE Trans. Systems, Man, and Cybernetics* 7(12), 868–871 (1977)
- Kapoor, C., Cetin, M., Tesar, D.: Performance based redundancy resolution with multiple criteria. In: *ASME Design Engineering Technical Conference*, Atlanta, Georgia, pp. 1–6 (1998)

# **Characteristics of the Robotic Arm of a 9-DoF Upper Limb Rehabilitation Robot Powered by Pneumatic Muscles**

Xianzhi Jiang, Caihua Xiong, Ronglei Sun, and Youlun Xiong

State Key Lab of Digital Manufacturing Equipment and Technology  
Institute of Rehabilitation and Medical Robotics  
Huazhong University of Science & Technology, Wuhan 430074, China  
Jiang\_sider@163.com

**Abstract.** A wearable exoskeleton upper limb rehabilitation robot developed in our lab consists of 8 active joints and one passive joint. Each active joint is powered by two pneumatic muscles (PMs) in opposing pair configuration through a steel wire with a flexible sleeve. A tension device is used to keep the steel wire in pulley slot. These factors incorporate with characteristics of PM make the rehabilitation robot system so complex (nonlinear, time-varying and time-delay) that it is difficult to model the robotic arm precisely. The characteristics of the shoulder-joint are experimentally studied in this article, which is helpful to design the controllers of the robot.

**Keywords:** Rehabilitation robot, Robotic arm, Pneumatic muscle (PM).

## **1 Introduction**

A patient post-stroke may lose control of upper limb. Fortunately, most of patients may restore motive functions and even the working abilities if they are treated with effective rehabilitation exercises. Generally, rehabilitation exercise is carried out by medical therapists on a person-to-person based physical therapy to the body of patient. But the efficiency of the manual rehabilitation exercise is limited by the therapist's physical ability. Rehabilitation robot can take place of the therapist to a certain extent and it can apply more intensive and efficiency rehabilitation exercises to the patients post-stroke.

Ideally, besides enough physical therapy, rehabilitation robot should provide high levels of safety and flexibility for the patients post stroke. In our lab, a 9-DoFs rehabilitation robot was designed and built with 8 active DoFs and one passive DoF. Each active DoF is a one axes robotic arm driven by two PMs in opposing pair configuration. The main advantage of a PM is that its behavior is very similar to that a human muscle [1], such as its manner of work, range of movement, high power/weight ratio and respond time.

Many previous researches have focused on the characteristics of joint powered by PMs. Sivakumar [2] studied on the dynamic property of PM for design of therapeutic

robot, their researches was based on a rehabilitation joint driven by only one PM (single-driven). John H [3] applied a sliding mode controller to a planar elbow manipulator actuated by PMAs in an opposing pair configuration (double-driven), and obtained good accuracy in elbow angle tracking. The studies of John H [4][5], Kishore [6] and Kyoung [7] are mainly focused on control algorithms and precision.

The main aim of this paper is to present the characteristics of the shoulder as a typical robotic arm of the 9-DoFs rehabilitation robot powered by PMs.

## 2 Experiment Setup

### 2.1 Characteristics of PM

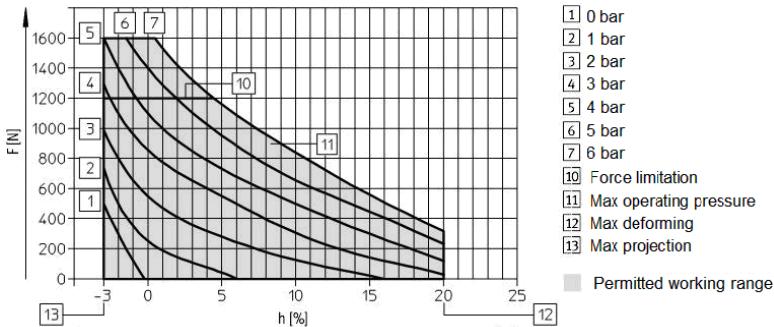
Fig. 1 shows the two types of PMs that were employed in the rehabilitation robot in our lab. The two types of the PM are MXAM-40-AA (Fig. 1. a) and MXAM-20-AA (Fig. 1. b) which are manufactured by Festo (German).



**Fig. 1.** The two types of PMs

PM consists of an expandable rubber tube and two layers of polymer based fibers which are bladed in different directions. It is noted that PM can only pull because when the bladder of the rubber tube is inflated by compressed air, it expands radially against the fibers and causes PM to contract.

Unlike the structure of the PM that studied by A. Manuello Bertetto and M.Ruggiu [8], the two layers of fibers are all built inside of the rubber tube. The advantage of this kind of design is that the friction force between the tube and the fibers is nearly zero. This leads to a good output performance of PM.



**Fig. 2.** Characteristic curves of PM (MXAM-20-AA)[9]

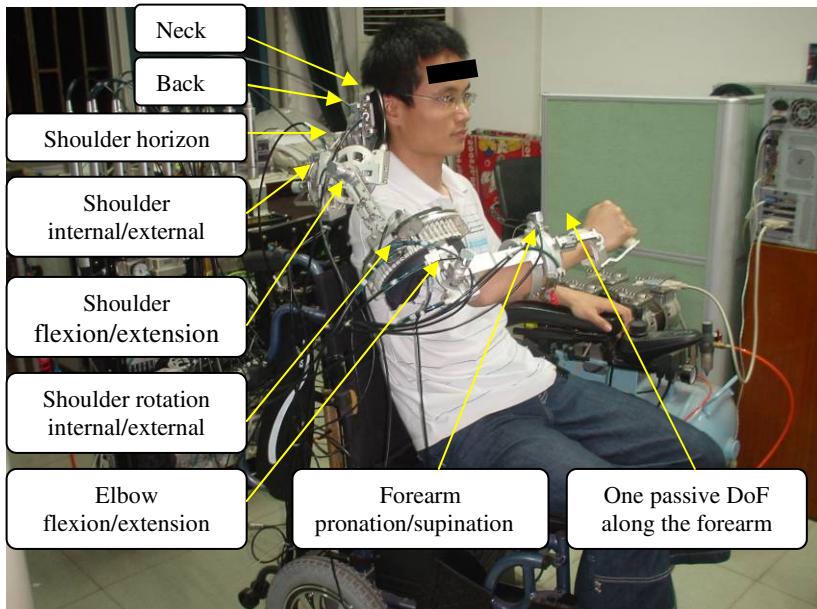
Fig. 2 is the PM's (MXAM-20-AA) characteristic curves. It shows that the output force of PM decreases as the PM contracts. Under a pressure of 0.6MPa (6bar), when

the PM's contraction reaches at a maximum of 20% of its initial length, the output force decreases to about 360N. This property of PM is important for the performance of rehabilitation robot.

Each active robotic arm of the rehabilitation robot is driven by two PMs which are arranged in an opposing pair configuration (double-driven). In the study of Sivakuma [2], the robotic arm was driven by only one PM (single-driven) and in the study of John H [3], the robotic arm was driven by two PMs in an opposing pair configuration. In a single-driven mode, the robotic arm can be raised by the PM and descends under the influence of gravity. But in a double-driven mode, one PM is in charge of uplift of the robotic arm and the other is in charge of pulling it down.

## 2.2 Specifications of the 9-DoFs Rehabilitation Robot

Fig. 3 is the real picture of the 9-DoFs exoskeleton rehabilitation robot designed and built in our lab. The robot was fixed to a wheelchair and the patient's torso was fixed to the wheelchair with straps and bands.

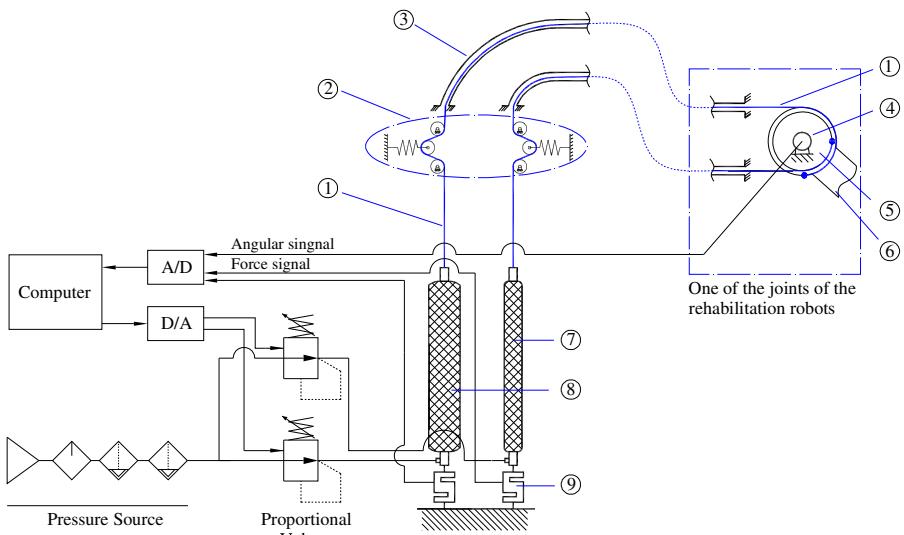


**Fig. 3.** The 9-DoFs rehabilitation robot

The rehabilitation robot had 8 active DoFs and one passive DoF (Fig. 3). Each active DoF is an one axes robotic arm which is driven by two PMs in opposing pair configuration (Fig. 4). Each PM can be operated individually by a proportional valve that controlled by computer.

Fig. 4 shows that each PM actuates the robotic by a steel wire through a tension device and a flexible sleeve of the steel wire. There are several advantages of this kind of mechanism: First, the arm-of-forces of the forces do not change while the

joint rotates; Second, the steel wire pulls the robotic arm to rotate when the PM contracts and there is no action of gliding between the steel wire and the pulley, so the angular information of the pulley would not lose. Third, the tension device could keep the steel in the slot of the pulley.



1-steel wire, 2-tension devices, 3- flexible sleeve, 4-angular sensor, 5-pully, 6-arm and load, 7-Slave-Side PM(MXAM-20-AA), 8-Master-Side PM(MXAM-40-AA), 9-force sensor.

**Fig. 4.** Schematic diagram of robotic arm

The initial length of master-side PM and the slave-side one is 450mm and their initial radius is 40mm and 20mm. The gas pressure of the pressure source is 0.7MPa (7bar). The radius of the pulley is 50mm.

The hardware included an IBM-compatible personal computer (AMD dual core 5000+, 2.60GHz), the proportional valve is ITV2031-212L-X26 (SMC, Japan) A/D card is USB2810A (Art, Beijing) and D/A card is PCI6216V (ADLINK, Taiwan). The control software was coded in Delphi 7 (Pascal program language).

### 3 Characteristics of the Robotic Arm

Generally, it is difficult to develop an accurate mathematical model of PM via academic analysis. This is due to the PM's nonlinearities and the compressibility of air [10]. This means that modeling of the robotic arm of the rehabilitation robot which is driven by two PMs will be more complex than modeling for only one PM. In addition, the employments of tension devices and flexible sleeves incorporate with PMs lead to more complex of the rehabilitation robot system.

### 3.1 Nonlinear Factors

There are three nonlinear factors in the robotic arm system: PM, tension device and flexible sleeves of the steel wire.

#### 3.1.1 Coupled Nonlinearities of PMs

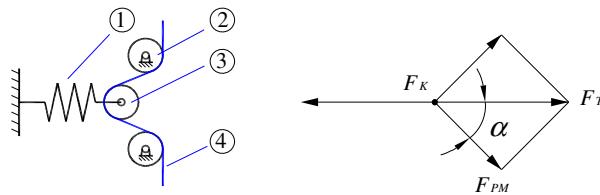
PM's nonlinearities is shown in Fig. 2. The model of PM can be described as following [11]:

$$\begin{aligned} F(\varepsilon, p) &= \pi \cdot r_0^2 \cdot p [a(1-\varepsilon)^2 - b] \\ \varepsilon &= (l_0 - l)/l_0, a = 3/\tan^2 \alpha, b = 1/\sin^2 \alpha \end{aligned} \quad (1)$$

where  $F$  is output force,  $\varepsilon$  is contraction ratio,  $p$  is the working pressure,  $l_0$  is PM's relaxed length,  $l$  is real length,  $r_0$  tube relaxed radius and  $\alpha$  is the fiber bias angle. The model is only used on theoretical analysis of performance of PM but is not precise enough in robotic arm of the rehabilitation robot. Furthermore, the coupling relationship between the output torques and angular displacements of the master-side PM and the slave-side one make the mathematical models of the robotic arm more complex.

#### 3.1.2 Time Delay of Tension Device

The function of tension device is to prevent the steel wire form jumping out of the slot of the pulley. But tension device brings time delay to the robotic arm.



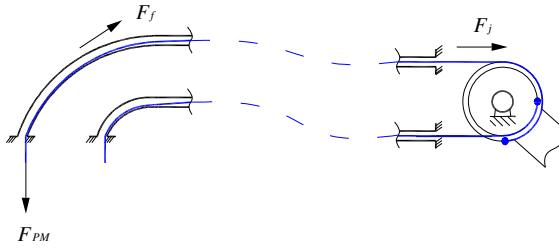
1-spring, 2-fixed pulley, 3-live pulley, 4-steel wire.

**Fig. 5.** Schematic diagram of tension device

Fig. 5 shows the Schematic diagram of tension device. In Fig. 5,  $F_K$  is force of the spring,  $\alpha$  is angle of the spring to horizontal direction.  $F_{PM}$  is the output force of PM. When the PM begins to contract, the live pulley would be pulled to the other side before robotic arm rotates. The time delay of tension device is affected by pressure of PM, spring stiffness, and the mechanical structure of tension device. We found out that a typical value of the time delay was 0.1~0.3 second.

#### 3.1.3 Friction Force between Steel Wire and Flexible Sleeve

As is shown in Fig. 4, the control box of rehabilitation robot connects to each robotic arm via two steel wires and each steel wire wears a flexible sleeve. The nonlinear and unpredictable friction force between steel wire and its flexible sleeve make it very difficult to obtaining good performance of the robotic arm.



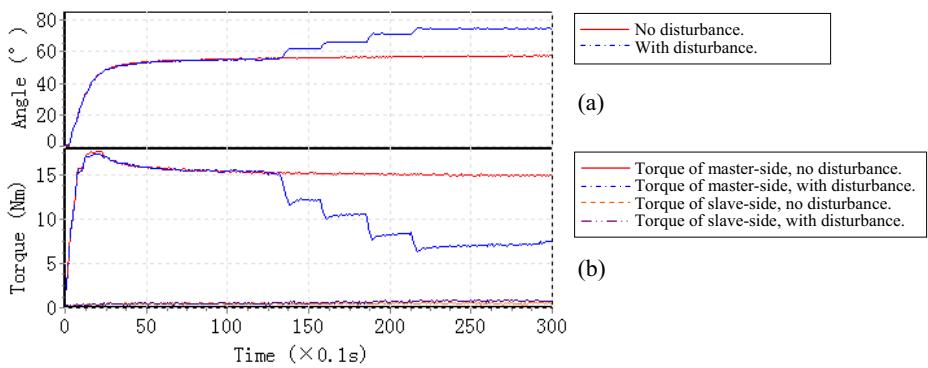
**Fig. 6.** Schematic diagram of friction force

In Fig. 6,  $F_{PM}$  is the force master-side PM,  $F_j$  is force of joint exerts on steel wire and  $F_f$  is friction force between steel wire and its flexible sleeve.  $F_f$  is affected mainly by friction factor of steel wire to its sleeve,  $F_{PM}$ , radius and length of sleeve. The friction force influences so heavily on the robotic arm that it is very difficult to control the movement and force of robotic arm especially when the direction of the friction force is changing.

### 3.2 Performance of Pressure Open-Loop

#### 3.2.1 Performance of Pressure Open-Loop under Single-Driven

For researching on performance of pressure open-loop of the shoulder as a typical robotic arm of the rehabilitation robot, we set the pressure of master-side PM at 0.15MPa and slave-side one at 0Mpa, we found that after a time delay of about 0.2 second, the shoulder rotated quickly until it reached at about  $50^\circ$  (Fig. 7) in 3 seconds and the angular displacement increased slowly in at least 15 minutes. And the torque of master-side PM reached at a peak value of 17.75Nm in 2 seconds and then decreased slowly all along.



**Fig. 7.** Angular and torque curves of robotic arm with the pressure of master-side PM was set to 0.15MPa and that of slave-side PM was set to 0MPa

In another experiment of the same condition, we gave disturbances (applied four torques on the robotic arm and each action sustained for about 0.5 second) to the robotic arm when the robotic arm was running, we found that the angular

displacement of robotic arm moved to a new position and when we removed disturbances, and the robotic arm kept stable at the new position. Simultaneously, the torque of master-side PM decreased quickly for a short time and increased slowly all along. The torque of slave-side PM did not change much in the two experiments.

Fig. 7 plotted the results of the two experiments. From Fig. 7, the robotic arm's angular displacement and PM's output torque are not completely depending on the pressure of PM, because the friction forces between the steel wires and their sleeves played important roles in angular displacement and torque of the robotic arm.

### 3.2.2 Performance of Pressure Open-Loop under Double-Driven

Ideally, when the pressure of master-side PM and that of slave-side one are set individually to specified values, the robotic arm would rotate to a unique angular position, either of torques of the two PMs. But it was not the case in reality at all which was verified by the following experiments. The offline analyses of data were used to analyzing the performance of pressure open-loop of robotic arm.

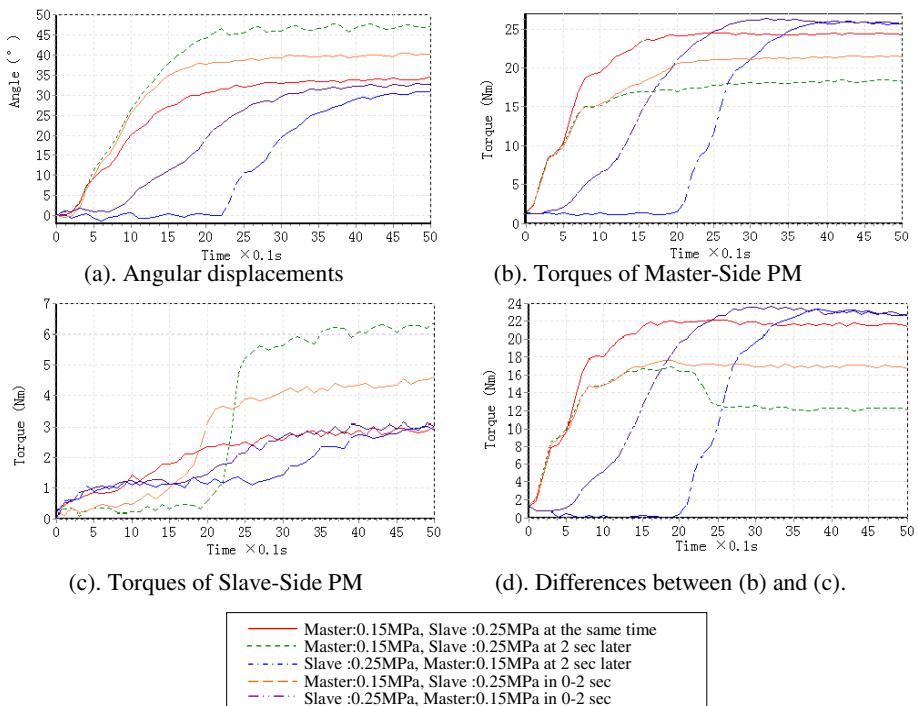
Case 1: The pressure of master-side PM was set to 0.15MPa and the slave-side one was set to 0.25MPa at the same time;

Case 2: Master-side of 0.15MPa, slave-side of 0.25MPa at 2 seconds later;

Case 3: Slave-side of 0.25MPa, master-side of 0.15MPa at 2 seconds later;

Case 4: Master-side of 0.15MPa, slave-side of 0~0.25MPa in 0~2 seconds;

Case 5: Slave-side of 0.25MPa, master-side of 0~0.15MPa in 0~2 seconds;



**Fig. 8.** Results of performances of robotic arm with the pressure of master-side PM was set to 0.15MPa and that of slave-side PM was set to 0.25MPa at different time

Fig. 8. (a) plotted angular displacements of robotic arm in the five cases. Fig. 8. (b) plotted torques of master-side. Fig. 8. (c) plotted torques of slave-side one. Fig. 8. (d) plotted the difference of the torque of master-side PM and that of slave-side one.

As shown in Fig. 8, the robotic arm is a nonlinear, time-varying and time-delay system with unpredictable factors such as friction forces of steel wires to the flexible sleeves. Therefore, modeling and control of the robotic arm would be challenges.

## 4 Position and Force Control of the Robotic Arm

The purpose of this section is not to control the robotic arm with high precision, but to observe the dynamic performance of the robotic arm. Since the two PM (master-side and slave-side) of robotic arm can be controlled individually, there are three modes of control of the robotic arm as shown in table 1:

**Table 1.** Three control modes of the two PMs

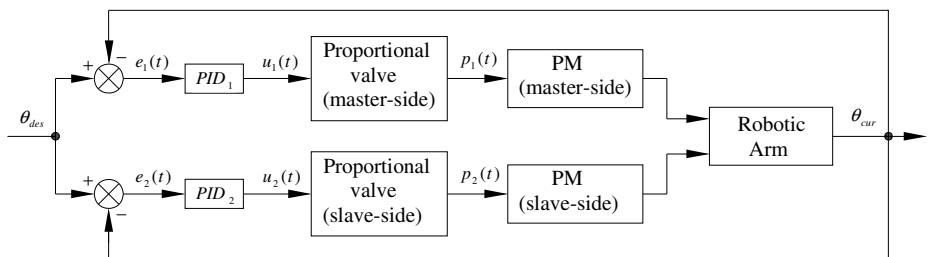
Mode	Master-Side PM	Slave-Side PM	Rehabilitation Exercise
1(P-P)	Position control	Position control	robot-active-patient-passive without stiffness control.
2(P-F)	Position control	Force control	robot-active-patient-passive with stiffness control.
3(F-F)	Force control	Force control	robot-passive-patient-active training.

Both the position and force controls were implemented in discrete time and the sampling rate was 10 Hz. The controllers of all the following experiments were conventional PID controllers with fixed parameters of  $K_p$  (proportional gain),  $K_i$  (integral gain) and  $K_d$  (derivative gain).

### 4.1 Performance of Position-Position Control

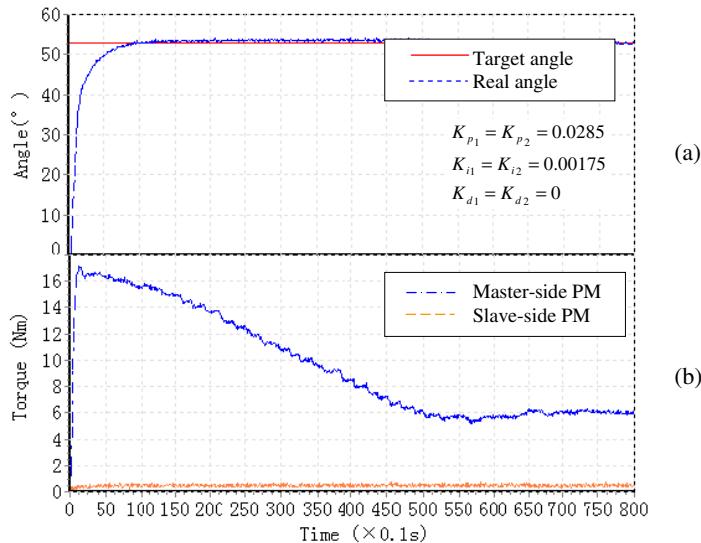
Position-Position (P-P) control mode is that both master-side PM and slave-side one are controlled based on angular position feedback of the angular sensor fixed in the robotic arm. This control mode is applied to robot-active-patient-passive rehabilitation exercises without stiffness control. The structure of P-P control mode is shown in Fig. 9.

Both PID controllers used the same parameters with  $K_p = 0.0285$ ,  $K_i = 0.00175$  and  $K_d = 0$  which are showed in Fig. 10.



**Fig. 9.** Structure of Position-Position control mode

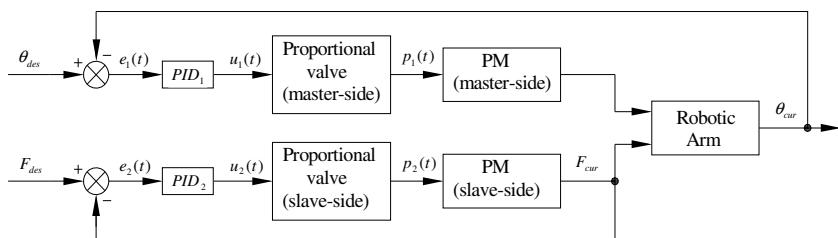
Fig. 10 shows the experimental results of P-P control mode of the robotic arm. And the destination angle was  $270^\circ$ . From Fig. 8. (a), the robotic arm reached at  $270^\circ$  after about 10 seconds. There was no vibration but an overshoot of about  $1^\circ$  for the following 40 seconds, and finally, the robotic arm kept stable at  $270^\circ$ . The torque of master-side PM reached at a peak torque of  $17.1\text{Nm}$  after 1.3 second from the start and then decreased for about 1 minute, and after that, the torque of master-side PM vibrated form  $5\text{Nm}$  to  $13.5\text{Nm}$  at a cycle of about 3 minutes (the data was not plotted in Fig. 10).



**Fig. 10.** Results of Position-Position control mode

## 4.2 Performance of Position-Force Control

In Position-Force (P-F) control mode, the master-side PM is controlled based on angular position feedback of the robotic arm and the slave-side one is controlled based on its force feedback (Fig. 11). This mode is applied to robot-active-patient-passive rehabilitation exercise with stiffness control.

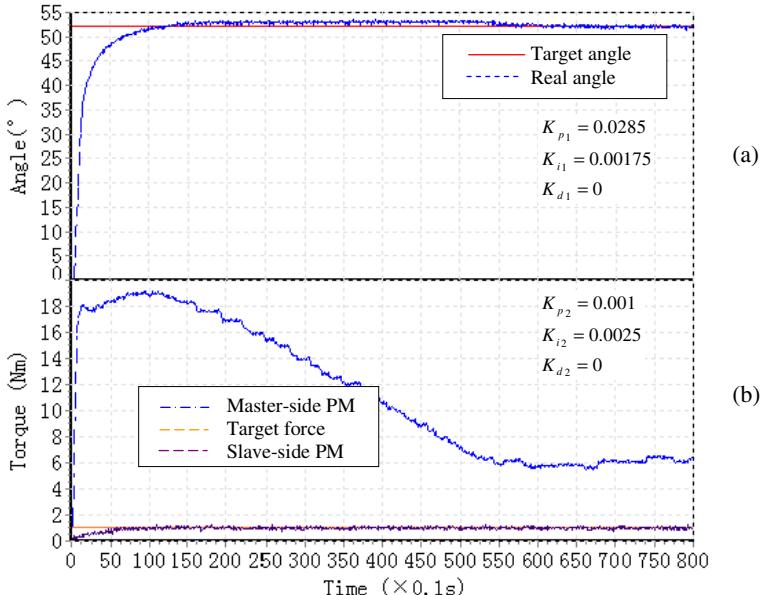


**Fig. 11.** Structure of Position-Force control mode

The parameters of the two PID controllers are showed in Fig. 12.

Fig. 12 shows the experimental results of P-F control mode of the robotic arm with a destination angle of  $270^\circ$  and the destination torque of slave-side PM was 1Nm.

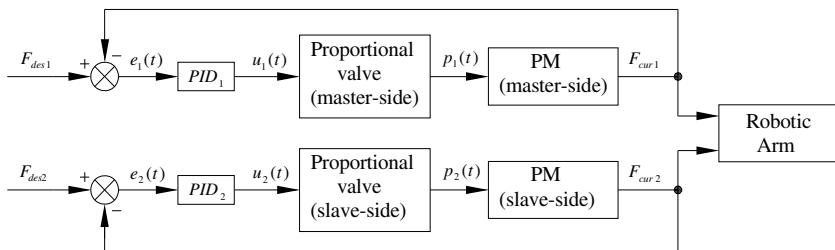
From Fig. 12. (a), the performance was similarity to that of Fig. 10. (a). But the torque of master-side PM reached at a peak torque of about 19Nm after 10 seconds from the start and then decreased for about 1 minute, and after that, the torque of master-side PM vibrated form 100N to 290N at a cycle of about 3 minutes (the data was not plotted in Fig. 12). The torque of slave-side PM reached at 1Nm (the destination force is 20N) after 5 seconds form the start and then kept stable.



**Fig. 12.** Results of Position-Force control mode

### 4.3 Performance of Force-Force Control

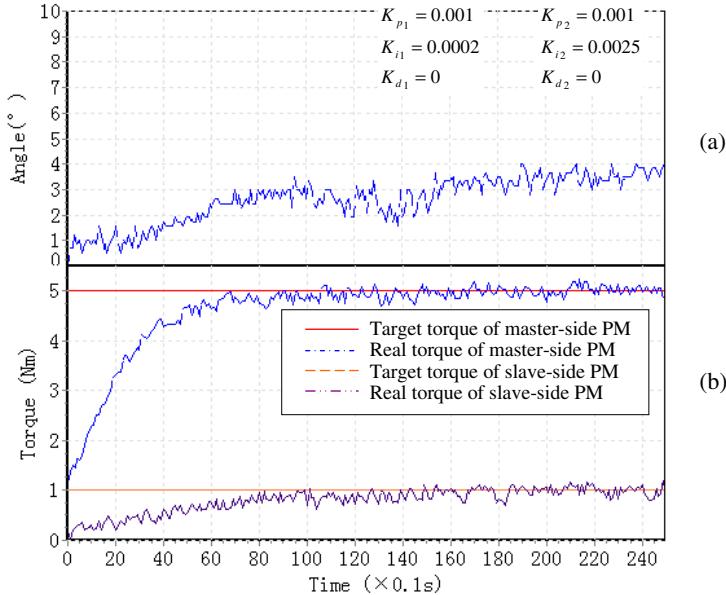
Force-Force (F-F) control mode is that both master-side PM and slave-side one are controlled based on its force feedback. This mode is applied to robot-passive-patient-active rehabilitation exercise. The structure of F-F control mode is shown in Fig. 13.



**Fig. 13.** Structure of Force-Force control mode

The parameters of the master-side PID controller were set to be  $K_{p1} = 0.001$ ,  $K_{i1} = 0.0002$  and  $K_{d1} = 0$  and that of the slave-side were set to be  $K_{p2} = 0.001$ ,  $K_{i2} = 0.0025$  and  $K_{d2} = 0$ .

Fig. 14 shows the experimental results of F-F control mode of the robotic arm. The destination force of master-side PM was 100N (the destination torque is 5Nm) and that of slave-side one was 20N (1Nm). From Fig. 14. (a), the torque of master-side PM reached at the target torque of 5Nm after about 6.8 second and the torque of slave-side one reached at the target torque of 1Nm in about 10 seconds.



**Fig. 14.** Results of Force-Force control mode

## 5 Conclusion and Future Works

The robotic arm of the 9-DoFs upper limb wearable exoskeleton rehabilitation robot is a very nonlinear, time-varying and time-delay system because the employments of PMs, tension devices and steel wire flexible sleeves.

When the pressure of both master-side PM and slave-side one are set individually to specified values, the robotic arm would rotate to a various angular positions under different operate time, either of forces of the two PMs. Especially when there is a disturbance, the angular position and forces of two PMs would shift to new values, which is due to friction force between steel wire and its flexible steel sleeve.

Conventional PID controllers show good performance in position and force control of the robotic arm. But due to the characteristics of the robotic arm system, there is always an overshoot in position control and the response time is not quick enough.

One of the interesting scopes for future work is in modeling of the robotic arms powered by two PMs with coupling, two tension devices and two steel wires with flexible sleeves. The other work is to design nonlinear/robust/adaptive controllers for achieving superior performance in the presence of model uncertainties and external disturbances with minimum reaching time, minimum distance error, and smooth control actions of the robotic arm for the rehabilitation robot.

## Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant No. 50775080), ‘863’ Hi-Tech Research and Development Program of China (Grant No. 2008AA040202 and 2007AA041501), and National Key Technology R&D Program of China (Grant No. 2008BAI50B04).

## References

1. Repperger, D.W., Phillips, C.A., Neidhard-Doll, A., Reynolds, D.B., Berlin, J.: Actuator design using biomimicry methods and a pneumatic muscle system. *Control Engineering Practice* 14, 999–1009 (2006)
2. Balasubramanian, S., Huang, H., He, J.: Quantification of Dynamic Property of Pneumatic Muscle Actuator for Design of Therapeutic Robot Control. In: 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2006 (2006)
3. Lilly, J.H.: Sliding Mode Tracking for Pneumatic Muscle Actuators in Opposing Pair Configuration. *IEEE Transactions on Control Systems Technology* 13(4) (2005)
4. Lilly, J.H.: Adaptive tracking for pneumatic muscle actuators in bicep and tricep configurations. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11(3) (2003)
5. Lilly, J.H.: A two-input sliding-mode controller for a planar arm actuated by four pneumatic muscle groups. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 12(2) (2004)
6. Balasubramanian, K., Rattan, K.S.: Trajectory tracking control of a pneumatic muscle system using fuzzy logic. In: Annual Meeting of the North American in Fuzzy Information Processing Society, NAFIPS 2005 (2005)
7. Ahn, K.K., Thanh, T.D.C.: Performance improvement of artificial pneumatic muscle manipulator. In: Proceedings of the 8th Russian-Korean International Symposium on Science and Technology, KORUS 2004, vol. 1 (2004)
8. Manuello Bertetto, A., Ruggiu, M.: Characterization and modeling of air muscles. *Mechanics Research Communications* 31(2), 185–194 (2004)
9. Festo catalog, Fluidic muscle DMSP/MAS, <http://www.festo.com>
10. Zhang, J.-F., Yang, C.-J., Chen, Y., Zhang, Y., Dong, Y.-M.: Modeling and control of a curved pneumatic muscle actuator for wearable elbow exoskeleton. *Mechatronics* 18(8), 448–457 (2008)
11. Chou, C.-P., Hannaford, B.: Measurement and Modeling of McKibben Pneumatic Artificial Muscles. *IEEE Transactions on Robotics and Automation* 12(1), 90–102 (1996)

# Optics Based Motion Measurement for a Catheter Navigation System: A Novel and Low Cost Approach

Yonghua Yan, Daguo Chen, and Hang Yin

Robot Research Institutue, Shanghai Jiaotong University,  
Dongchuan Rd.800, Shanghai, China  
[yhyan@sjtu.edu.cn](mailto:yhyan@sjtu.edu.cn)

**Abstract.** Robot-assisted therapy has been investigated to protect physicians from the radiation exposure during fluoroscopic x-ray image guided catheter intervention. This paper introduces an optics based approach to measure catheter motion input to build a catheter navigation system. The principle of measurement, which forms foundation of catheter sensor system, is presented in terms of geometry relationship and approximation. Then calibrated motion measurement is achieved, for both translational and rotational components. A common optical mouse device is used to cheaply implement such a sensor system prototype with other necessary mechanical stages. A computer running corresponding software acquires and processes motion sense via a USB port. Partial experimental results on motion measurement are analyzed, to show proposed approach has impressive motion sense resolution. Finally, comparison between optical mouse sensor and traditional optical encoder are demonstrated and discussions on this technology and catheter manipulator system are also made for future research and work.

**Keywords:** Catheter intervention, robot-assisted therapy, remote navigation, motion measurement.

## 1 Introduction

Fluoroscopic x-ray guided catheter intervention has become an important method for medical practice such as diagnosis and treatment of hepatic tumor [1][2]. During the procedures of intervention, Fluoroscopic x-ray images are employed to assist physicians with exact vascularity, while DSA (Digital Subtraction Angiography) is used to enhance the image when the physicians encounter confusing areas. Unfortunately, traditional catheter intervention which is often conducted manually by physicians may result in potential diseases to them due to long time radiation exposure. For the sake of safety, physicians have to wear protective suits which, however, are believed to raise chronic back and neck pain [3]. An alternative solution to reduce radiation exposure is performing catheter intervention at a remote location. To achieve the goal, robot-assisted techniques are believed to play an active role to build such kind of remote catheter navigation system.

Robot-assisted surgery had an impressive development during the last decades. Since the 1980s, the Probot, Robodoc explored the field, after the da Vinci system launched by Intuitive in 1999, the da Vinci, GTS, obtained FDA clearance in succession [4]. Those systems are mainly designed as assistance of human hands. Generally, two sub-systems constitute a whole catheter navigation system. Catheter sensor system is used to measure the motion from operators and catheter manipulator system replicates corresponding movement to drive actual catheter intervention. Based on this architecture, successful products and system have been developed. Negoro et al. presented their remote catheter which allows physicians using a joystick to control the motion of catheter [5]. While Corindus Inc. employed both a joystick and a touch screen to sense continuous and discrete motion respectively [6]. Another system, proposed by Niobe, Stereotaxis, Inc., navigates a relatively small magnetic guide wire by generating the path of another permanent magnet [7]. Some of the systems utilize standard input devices like joystick to detect the interventional motion, which however, also have physicians conduct this procedure in an unfamiliar way instead of an intuitive one. Some study has implied that considerable training time is needed for experienced physicians to master Niobe system [8]. To address this problem, Yogesh et al. proposed their powerful catheter sensor system, with which physicians are able to pull, push or rotate the input catheter directly rather than withdraw their dexterous skills [9]. For manipulator sub-system, force sensors are often used to provide better performance and other ways for actuation like SMA (Shape Memory Alloy) have also been investigated [10][11][12].

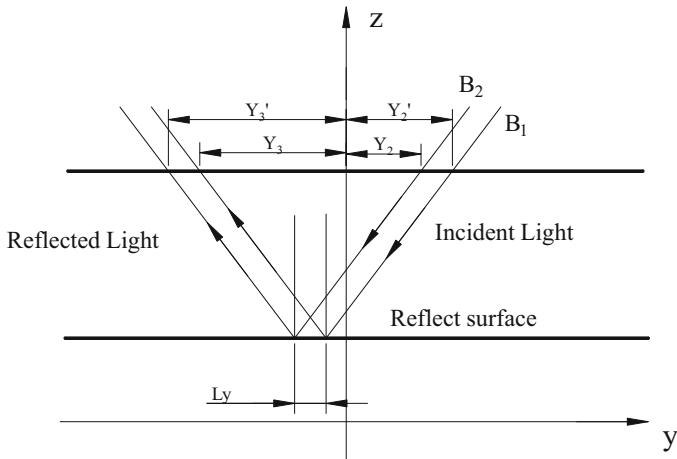
In this paper, we introduce a novel design of catheter sensor system based on the sensor used by optical mouse. The motion measuring method we proposed can be utilized to build a catheter sensor system enabling physicians to work in their conventional way. Relevant optics principle and some key issues for implementation are discussed and addressed. A series of experiments are introduced to evaluate the precision of a prototype sensor system and the results demonstrate pretty good performance. What's more, it's easy to accept that the approach, which requires lower expense and less complexity compared with measurement based on special electromechanical design, paves the way to a low cost catheter navigation system.

## 2 Methods

In catheter intervention procedure, sensor system is used to measure the motion of input catheter, which mainly includes translational and rotational components. The system must also be compatible with the dimensions of general catheters whose diameter is near 2 mm.

Optics based methods are frequently used for motion sensation for the high accuracy as a kind of non contact measurement. Optical mouse is a common but acute and effective device for motion measuring on a 2-dimensional surface. These days most of the laser based optical mouse can achieve the resolution of 1000 dots per inch(DPI), for example the Cypress CYONS1001U has programmable resolution of 400, 800, 1600 and 3200 DPI. According to experience of interventional physicians, a resolution of 0.5mm for translational movement and 1 degree for rotational movement

would be sufficient because the catheter is relative flexible and human hands' manipulation are not so accurate. So we can simply apply relevant techniques to implement a catheter sensor system. However, additional analysis and enhancement are still necessary before the implementation.



**Fig. 1.** Principle of relative movement sensation of optical mouse

As is shown in Fig. 1, Optical mouse uses LED or other light sources to illuminate the surface it tracks over. A tiny video camera is settled to receive reflected light to take a series of images of the surface. The images are processed by a special microchip embedded in the mouse. Thus relative motion between surface and mouse can be detected by comparing collected images. In this way, great accuracy can be achieved with the help of high resolution of the device, which is usually 1000 dpi or even more, for modern optical mice.

However, since the basic geometry of common catheter is a surface of cylinder instead of a plane for tracking on, fundamental analysis has to be done before applying this method directly to build the catheter sensor system.

## 2.1 Translational Motion Measurement

When physicians pull or push input guide wire, motion along translational direction is measured. Since the track is a straight line, translational measurement can be performed as on a surface. It is inevitable that the installation of sensor is not as expected so calibration is needed to address this problem. Assuming catheter sensor is placed with an oblique angle  $\alpha$  as illustrated in Fig. 2. Take displacement from the origin point O to the point A as example, we denote  $L_{xw}$ ,  $L_{yw}$  and  $L_{xs}$ ,  $L_{ys}$  as actual and measured displacement in different directions respectively. Since the two coordinates only have a relationship of rotation transformation, we have relations as follows:

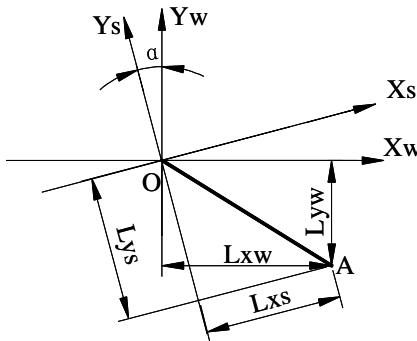
$$\begin{bmatrix} L_{xs} \\ L_{ys} \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} L_{xw} \\ L_{yw} \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} L_{xw} \\ L_{yw} \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}^{-1} \begin{bmatrix} L_{xs} \\ L_{ys} \end{bmatrix} \quad (2)$$

With the equations above, calibration can be done to compensate possible error from imperfect placement. Let  $L_{xw}$  or  $L_{yw}$  be zero and substituted into equation.2, we may evaluate the oblique angle  $\alpha$  as follows:

$$\alpha = \arctan\left(\frac{L_{xs}}{L_{ys}}\right) \quad (3)$$

Then the movement along translational direction can be easily solved from sensed displacement as oblique angle  $\alpha$  is known.



**Fig. 2.** Calibration for offsetting oblique placement of optical sensor

## 2.2 Rotational Motion Measurement

In the case of rotational movement, the sensing procedure can no longer be simply treated as tracking on a planar surface. Interestingly, however, we may make such an approximation as long as high sampling rate is guaranteed. Fortunately, common optical mouse device indeed meet this requirement. Considering a time period short enough, it is reasonable to approximate the relation between actual motion track, an arc, and sensed distance, a straight line, with a relatively simple function.

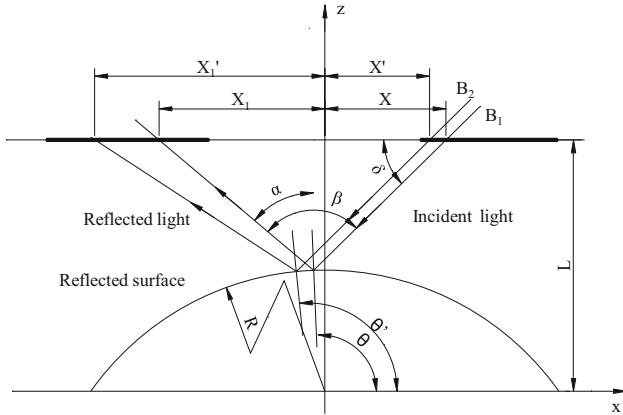
First, we give notations to some useful parameters as shown in Fig. 3.  $B_1$  and  $B_2$  are incident light.  $X$  and  $X_1$  are used to indicate horizontal distance of incident and reflected light point respectively. Related incidental and reflected angle are denoted by  $\alpha$ ,  $\beta$  and  $\delta$ . Corresponding angle of rotation is denoted by  $\theta$ . Given radius of the catheter  $R$ , we may get the length of arc by calculating

$$L_1 = (\theta - \theta') \cdot R = \Delta\theta \cdot R \quad (4)$$

While actual sensed distance is

$$\Delta X = X - X_1 \quad (5)$$

Obviously, if the surface being tracked on is planar, sensed displacement will equal to actual move distance  $L_1$ . Therefore, what we should evaluate is the function relation between  $\Delta X_1$  and  $L_1$ .



**Fig. 3.** Approximation of rotational motion measurement

According to Fig. 3, we may have such statements

$$X_1 = -R\cos\theta + (X - R\cos\theta)\cot(\delta - 2\theta)\tan\delta \quad (6)$$

$$X\tan\delta + R\sin\theta - R\cos\theta\tan\delta = L \quad (7)$$

Apply  $\alpha = 2\theta - \frac{\pi}{2} - \delta$  and  $\beta = 2(\theta - \delta)$  to statements above, substitute (6), (7) into

(5) and transform the equation. We will have

$$\begin{aligned} \Delta X_1' &= X_1' - X_1 \\ &= -R\cos\theta + (L - R\sin\theta)\cot(\delta - 2\theta) + R\cos\theta - (L - R\sin\theta)\cot(\delta - 2\theta) \\ &= -R(\cos\theta - \cos\theta) + \frac{L[\tan(\delta - 2\theta) - \tan(\delta - 2\theta)] - R[\sin\theta \tan(\delta - 2\theta) - \sin\theta \tan(\delta - 2\theta)]}{\tan(\delta - 2\theta) \tan(\delta - 2\theta)} \end{aligned} \quad (8)$$

Similarly, we may have  $\Delta\theta = \theta - \theta'$ . Assume the physician manipulate the catheter at a speed of 2 round per second which is fairly fast for the intervention, the optical mouse has a sampling rate of 1000Hz, so we have an increment of  $\pi/250$  rad, which

is only 1% of  $\theta = \pi/2$  rad. Thus we may take  $\Delta\theta$  as a variable minor enough and apply approximated relations as follows

$$\tan\Delta\theta = \sin\Delta\theta = \Delta\theta \quad (9)$$

Substitute (9) into Statement (8), then let  $\theta$  and  $\delta$  be  $\pi/2$  and  $\pi/4$  respectively, which means an incident light shines on the top point of arc with an incident angle of  $\pi/4$ . We finally get approximated  $\Delta X_1'$  as the expression below

$$\Delta X_1' = R\Delta\theta \left[ 1 + \frac{4L}{R(1+2\Delta\theta)} - \frac{4}{1+2\Delta\theta} \right] = R\Delta\theta \left[ 1 + \frac{4(L-R)}{1+2\Delta\theta} \right] \quad (10)$$

Having the relation of  $2\Delta\theta \ll 1$ , we may conclude that the distance sensed on cylindrical surface is approximately proportional to the result on planar surface. The factor can be solved from an experimental calibration as what we discussed in the section of translational motion measurement.

One thing should be noted is that, in practice,  $R$  is unlikely to be a constant as perfect cylindrical surface is not always guaranteed. Variant radius  $R$  must be taken into account and corresponding effect on measurement should be evaluated.

To address this problem, we introduce and investigate the derivative of function  $\Delta\theta = f(R)$ . From (10), we will have an equation below

$$2R\Delta\theta^2 + (4R(L-R) + R - 2\Delta X_1')\Delta\theta - \Delta X_1' = 0 \quad (11)$$

Noting that  $\Delta\theta^2$  is small of higher order which can be ignored for approximation, we may have a less complicated function form by simplifying (11)

$$\Delta\theta = \frac{\Delta X_1'}{4R(L-R) + R - 2\Delta X_1'} \quad (12)$$

Then corresponding derivative can be solved. Taking the fact  $\Delta X_1' \ll R$  into consideration, the derivative is given and simplified as follows

$$\frac{d\Delta\theta}{dR} = \frac{-\Delta X_1'[4(L-2R)+1]}{[4R(L-R)+R-2\Delta X_1']^2} \approx \frac{\Delta X_1'[4(L-2R)+1]}{R^2[4(L-R)+1]^2} \quad (13)$$

Noting  $L > R$  and the following relation

$$\frac{4(L-2R)+1}{[4(L-R)+1]^2} < \frac{4(L-R)+1}{[4(L-R)+1]^2} = \frac{1}{4(L-R)+1} < 1 \quad (14)$$

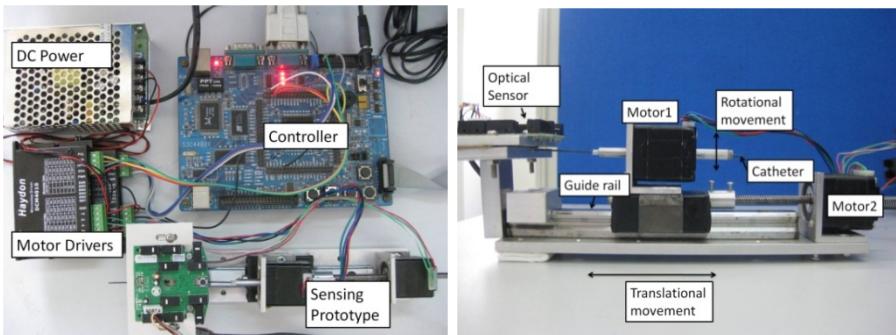
We may have

$$\left| \frac{d\Delta\theta}{dR} \right| = \left| \frac{-\Delta X_1' [4(L-2R)+1]}{[4R(L-R)+R-2\Delta X_1']^2} \right| \approx \left| \frac{-\Delta X_1'}{R^2} \right| \cdot \frac{[4(L-2R)+1]}{[4(L-R)+1]^2} < \left| \frac{-\Delta X_1'}{R^2} \right| \quad (15)$$

We may conclude that, the error brought by imperfect surface has relatively small effect on the result of measurement, so the method does have the potential to be applied in practice.

### 3 Description of Experiment

To implement the measuring method proposed and evaluate corresponding performance, a prototype of catheter sensor system is built. As is shown in Fig. 4, the prototype is mainly consisted of a mechanical guide rail and bearing, for supporting translational and rotational movement respectively. Sensor derived from an optical mouse is installed just above the guide wire to take the measurement. As we discussed in translational section, calibration should be conducted first to offset the errors from possible oblique placement.



**Fig. 4.** Principle of experiment setup (left) and prototype equipment for intervention sensing experiment (right)

The optical mouse sensor is connected to a personal computer via a USB port. The computer is utilized for measuring data acquisition and processing. Corresponding software for following experiment is developed in C++ based on common APIs related to mouse operation under Microsoft Windows platform.

During the experiment procedure, step motors are used to generate translational and rotational movement input. Every time we move the input catheter by a fixed short distance and then evaluate sensing performance from actual data of measurement we got. While the input catheter is moving, the computer simultaneously processes collected data in real time, by doing necessary adjustment and offset.

## 4 Experiment Result and Discussion

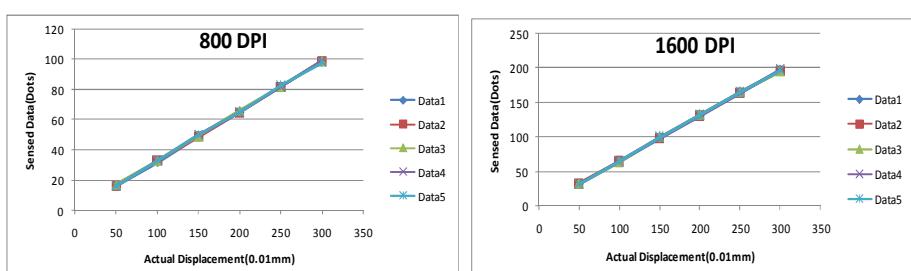
Following the procedure above, we carried out a series of experiments to evaluate the performance of the proposed method as well as our sensor system prototype. First, we move the input catheter along translational direction by a fixed incremental distance. This procedure is repeated by 5 times and corresponding diagrams and figures are shown as below, We also quantified the linearity of measurement by doing regression analysis and correlation coefficient R2 is given in Table 1, 2 and Fig. 5.

**Table 1.** Experiment Data for translational movement at sensor resolution of 800 DPI

DPI	Dis.(*0.1mm)	1		2		3		4		5	
		X	Y	X	Y	X	Y	X	Y	X	Y
800	50	0	16	0	16	-1	17	0	16	0	16
	100	-1	32	0	33	-1	33	0	33	-1	33
	150	-2	49	-1	49	-2	49	-1	50	-2	50
	200	-1	65	-1	65	-1	66	-2	65	-1	65
	250	-1	82	-2	82	-1	82	-1	82	-2	83
	300	-2	99	-1	99	-1	99	-1	99	-1	98
	Correlation Coefficient	0.99989	0.99985	0.99989	0.99971	0.99951					

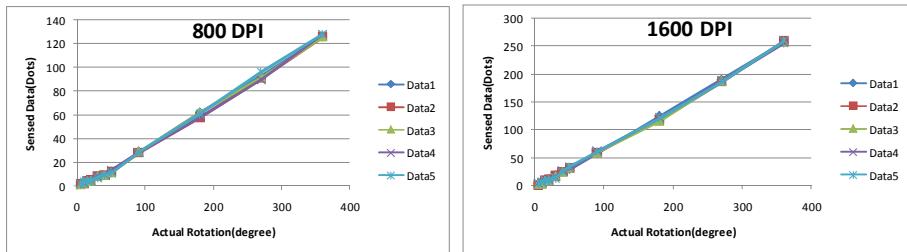
**Table 2.** Experiment Data for translational movement at sensor resolution of 1600 DPI

DPI	Dis.(mm)	1		2		3		4		5	
		X	Y	X	Y	X	Y	X	Y	X	Y
1600	50	-1	32	-1	33	-1	32	-2	32	-2	32
	100	-3	65	-3	65	-3	64	-3	65	-1	65
	150	-2	99	-2	99	-1	100	-1	98	-2	100
	200	-2	131	-3	132	-3	132	-2	131	-3	132
	250	-4	165	13	165	-3	165	-3	164	-3	165
	300	-4	197	-4	196	-3	195	-3	198	-4	197
	Correlation Coefficient	0.99993	0.99985	0.99944	0.99998	0.99983					



**Fig. 5.** Function relation of input and sensed translational movement, horizontal axis and vertical axis stand for actual displacement and sensed dots respectively

Similarly, we conduct experimental evaluation for another motion dimension by rotating the catheter in fixed incremental angle degrees. The typical range of rotational movement is a full circle and this procedure is also duplicated by 5 times as what we did for translational testing. Corresponding figures are illustrated as Fig. 6 follows.



**Fig. 6.** Function relation of input and sensed rotational movement, horizontal axis and vertical axis stand for actual rotation degrees and sensed dots respectively

The average correlation coefficients of regression analysis are 0.99792 and 0.99893 respectively. From the diagrams and figures above, we can find impressive linear relations between input and sensed movement, in both translational and rotational directions. Great accuracy and repeatability are shown by these results. Thus, the precise of sensation is sure to meet the requirement of measuring catheter motion input exerted by physicians dynamically.

We performed experiments with different DPIS of optical mouse which include levels of 400, 800, 1600 and 3200. For sake of simplicity, we only give experiment data sensed with 800 and 1600 DPI. Interestingly, our experiment shows that higher DPI does not necessarily mean higher precision. One reason for this phenomenon is that, the increase of DPI in optical mouse is achieved by promoting the magnification of filter. However, it's not free because higher magnification makes captured image blurred which harms the precision of system. For extremely high DPI like 3200, the measurement suffers from inappropriate sensitivity of the sensor so a tradeoff should be made and it's better for the system to run at a moderate DPI level.

Another thing should be noted is that, the method and prototype does not show very good performance for some first data points in rotational direction. We find that this nonlinear phenomenon possibly comes from backlash of mechanical assembly. More precise mechanical design and assembly are desired to achieve higher sensing accuracy.

Compared to the traditional ways, we find this method gain advantage on cost, space requirement and other items, as Table 3 shows. The resolution comparison is made under the assumption of the diameter of catheter and friction roller for encoder is 1mm and 2mm respectively. The new method doesn't contact with the catheter, which makes sterilization easier. As human hands have less control accuracy on rotation than translation and taken that most of the rotational torque is applied on the

catheter in the form of twisting into consideration, the rotational resolution 1 degree per count would be enough. These results suggest that optical mouse sensor is a low cost and competitive substitution for the traditional method.

**Table 3.** Comparison between optical mouse sensor and optical encoder

Compared Items	Optical Mouse Sensor	Optical Encoder	
<b>Space Requirement</b>	Only one chip with essential peripheral circuit	Need space for itself and installation	
<b>Contact with catheter</b>	No	Yes	
<b>Model</b>	Cypress CYONS1001U	Omron E6B2-CWZ6C	
<b>Specification</b>	1600DPI	3200DPI	360CPR
<b>Resolution</b>	Rot.(deg/count) Trans(mm/count)	1.8 0.016	0.9 0.008
		1 0.017	0.36 0.006
<b>Cost(Aprox)</b>	US\$10	US\$30*2	

## 5 Conclusions

In this paper, a novel optics based motion measuring method is proposed and described. The method is used to build a catheter sensor system to help physicians conduct intervention intuitively and safely. Principles and approximating conditions are discussed to ensure the feasibility theoretically. Prototype of sensor system is built and experiments are carried out to validate the performance. Results have shown good linear relation between input and sensed motion data, thus impressive precision can be expected. For future work, we will improve mechanical base to pursue better sensing accuracy. Moreover, corresponding catheter manipulator system and force feedback mechanism will be investigated and implemented to incorporate this sensor system for realizing a powerful but low cost catheter navigation system.

## References

1. Marescaux, J., Leroy, J., Gagner, M., Rubino, F., et al.: Transatlantic robot-assisted tele-surgery. *Nature* 413(6854), 379–380 (2001)
2. Tendick, F., Shankar Sastry, S., et al.: Application of micromechatronics in minimally invasive surgery. *IEEE/ASME Transactions on Mechatronics* 3(1), 34–42 (1998)
3. Ross, A.M., Segal, J., Borenstein, D., Jenkins, E., et al.: Prevalence of spinal disc disease among interventional cardiologists. *American Journal of Cardiology* 79(1), 68–70 (1997)
4. Gomes P.: Surgical robotics: Reviewing the past, analyzing the present, imaging the future. *Robot Comput. Integr. Manuf.* (2010), doi:10.1016/j.rcim.2010.06.009
5. Negoro, M., Tanimoto, M., Arai, F., Fukuda, T., et al.: An intelligent catheter system robotic controlled catheter system. *Interventional Neuroradiology* 7(supp1. 1), 111–113 (2001)
6. Carpi, F., Pappone, C.: Stereotaxis Niobe magnetic navigation system for endocardial catheter ablation and gastrointestinal capsule endoscopy. *Expert Review of Medical Devices* 6(5), 487–498 (2009)

7. Beyar, R., Wenderow, T., Lindner, D., et al.: Concept, design and pre-clinical studies for remote control percutaneous coronary intervention. *EuroIntervention* 1(3), 340–345 (2005)
8. Schiemann, M., Killmann, R., Kleen, M., et al.: Vascular guide wire navigation with a magnetic guidance system: experimental results in a phantom. *Radiology* 23(2), 475–481 (2004)
9. Thakur, Y., Cakiroglu, J.H., et al.: A device for real-time measurement of catheter-motion and input to a catheter navigation system. *Progress in biomedical optics and imaging - Proceedings of SPIE* 6509 (PART 1), article. no. 65090G (2007)
10. Wang, J., Guo, S., Kondo, H., et al.: A novel catheter operating system with force feedback for medical applications. *International of Information Acquisition* 5(1), 83–92 (2008)
11. Marcelli, E., Cercenelli, L., Plicchi, G.: A novel telerobotic system to remotely navigate standard electrophysiology catheters. *Computers in Cardiology* 35(1), 137–140 (2008)
12. Kim, C.-W., et al.: Development of 3-dimension controllable catheter to show the immediate response using both the thermoelectric module and the shape memory alloy. *Modern Physics Letters* 22(11), 1099–1104 (2008)

# Visual-Audio Integration for User Authentication System of Partner Robots

Jiangtao Cao<sup>1</sup>, Naoyuki Kubota<sup>2</sup>, Ping Li<sup>1</sup>, and Honghai Liu<sup>3</sup>

<sup>1</sup> School of Information and Control Engineering, Liaoning Shihua University, China

<sup>2</sup> Graduate School of System Design, Tokyo Metropolitan University, Japan

<sup>3</sup> School of Creative Technologies, University of Portsmouth, UK

**Abstract.** As one of non-contact biometric way for user authentication system of partner robots, visual-based recognition methods still suffer from the disturbance of light noise in the applications. Inspiring from the human’s capability of compensating visual information (looking) with audio information(hearing), a visual-audio integrating method is proposed to reduce the disturbance of light noise and to improve the recognition accuracy. Combining with the PCA-based face recognition, a two-stage speaker recognition algorithm is used to extract useful personal identity information from speech signals. With the measurement of visual background noise, the visual-audio integrating method is performed to draw the final decision. The proposed method is evaluated on a public visual-audio dataset VidTIMIT and a partner robot authentication system. The results verified the visual-audio integrating method can obtain satisfied recognition results with strong robustness.

## 1 Introduction

By the earliest motivation of building robots to take care of human being in the daily life, the researches on partner robots have been explored and developed over the recent decades. As one type of humanoid robots, its capability of communication with people is very important. With the recent innovations of hardware and software in robotics and artificial intelligence, the parter robots are growing to be more and more “smart”, especially on the way of organizing the environment information and the using of whole informational space between human being and robots. Recently, some researches have been studying on building the informational structured space for the partner robots [1][2].

In general, there are lots of information sources within the daily living environment, such as visual information, audio, temperatures, vibrations, etc. Among these information sources, visual and audio information have been widely used to teach robots to recognize or understand human being. With visual information, some impressive results have been achieved in the computer vision fields, such as face detection and recognition [3]. The function of computer vision has been one of critical part of advanced intelligent robots. With audio information, one of example is speech recognition, which attempts to understand the personal

speaking. There have been many applications, such as voice dialing, speech-to-text processing and health care of disable people, *etc.* However, based on each single information source, most of existed methods make strict assumptions on the environment characteristics because of the background noise [4]. That means, methods of computer vision are sensitive to the light noise and speech recognition systems are perturbed to the acoustic environment and sound noise. The interesting thing is, the light noise hardly affects the speech information and sound noise can not change the visual information, neither. Then a new research direction of visual-audio integrating methodologies, just like the integrating capability of human being, has been attracted more and more attention, such as lip-reading for speech recognition [5].

Inspiring from the idea of visual-audio integration, herein, a novel user authentication system for partner robots is designed to recognize the users with their visual and audio information. The authentication system is one of important function for partner robots to recognize its owner and actively contact with the user by using the related personal background knowledge. The prevailing techniques of user authentication, which involve the use of either passwords and user IDs (identifiers), or identification cards and PINs(personal identification numbers), suffer from several limitations. With these methods, there is no way to positively know who the actual user is. With the requirements of partner robots, by using authentication system, the robots will obtain the related user information which includes some special requirements or individual preference to take care. The traditional authentication policy has become inadequate. Recently, some biometrics methods have been studied for user authentication system of robots. For instance, there are some touched methods with fingerprint and untouched methods with face or speech information. This research will focus on the untouched methods with the integration of the visual information( *e.c.*, face information) and the audio information ( *e.c.*, speech) to improve the accuracy of authentication system for partner robots.

According to the requirements of real-time computing and accuracy of authentication system for partner robots, a novel visual-audio integration method is designed in this paper. Firstly, based on our previous research results on speech recognition, the two-stage features matching method is used for dealing with speech information. Secondly, one of efficient face recognition method based on PCA (principle component analysis) is implemented to deal with user's face data. By building an adaptive switching structure, the speech and face recognition process can be adaptively tuned to reduce the disturbance from background noise and improve the final accuracy rate of authentication system. For evaluating the proposed structure, a public visual-audio dataset VidTIMIT and a partner robot authentication system are implemented.

The rest of the paper is organized as follows. Section 2 presents the two-stage feature matching method for speaker recognition. Section 3 explains the face recognition process with PCA algorithm. A novel tuning structure is proposed to integrate visual-audio information in Section 4. Also the whole authentication system for partner robot is issued. Some experiments on a public dataset and

a partner robot are given in Section 5. Concluding remarks, future works are discussed in Section 6.

## 2 Speaker Recognition with Two-Stage Features Matching Method

The early automatic speaker recognition systems were made in the 1960s at Bell Labs [6]. Up to date, there have been many research projects and applications on speaker recognition systems, more details about development process can be found in several review papers [7,8,9]. With the real-time computing requirements of partner robots, one of the text-independent methods for speaker recognition, vector quantization(VQ) is used in this research. Based on the previous research results, a two-stage feature matching structure is rebuilt to improve the accuracy of speaker recognition. The main idea is represented in this section.

A basic structure for most of speaker recognition systems can be found in [8]. The first step is feature extraction. The feature extraction will transform the speech signal into feature vectors which present the specific properties of each speakers. With these feature vectors, the template models of claimed speakers(or identified speakers) are trained. In the pattern matching part, the feature vectors from the test speaker are compared with the template models. By using the similarity score and related decision module, the final identified speaker is decided. Under this general recognition framework, the two-stage feature matching method used MFCC(mel-frequency cepstral coefficients) and other three long-term speech features to recognize the speech signal. With a switching function, the feature matching process can be adaptively tuned between two different kinds of speech features. Since the recognition method can use the short-term and long-term properties of speech signal, the accuracy rate of proposed speaker recognition system is improved. The details of speech features and the two-stage matching structure are described in Section 2.1 and Section 2.2.

### 2.1 Additional Speech Feature Extraction

Generally, the MFCC features, one kind of short-term speech features have dominated most of speaker recognition systems because of their general presentation of human vocal and auditory systems [8]. However, this kind of features will ignore some dynamic changing aspects of speech signal that convey speaker information, such as prosodic information from pitch and energy contours. From the view point of speech signal changing properties, three additional speech features are extracted as further expression to convey useful speaker information and are used to the speaker recognition systems.

- **Energy entropy.** The entropy is used as a measure of energy dispersal or distribution. For speech signal, it will present possible energy changing regularity for different speakers. There are lots of implementation of entropy in different research fields. One method in [10] is represented as below. The entropy is computed by the equation 11.

$$E(x) = - \sum_{x \in X} x_i \cdot \log_2 x_i \quad (1)$$

where,  $x_i$  is the probability mass function by normalizing the spectrum in each sub-band and it can be calculated by the equation [2].

$$x_i = \frac{X_i}{\sum_{i=1}^N X_i}, \quad i = 1, 2, \dots, N \quad (2)$$

where  $X_i$  presents the energy of the  $i$ th frequency component of the spectrum.

- **Spectral flux.** As the definition in [1], spectral flux is defined as the spectral correlation between adjacent information windows. It is often used as an indication of how quickly the power spectrum of a signal is changing. For comparing the dynamic changes of speech signals, the spectral flux is calculated as one of additional speech features. Normally, it can be obtained by the Euclidean distance between two normalized spectrum. It is noted that the calculated spectral flux does not dependent on overall energy, nor present the phase information.
- **Spectral centroid.** It is used as one of properties of a spectrum in digital signal processing. It indicates where the “center of mass” of the spectrum is. For audio information, it is a simple quantification of the timbre quality of brightness [2]. One of its calculation methods is as the weighted mean of the power spectrum, with their magnitudes as the weights, which is described as below.

$$\text{Centroid} = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)} \quad (3)$$

where  $x(n)$  represents the power spectrum of component number  $n$ , and  $f(n)$  represents the center frequency of that component.

Though these features can hardly be used as independent speech feature vectors to do speaker recognition, they can play a great role in the features rematch stage which aims to classify the results after using the MFCCs features. For efficiently using these two kinds of speech features on speaker recognition, a two-stage pattern matching algorithm is designed and described in Section [2.2].

## 2.2 Two-Stage Features Matching Structure

In this scenario, two different kinds of speech features are employed to build different speech models for features matching task. In each stage, the speaker identification and verification will be carried out with related evaluating criteria.

- **The First Stage:**

After the process of features extraction, the input speech signal can be presented by its MFCCs and additional features. In the training module, the MFCCs are

used to build claimed speaker models by fuzzy c-means declustering VQ algorithm. In the speaker recognition systems, the LBG (Linde, Buzo, Gray) VQ method has been widely used because of its efficient codebook updating algorithm [13]. Herein, an improved LBG VQ method with fuzzy declustering is presented for the first stage pattern matching [4]. With the above designed VQ method, the speech features MFCCs are used to build the template models for all claimed speakers. Then the similarity between the test speaker speech signal and template models  $\theta_r$  ( $r = 1, 2, \dots, I$ ) are derived by Euclidean distance. Here,  $I$  is the number of claimed speakers.

### – The Second Stage:

After the first stage, the test speech will do the further match with limited amount of templates which satisfy the constraints of the first stage. In order to justify the impostor in this stage, a threshold from the additional features is defined. If the additional features matching distance between the test speech signal and the templates is higher than the threshold, the test speaker will be identified as the impostor. On the contrary, the test speech will do the rematch by using additional features.

Considering the characterizes of additional features and the computational cost, we only use the standard deviation and mean value of each additional feature to build the second stage feature vectors for the pattern rematch. That is, for each speech signal, its second stage feature vector only includes 8 values. The similarity of test speaker and the claimed speakers is calculated by using the score normalized Euclidean distance. The rematching process is summarized as below.

- Feature vector calculation. For the test speech signal, its energy entropy, spectral flux and spectral centroid are calculated. Then the standard deviation and mean value of these features are obtained and the second stage feature vector is obtained.
- Pattern rematch. By using the Euclidean distance, the matching distance among the test speech feature vector and the claimed speakers' feature vectors are calculated.
- Score normalization. For avoiding the score variability from the phonetic content, the duration and the environment noise, one of the score normalization techniques is employed as equation 4

$$\hat{D}_x(\lambda_i) = \frac{D_x(\lambda_i) - \bar{\mu}_{\lambda_i}}{\bar{\sigma}_{\lambda_i}} \quad (4)$$

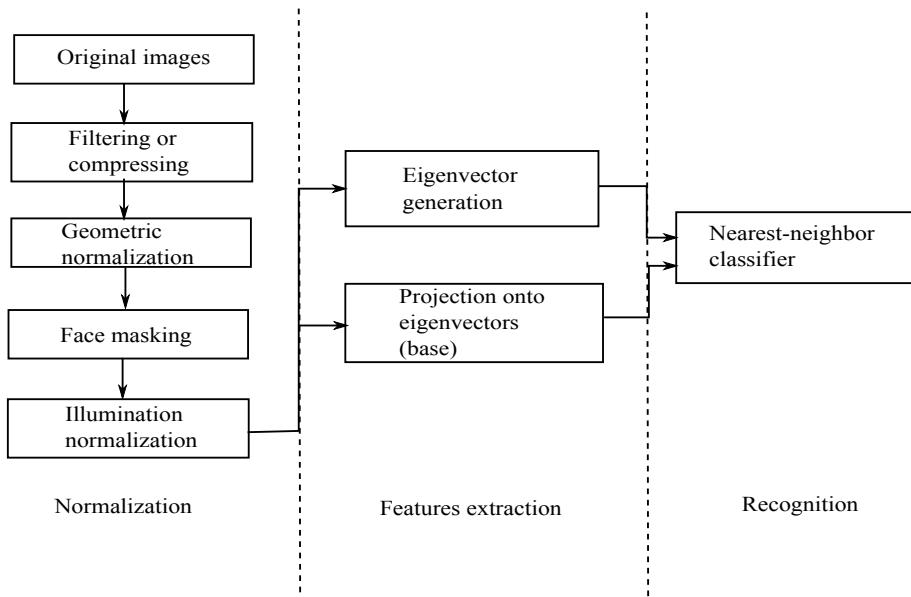
where,  $D_x(\lambda_i)$  is the matching distance between the test speech  $x$  and the template  $\lambda_i$ ,  $i$  belongs to the number of second stage templates.  $\bar{\mu}_{\lambda_i}$  and  $\bar{\sigma}_{\lambda_i}$  are the mean values of the template features' standard deviation and mean, respectively.

- Final decision. Choosing the template with minimum normalized score as the best matched speaker, the final recognition task is done.

### 3 Face Recognition Algorithm

As one of the few biometric methods that possess the merits of both high accuracy and low intrusiveness, face recognition technology has a variety of potential applications in information security, law enforcement and surveillance, smart cards, access control, *etc.* Also it has received significantly increased attention from both the academic and industrial communities during the past 20 years. Among these different methods, PCA-based face recognition method (also known in the literature as eigenfaces method) is used in this research.

PCA-based face recognition methods are quite popular and have been a benchmark algorithm, because of the ease of implementing them and their reasonable performance level [15]. Further details and new researches on PCA-based face recognition can be found in [16]. A brief introduction of PCA-based face recognition is as follows. The face recognition structure is as Figure 1.



**Fig. 1.** The general structure of PCA-based face recognition

With the PCA method, the input is a training set of  $N$  facial images. Each image is interpreted as a point in the space  $\mathbb{R}^{n \times m}$ , where each image is  $n$  by  $m$  pixels. PCA method finds the optimal linear least-squares representation in the  $N - 1$  dimensional space, with the representation preserving variance. Then the face information can be represented by a set of  $N - 1$  eigenvectors( normally be referred to eigenfaces,  $e_1, \dots, e_{N-1}$  ) and eigenvalues( $\lambda_1, \dots, \lambda_{N-1}$ ). By normalizing the eigenvectors to be orthonormal, the eigenvectors are arranged as  $\lambda_i > \lambda_{i+1}$ . Since the low-order eigenvectors encode the larger variations of

face information and the higher-order eigenvectors encode smaller variations, the higher-order eigenvectors are excluded from the original representation space as background noise. Then the training face image can be represented by a low-dimension eigenvectors  $M < N - 1$ . All  $K$  training facial images are compressed as  $g_1, \dots, g_K$  in face space. The identity of test image is finally decided by minimizing the Euclidean distance between the test eigenvector  $p$  and the training templates  $g_j$ , ( $j = 1, \dots, K$ ).

## 4 Visual-Audio Integration for Authentication System

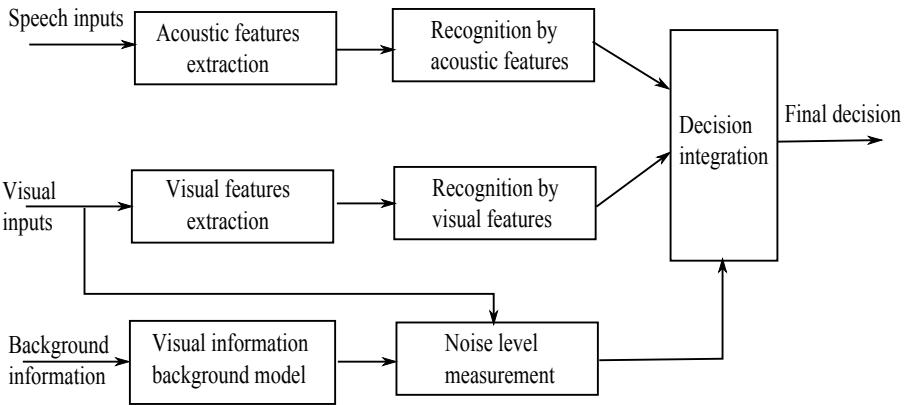
As mentioned in [4], the nature of human's perception demonstrates a statistical advantage of bimodality: when humans have estimates of an environmental property from two different sensory systems, any of which is possibly corrupted by noise, they combine the two signals in the statistically optimal way so that the variance of the estimates for the property is minimized after integration. Inspired from this capability of human, some visual-audio integration methods have been proposed. These methods can be roughly categorized into two types: feature integration and decision integration. In the feature integration, the features from the two information sources are combined and the final task is based on the combined features. In the decision integration, the features of each modality are used for recognition separately and the related outputs from different recognizers are integrated into the final decision [17,4]. Considering the requirements of partner robots, the second type of integration methodology is used in this research.

The primary task of integration is to obtain the performance which is equal to or better than the performance of any modality for various noise conditions. When the visual background noise level is high, the acoustic modality performs better than the visual one, thus the integration performance should be better than visual recognition performance, vice versa. With the assumption of class-conditional independence, *i.e.*, the two information sources are evaluated independently, the proposed visual-audio information integration framework is shown in Figure 2. The background noise of visual information can be estimated based on a piecewise smooth image prior model and measured CCD camera response functions [18]. With the estimated upper bound of image noise, the adaptive weighted integration in final decision can be performed as the following rule:

$$C^* = \arg \max_i \{ \lambda f(O_{audio,i}) + (1 - \lambda) f(O_{visual}, i) \} \quad (5)$$

where,  $O_{audio}$  and  $O_{visual}$  are the extracted features from audio and visual inputs, respectively, and the  $f(O_{audio,P_i})$  and  $f(O_{visual}, P_i)$  are the normalized likelihood outputs for the  $i$ -th class. The integrating weight  $\lambda \in [0, 1]$  determines how much the final decision relatively depends on each modality. The weight value can be tuned according to the noise level in the visual information.

The whole visual-audio integration method for authentication system of partner robots can be summarized as below.



**Fig. 2.** The general structure of visual-audio integration

- Capturing the visual information, the face detection is performed.
- Estimating the image noise and comparing the noise level with the defined threshold value. If the noise is higher, then turn to next step, otherwise, the PCA-based face recognition is performed. The final result is decided.
- Capturing the audio information, the speech features are extracted.
- Performing the speaker recognition and face recognition by using speech features (MFCC) and visual features (eigenfaces), separately.
- Following the tuning rule in equation 5, the weighted likelihood values are obtained.
- By reordering all the integrated likelihood values, the final task can be done.

## 5 Experiments

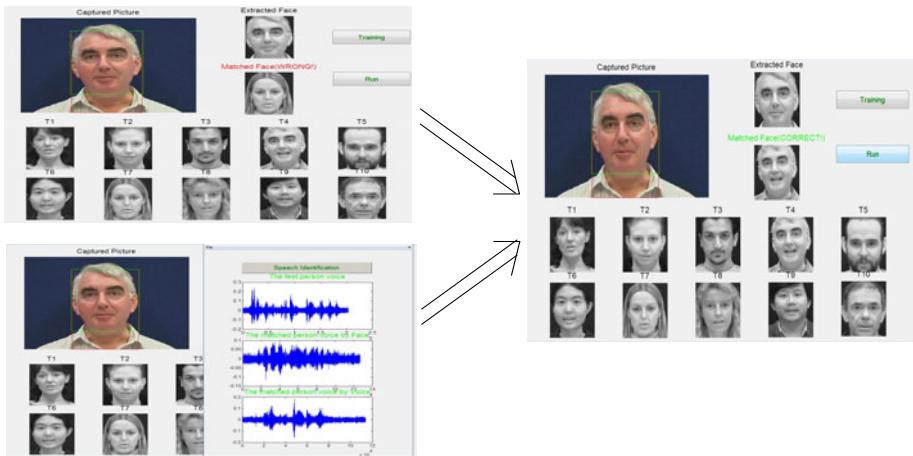
In this section, the proposed visual-audio integration method is tested on the public dataset: VidTIMIT audio-video dataset. The main aim of this test is to verify the performance of visual-audio integration is better than each single modality. Since there is no camera details, the speech background noise from NOISEX database is added to test the robustness on noise disturbance. From application point, the proposed method is implemented on the authentication system of partner robot: ApriPoco. Meanwhile, the separated face recognition and speaker recognition methods are compared with the proposed method.

### 5.1 Tests on VidTIMIT Dataset

The VidTIMIT dataset is comprised of video and corresponding audio recordings of 43 people, reciting short sentences [19]. The dataset was recorded in 3 sessions, with a mean delay of 7 days between Session 1 and 2, and 6 days between Session 2 and 3. The sentences were chosen from the test section of the TIMIT corpus. There are 10 sentences per person. The first six sentences (sorted

alphanumerically by filename) are assigned to Session 1. The next two sentences are assigned to Session 2 with the remaining two to Session 3. In the video, each person performed a head rotating sequence in each session. The sequence consists of the person moving their head to the left, right, back to the center, up then down and finally return to center. All recordings were done in an office environment using a broadcast quality digital video camera. The video of each person is stored as a numbered sequence of JPEG images with a resolution of  $512 \times 384$  pixels. The corresponding audio is stored as a mono, 16 bit, 32 KHz WAV file.

We used the 5 females and 5 males from total 43 people as the training and test dataset. For each person, 7 images from Session 1, 7 images from Session 2 and 6 images from Session 3 are chosen as training data for face recognition. Each sentence is divided into 3 seconds fragments. And the first 2 fragments are used as training data for speaker recognition. From the other images, 200 images from Session 1, 2 and 3 are chosen to be tested. All the other speech signals are used as the related audio inputs. For face recognition, the PCA-based algorithm in [3] is used. For speaker recognition, the improved VQ method in Section 2 is implemented. The MFCCs order is 20, and the codebook size of VQ method is 32. One example of proposed method on VidTIMIT is shown in Figure 3. In this figure, by using speech information, the final recognition result is corrected from the wrong result of face-based recognition.



**Fig. 3.** One example test of visual-audio integration method on dataset VidTIMIT

The final recognition accuracy rate(RAR) of different methods were shown in Table 1.

The comparison results in Table 1 showed the proposed method obtained the higher accuracy rate with or without additional speech noise. When there was no additional background noise, the integration method achieved the best results for the recognition task.

**Table 1.** SRR performance of different recognition methods

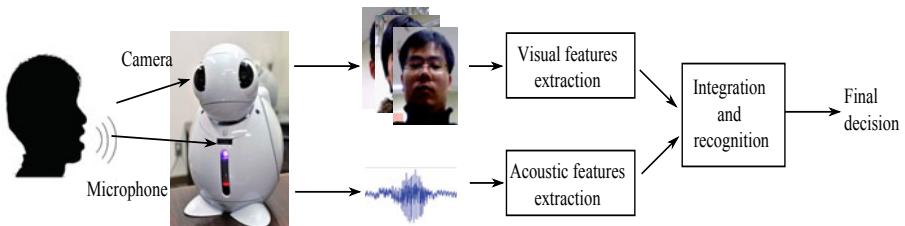
Speech Noise	No additional noise	White noise (NOISEX)	Speech babble(NOISEX)
Image only	84.5%	84.5%	84.5%
Speech only	77.2%	68.1%	60.1%
Proposed method	96.5%	87.5%	84.5%

For testing the robustness of integrated methods, two background noise signals (White noise and Speech babble) were extracted from the samples of NOISEX database and added into the speech files. The test results with additional speech noise indicated the noise significantly disturbed the speech recognition, and there was no effect on the face recognition. The RAR of proposed method also decreased with the added background noise. However, its results were better or equal to the other two single modalities.

## 5.2 Test on a User Authentication System of Partner Robot

The proposed method was implemented in a partner robot project which aimed to take care elderly people by building a general informational structured space with many kinds of sensors. As part of this partner robot project, there have been a informational database which included visual and audio information from different users. The framework of visual-audio integration for authentication system of partner robot was shown in Figure 4.

In this experiment, the database included 300 face images and related 150 speech files(fixed content) from 5 different users. The camera and microphone were integrated into the partner robot body. The frame rate of camera was 25 fps. Each frame was  $640 \times 480$  pixels. The sampling frequency of microphone was 16kHz. For simulating the changing light conditions, the strong light level(at noon) was defined as the standard working condition. The weak light levels(in the morning and in the evening) were used to compare with the standard condition. The RAR of three recognition methods on these test visual and audio information were summarized in Table 2.



**Fig. 4.** The framework of visual-audio integration for authentication system of partner robot

**Table 2.** RAR of authentication system by different methods

Background light condition	In the morning	At noon	In the evening
Image only	76.5%	90.5%	80.5%
Speech only	82.6%	82.6%	82.6%
Proposed method	94.5%	98.5%	96.5%

It can be seen from Table 2 that the proposed visual-audio integration method consistently outperformed the other two methods on the different background. And its performance was good enough to be used for real-time application of partner robot. It was also shown that proposed method can efficiently get over the disturbance from the change of background light.

## 6 Concluding Remarks

The experiment results have shown that the proposed method was succeeded in integrating visual and acoustic information to improve the accuracy of user authentication system. By using the visual-audio integrating structure, the user's face and speech features are extracted and combined to deal with the disturbance of background noise.

As mentioned before, the visual-audio information integration can also be performed from features level, not the decision level. In future, the direct integration of visual features and acoustic features will be considered into the partner robot authentication system. More flexible and robust reasoning and recognition methods will also be considered in the future research [20][21].

## Acknowledgment

The authors would like to thank Mr. Tang Dalai and Mr. Oboo Takenori for providing the visual-audio dataset of the partner robot ApriPoco.

## References

1. Kubota, N., Shimizu, T., Abe, M.: Join attention between a human being and a partner robot based on computational intelligence. *Journal of Advanced Computational Intelligence and Intelligent Informatics* 11, 1274–1280 (2007)
2. Kubota, N., Nishida, K., Masuta, H.: The role of prediction in structured learning of partner robots. In: *IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, pp. 1901–1906 (2007)
3. Moon, H., Phillips, P.: Computational and performance aspects of PCA-based face-recognition algorithms. *Perception* 30(3), 303–322 (2001)
4. Lee, J.S., Park, C.H.: Adaptive decision fusion for audio-visual speech recognition. *Speech recognition, technologies and applications*, I-tech, 275–296 (2008)

5. Kaucic, R., Dalton, B., Blake, A.: Real-time lip tracking for audio-visual speech recognition applications. In: Buxton, B.F., Cipolla, R. (eds.) ECCV 1996. LNCS, vol. 1065, pp. 376–387. Springer, Heidelberg (1996)
6. Pruzansky, S.: Pattern-matching procedure for automatic talker recognition. *The Journal of the Acoustical Society of America* 35, 354–358 (1963)
7. Furui, S.: Fifty years of progress in speech and speaker recognition. *The Journal of the Acoustical Society of America* 116, 2497–2506 (2004)
8. Bimbot, F., Bonastre, J., Fredouille, C., Gravier, G., Magrin-Chagnolleau, I., Meignier, S., Merlin, T., Petrovska-Delacrétaz, D., Reynolds, D.: A tutorial on text-independent speaker verification. *EURASIP Journal on Applied Signal Processing*, 430–451 (2004)
9. Kinnunen, T., Li, H.: An overview of text-independent speaker recognition: From features to supervectors. *Speech Communication* 52, 12–40 (2010)
10. Toh, A., Tognetti, R., Nordholm, S.: Spectral entropy as speech features for speech recognition. *Proceedings of PEECS*, 22–25 (2005)
11. McAdams, S.: Perspectives on the contribution of timbre to musical structure. *Computer music journal* 23(3), 85–102 (1999)
12. Schubert, E., Wolfe, J.: Does timbral brightness scale with frequency and spectral centroid? *Acta Acustica united with Acustica* 92(5), 820–825 (2006)
13. Linde, Y., Buzo, A., Gray, R., et al.: An algorithm for vector quantizer design. *IEEE Transactions on communications* 28(1), 84–95 (1980)
14. Pham, T., Brandl, M., Beck, D.: Fuzzy declustering-based vector quantization. *Pattern Recognition* 42(11), 2570–2577 (2009)
15. Turk, M., Pentland, A.: Face recognition using eigenfaces. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, vol. 591, pp. 586–591 (1991)
16. Yang, J., Zhang, D., Frangi, A., Yang, J.: Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(1), 131–137 (2004)
17. Rogozan, A., Deléglise, P.: Adaptive fusion of acoustic and visual sources for automatic speech recognition. *Speech Communication* 26(1-2), 149–161 (1998)
18. Liu, C., Freeman, W.: Noise estimation from a single image. In: Proceedings of IEEE Conference on CVPR, pp. 901–908 (2006)
19. Sanderson, C.: Biometric person recognition: Face, speech and fusion. VDM Verlag Dr. Müller (2008)
20. Ji, X., Liu, H.: Advances in view-invariant human motion analysis: A review. *IEEE Transactions on Systems, Man and Cybernetics, Part C* 40(1), 13–24 (2010)
21. Kubota, N., Liu, H., Mori, Y.: Visual perception for emotion-based communication of a partner robot. *Kansei Engineering International* 8(1), 41–50 (2009)

# Synthetic Finger Phalanx with Lifelike Skin Compliance

John-John Cabibihan and Shuzhi Sam Ge

Social Robotics Laboratory (Interactive and Digital Media Institute) &  
Dept. of Electrical and Computer Engineering,

National University of Singapore

{elecjj, samge}@nus.edu.sg

**Abstract.** Touching for prosthetics and social robotics is important for functional and social reasons. Current synthetic finger designs, however, are too stiff when compared to the human touch. We ask whether it would be possible for synthetic skins to mimic the skin compliance of the human finger phalanx. In this paper, finite element simulations were conducted on finger phalanx models with open pockets of 1 mm and 2 mm heights in the internal geometry. With an open pocket design, simulation results show that it is possible to have force-displacement results that are comparable to the published data of the human finger phalanx under cyclic loading.

**Keywords:** Synthetic Skins, Prosthetics, Social Robotics, Rehabilitation Robotics, Finite Element Analysis.

## 1 Introduction

Skin compliance is the property that allows large displacements to occur under low contact forces. This paper investigates the effect on skin compliance by having open pockets on the internal geometry of a finger phalanx model as can be observed in the force-displacement results. The structure of the human fingertip is a highly integrated and layered structure [1]. The layers of the finger skin consist of the stratum corneum, epidermis, dermis, a layer of fat, the bone and the nail. These layers can be modelled with hyperelastic, viscoelastic and poroelastic behaviors [2].

A compliant synthetic skin will have significant functional and social implications. For grasping objects by a prosthetic or robotic hand, large areas can be achieved even with low contact forces. Biomechanical studies on the human fingers [3, 4] suggest that a 1 N indentation force results in about 2 mm displacement of the skin tissue. As the contact force is increased, the skin tissue becomes stiff and the contact area changes minimally. This permits a more stable grasp.

As social beings, humans acknowledge the presence of other humans through greetings. Apart from gaze, nodding and hellos, humans also exchange greetings by touching. These are typically in the form of a handshake, a hug or a touch on the arm. It can be expected that prosthetic and humanoid hands will touch and will be touched by others. With the skin as the main interface during the touching interactions, soft

contact would be a reasonable requirement for reasons of safety and comfort [5, 6]. The perception of softness can be attributed to skin compliance [7].

The next section describes the methods used for the simulations. This consists of the synthetic materials used and the constitutive equations to model them. The finger phalanx geometries and the indentation procedures are then described. Following these are the results and lastly, the conclusions.

## 2 Methods

### 2.1 Materials

Silicone and polyurethane materials were used. These two materials have been earlier described in the fingers of robotic [8] and prosthetic hands [9]. The characterization data for these materials have been described elsewhere [5, 10, 11]. The silicone sample has a Shore A durometer value of 11 while the polyurethane has a value of 45.

### 2.2 Constitutive Equations

The constitutive equations used here were the same as the one used in [5], which was modified from simulation models of the human fingertip in [2]. The total stress is equal to the sum of the hyperelastic (*HE*) stress and the viscoelastic (*VE*) stress such that:

$$\sigma(t) = \sigma_{HE}(t) + \sigma_{VE}(t) \quad (1)$$

where  $t$  is the time. The hyperelastic behavior was derived from a function of strain energy density per unit volume,  $U$ .

$$U = \sum_{i=1}^N \frac{2\mu_i}{\alpha_i^2} \left[ \lambda_1^{\alpha_i} + \lambda_2^{\alpha_i} + \lambda_3^{\alpha_i} - 3 + \frac{1}{\beta} (J^{-\alpha_i \beta} - 1) \right] \quad (2)$$

$$\sigma_{HE} = \frac{2}{J} F \frac{\partial U}{\partial C} F^T \quad (3)$$

where  $J = \lambda_1 \lambda_2 \lambda_3$  is the volume ratio,  $\alpha_i$  and  $\mu_i$  are the hyperelastic material parameters,  $\beta = v/(1 - 2v)$  where  $v$  is the Poisson's ratio,  $N$  is the number of terms used in the strain energy function, and  $F$  and  $C$  are the deformation gradient and the right Cauchy-Green deformation tensors, respectively. It was assumed that the candidate materials were incompressible, and therefore  $J$  was set to unity. In the case of uniaxial compression tests, the following were used for the identification:  $\lambda_1 = \lambda$ ,  $\lambda_2 = \lambda_3 = \lambda^{1/2}$ .

The viscoelastic behavior was defined as follows, with a relaxation function  $g(t)$  applied to the hyperelastic stress:

$$\sigma_{VE} = \int_0^t \dot{g}(\tau) \sigma_{HE}(t - \tau) d\tau \quad (4)$$

In order to describe several time constants for the relaxation, the stress relaxation function  $g(t)$  was defined using the Prony series of order  $N_G$ , where  $g_i$  and  $\tau_i$  are the viscoelastic parameters:

$$g(t) = \left[ 1 - \sum_{i=1}^{N_G} g_i (1 - e^{-t/\tau_i}) \right] \quad (5)$$

The number of coefficients to identify is equal to  $2(N+N_G)+1(v)$  for each material type. The coefficients for hyperelastic ( $N$ ), stress relaxation ( $N_G$ ) and Poisson's numbers ( $v$ ) are given in the Table 1.

**Table 1.** Coefficients of the Synthetic Materials [5]

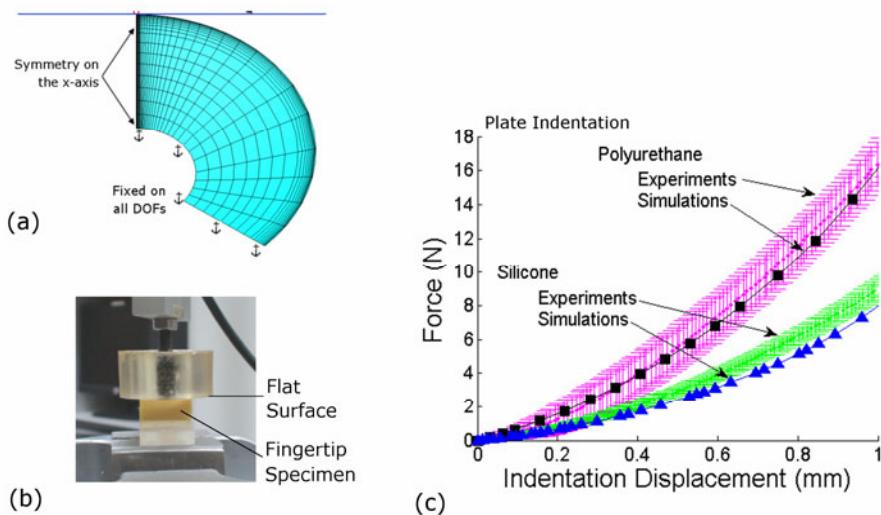
<i>i</i>	1	2	3
Silicone ( $v = 0.49$ )			
$g_i$	0.015	0.044	0.029
$\tau_i$ (s)	0.025	0.150	0.300
$\mu_i$ (MPa)	0.080	0.010	-
$\alpha_i$	0.001	15.500	-
Polyurethane ( $v = 0.47$ )			
$g_i$	0.167	0.158	0.113
$\tau_i$ (s)	0.100	1.380	25.472
$\mu_i$ (MPa)	0.100	0.063	-
$\alpha_i$	5.500	8.250	-

The indentation and imaging tests demonstrated that the constitutive equations can well describe the behaviour of the silicone and polyurethane samples [5]. Results from the finite element models and actual experiments on synthetic finger phalanges were compared in the earlier paper. The indentation set-up and the plate indentation experiments are provided in Fig. 1 for the reader's convenience.

### 2.3 Finite Element Modeling

Considering the good agreement of the simulation results with the experimental results in [5], two types of geometries were prepared for the simulations in this paper. The geometry of the fingertip model with silicone material behaviour has an arc-shaped pocket with a 1 mm height (Fig. 2a). Polyurethane, which is the stiffer among the two, was modified to have a pocket height of 2mm (Fig. 2b).

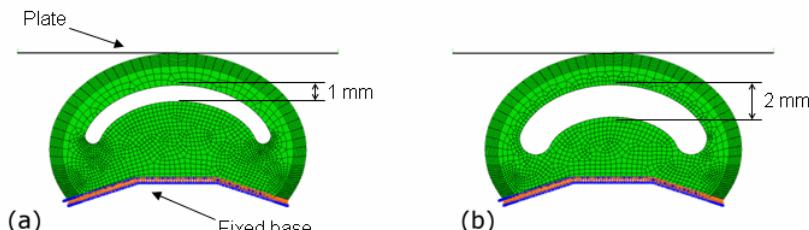
The eight-node plain strain bi-quadratic elements were used to model the fingertip. A thickness of 10 mm (dimension towards the page) is set. The base of the finger geometry was constrained in all degrees of freedom. This base region corresponds to where the bone and fingernail are located and where the structural and electronic components can be placed. The geometries were modelled using the commercial finite element analysis software Abaqus™ /Standard 6.7-1 (Simulia, Providence, RI). The simulations were run at the Supercomputing and Visualisation Unit of the Computer Centre, National University of Singapore.



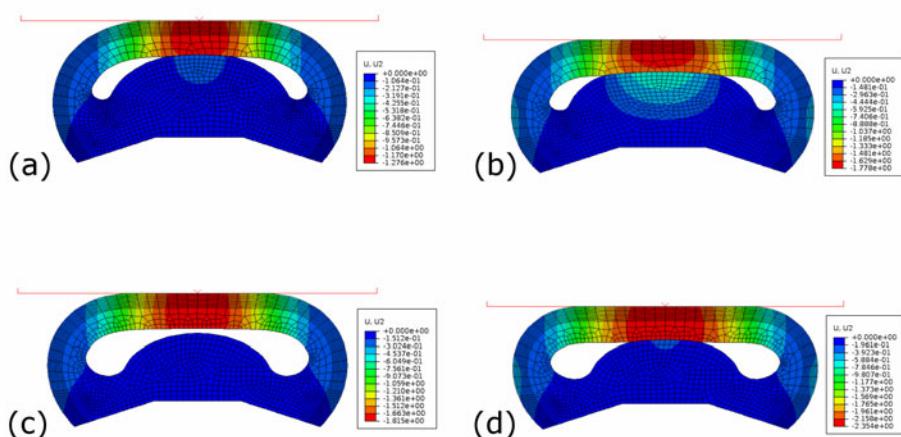
**Fig. 1.** Simulation and validation experiments for the solid internal geometry model. (a) Finite element model. (b) Indentation set-up of the finger phalanx specimen being compressed by a flat surface in a universal testing machine. (c) Validation results for the simulation model and the indentation experiments. (With kind permission from Springer Science+Business Media: Int J Soc Robot, Towards Humanlike Social Touch for Sociable Robotics and Prosthetics, v1, 2009, 29-40, Cabibihan et al, Figures 2, 3 and 5).

## 2.4 Simulation Conditions

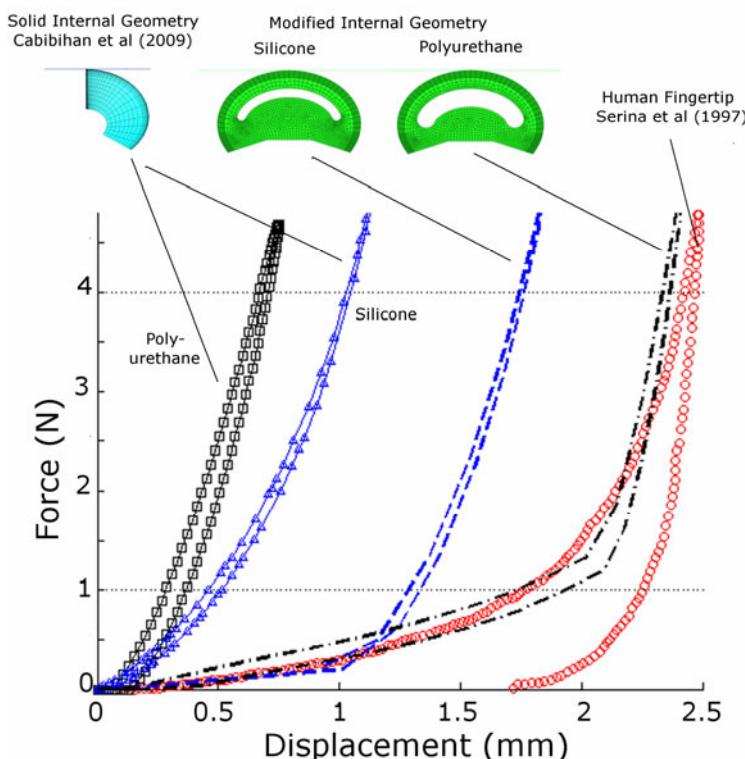
These loading conditions are similar to the conditions used in obtaining the human fingertip skin compliance in [3]. Similar loading conditions were applied for synthetic skin in [5]. In the experiment with the human fingertip, the human subjects repeatedly tapped their index finger on a flat plate while the contact force and pulp displacement were measured. Subjects were instructed to match the applied force history to the positive half of a 0.5 Hz sine wave with amplitude of  $5 \pm 0.5$  N. Displacements at 1 N and 4 N were obtained. In the simulations for this paper, the models were loaded in similar cyclic loading conditions.



**Fig. 2.** The finite element model of the finger phalanges. The geometries of the 2D finite element fingertip models with (a) silicone and (b) polyurethane behaviors, respectively.



**Fig. 3.** The displacement contour plots of the finger phalanges when indented with a cyclic loading having 5 N amplitude. (a and b) silicone with 1 mm-height open pocket at 1 N and 4 N, respectively. (c and d) polyurethane with 2 mm-height open pocket at 1 N and 4 N, respectively.



**Fig. 4.** Force-displacement curves of synthetic and human fingertips

### 3 Results

The displacement contours at the 1 N and 4 N indentations are shown on Fig. 3. From these figures, it is apparent that there was a large contact surface under the plate. At 1 N load, the silicone finger phalanx model with 1 mm height pocket achieved a 1.28 mm displacement (Fig. 3a); for the polyurethane model with 2 mm height pocket, the displacement is 1.82 mm (Fig. 3c). At 4 N load, silicone has a 1.78 mm displacement (Fig. 3b) while polyurethane has a 2.35 mm displacement (Fig. 3d). These results are plotted in Fig. 4.

These results can be appreciated with reference to the published force-displacement outcomes of the homogeneous geometry conditions for synthetic fingers and the human finger. In [5], a solid finger phalanx geometry was modelled as symmetric in the vertical axis. Fig. 4 also shows the simulation results and the human finger data.

For the model with the solid internal geometry, a 1 N indentation resulted in a 0.45 mm displacement for silicone and a 0.28 mm displacement for polyurethane. For the human finger, the displacement for the solid geometry was 1.83 mm [3]. At 4 N indentation, the displacements are 1.03 mm and 0.67 for silicone and polyurethane, respectively. For the human finger, it was 2.91 mm [3].

### 4 Conclusions

A method for improving the skin compliance of synthetic skins was presented in this paper. Through finite element simulations, we altered the internal geometry of finger phalanx models by introducing open pockets with 1 mm height for the silicone material and 2 mm height for the polyurethane material. Overall, there was an increase in displacement results when open pockets were introduced. Notably, it was possible to make a synthetic skin more compliant even with a stiff synthetic material such as polyurethane. Recall that the polyurethane material has a Shore durometer value of 45 while the silicone has a value of 11; the lower value signifies the lower resistance of an indenting probe to the material. More importantly, the simulation results showed that it is possible to mimic the skin compliance of the human finger phalanx with synthetic materials. It remains to be seen whether the synthetic skin with open pocket designs will be able to trick a blindfolded human subject in perception experiments into believing that the synthetic skin is lifelike.

### Acknowledgements

This work has been funded under the project “Social Robots: Breathing Life into Machines” by the National Research Foundation, Singapore.

### References

1. Lanir, Y.: Skin mechanics. In: Skalak, R., Chien, S. (eds.) *Handbook of Bioengineering*, pp. 1–25. McGraw-Hill, New York (1987)
2. Wu, J.Z., Dong, R.G., Rakheja, S., Schopper, A.W., Smutz, W.P.: A structural fingertip model for simulating of the biomechanics of tactile sensation. *Medical Engineering and Physics* 26, 165–175 (2004)

3. Serina, E., Mote, S., Rempel, D.: Force response of the fingertip pulp to repeated compression - effects of loading rate, loading angle and anthropometry. *Journal of Biomechanics* 30, 1035–1040 (1997)
4. Pawluk, D.T., Howe, R.D.: Dynamic lumped element response of the human fingerpad. *Journal of Biomechanical Engineering* 121, 178–183 (1999)
5. Cabibihan, J.J., Pattofatto, S., Jomaa, M., Benallal, A., Carrozza, M.C.: Towards humanlike social touch for sociable robotics and prosthetics: Comparisons on the compliance, conformance and hysteresis of synthetic and human fingertip skins. *International Journal of Social Robotics* 1, 29–40 (2009)
6. Cabibihan, J.J., Pradipta, R., Chew, Y.Z., Ge, S.S.: Towards humanlike social touch for prosthetics and sociable robotics: Handshake experiments and finger phalange indentations. In: Kim, J.H., Ge, S.S., Vadakkepat, P., Jesse, N. (eds.) *FIRA RoboWorld Congress 2009*. LNCS, vol. 5744, pp. 73–79. Springer, Heidelberg (2009)
7. Friedman, R.M., Hester, K.D., Green, B.G., LaMotte, R.H.: Magnitude estimation of softness. *Exp. Brain Res.* (2008)
8. Beccai, L., Roccella, S., Ascari, L., Valdastri, P., Sieber, A., Carrozza, M.C., Dario, P.: Development and experimental analysis of a soft compliant tactile microsensor for anthropomorphic artificial hand. *Mechatronics, IEEE/ASME Transactions on Mechatronics* 13, 158–168 (2008)
9. Edin, B.B., Ascari, L., Beccai, L., Roccella, S., Cabibihan, J.J., Carrozza, M.C.: Bio-inspired sensorization of a biomechatronic robot hand for the grasp-and-lift task. *Brain Research Bulletin* 75, 785–795 (2008)
10. Cabibihan, J.J., Pattofatto, S., Jomaa, M., Benallal, A., Carrozza, M.C., Dario, P.: The Conformance Test for Robotic/Prosthetic Fingertip Skins. In: *Proceedings of the First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics*, pp. 561–566 (2006)
11. Cabibihan, J.J., Carrozza, M.C., Dario, P., Pattofatto, S., Jomaa, M., Benallal, A.: The Uncanny Valley and the Search for Humanlike Skin Materials for a Prosthetic Fingertip. In: *Proceedings of IEEE-RAS Humanoid Robotics*, Genova, Italy, pp. 474–477 (2006)

# **Study on Educational Application of Android Robot SAYA: Field Trial and Evaluation at Elementary School**

Tskuya Hashimoto, Naoki Kato, and Hiroshi Kobayashi

Department of Mechanical Engineering, Faculty of Engineering,  
Tokyo University of Science,  
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan  
[{tak,kato,hiroshi}@kobalab.com](mailto:{tak,kato,hiroshi}@kobalab.com)

**Abstract.** Android robots that have a human-like appearance have been developed recently. We have been also developing an android-type communication robot SAYA equipped with the face robot that can express facial expressions. This paper shows an application of the android robot SAYA in educational fields as medium of a remote class support system. In this system, SAYA plays a role of a teacher, and the principle of leverage is adopted as a topic of science class for elementary school students. Moreover, we conducted the field trial at elementary schools to verify the effectiveness of the system. As a result, we confirmed its positive effects and possibility in elementary schools.

**Keywords:** Human-robot interaction, Android robot, Tele-operation, Educational application, Field trial.

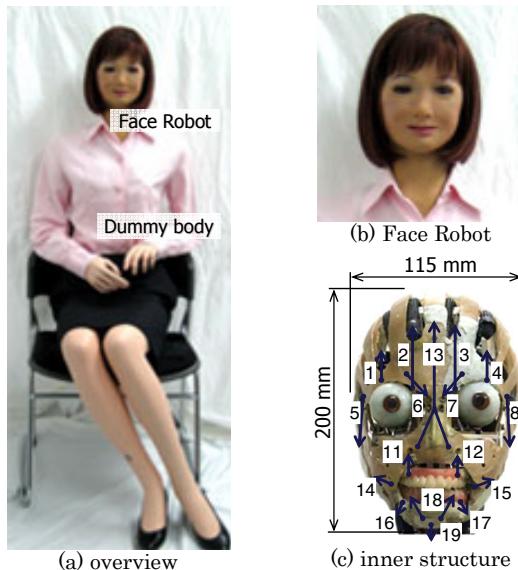
## **1 Introduction**

Coexistence of humans and robots might be realized because of progress of science and technology in recent years. Therefore investigating what kind of intelligence, functions, and mechanisms that robots should possess to realize smooth communication with humans is very interesting. In order to toward that purpose, many kinds of communication robots have been developed. They also have been applied to our daily lives and interaction manners between humans and robots have been investigated. For instance, some animal-type robots were commercialized and they were actually used in our living space for entertainment [1] or mental therapy [2]. In particular, effectiveness of seal type robot [2] for elderly people was verified in nursing-care facilities. Animal type robots realize emotional communication with human by taking adorable actions. On the other hand, humanoid-type robots with a head and arms also have been developed to realize more human-like communication. Researchers of humanoid robots have studied on bodily motions such as gestures, nodding, eye-direction, facial expressions and so on actively as non-verbal behaviors to realize natural communication with human [3]-[6]. For example, Robovie [3][4] was actually used at a station and a museum, and it was able to interact with human and offer services. Furthermore, communication robots are also used in educational field [7]-[9], and they can teach students and learn with students through communication. A one of

merits of these challenges is that children might be interested in science and technology through interaction with robots. In particular, since Japanese young people are moving away from the science and the technology by PISA's and TIMSS's survey, it seems that having elementary school children interact and communicate with robots is effective to encourage children to be interested in science and technology.

In this study, we propose educational system with communication robot for educational fields, particularly for elementary schools. In addition, we use android robots that have human like appearance as an interface of the system. Android robots also have been developed as a new medium of communication robots [10]-[14]. For example, Kobayashi et al. focused on importance of facial expressions in non-verbal communication and realized human-like facial expressions with the face robot that resembled human-like face [10][11]. And also, we have developed the android robot SAYA with that face robot [11] (Fig. 1), and it works at university entrance as a receptionist robot [12]. Then, we use SAYA as a medium of educational system. In addition, although intelligence technologies for communication robots have not been practical yet, it seems that tele-communication robot is more practical because its intelligence is handled by the operator. Thus, we study on remote class support system with the android robot SAYA as a practical educational system. Furthermore, we adopt the principle of leverage as a topic of science class that is offered to elementary school students. Then, we carried out field trials at elementary schools, and we evaluate effectiveness of the education with the android robot SAYA.

Chapter 2 describes the mechanical structure and capability of android robot SAYA. Chapter 3 shows the remote-class support system with SAYA and detailed contents and flow of the class. In chapter 4, detailed procedure of field trial and its results are shown in chapter 5. Finally, conclusions are described in chapter 6.



**Fig. 1.** Android Robot SAYA



**Fig. 2.** Examples of SAYA's facial expressions

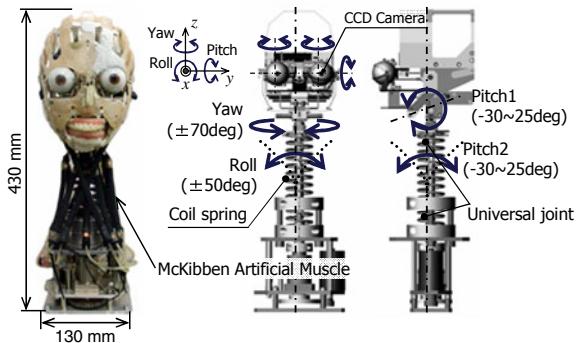
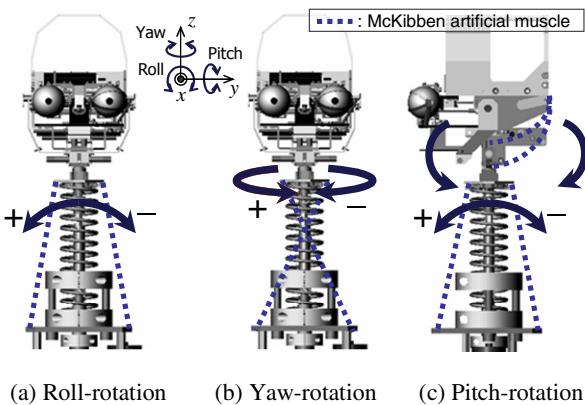
## 2 Android Robot SAYA

We have developed the android robot SAYA as a communication robot. In this robot, we equip the face robot that can express human-like facial expressions on a dummy body as a head part. Therefore, it looks like a real human as shown in Fig. 1(a). And also, we apply SAYA to not only the reception in our university entrance by implementing speech dialogue system [12] but also educational applications.

### 2.1 Face Robot

In our face-to-face communication, it is said that facial expressions play a very important role as a one of non-verbal communication [15]. Thus, facial expressions are assumed to be an important media for realizing smooth communication between human and robot. Therefore, we have developed highly realistic face robots that can express human-like facial expressions [10][11]. In the face robot, Facial Action Coding System (FACS) [16] proposed by P. Ekman et al is referred as an practical method for generating human-like facial expressions. Then, 19 control points on the face of the robot are selected based on FACS as shown in Fig. 1(c). By combining movements of these control points, various facial expressions are realized with the face robot. In fact, a high correct recognition rate of 6 typical facial expressions that are possible to be recognized and express universally was achieved in previous research [11]. Fig. 2 shows examples of 6 typical facial expressions that the face robot can express.

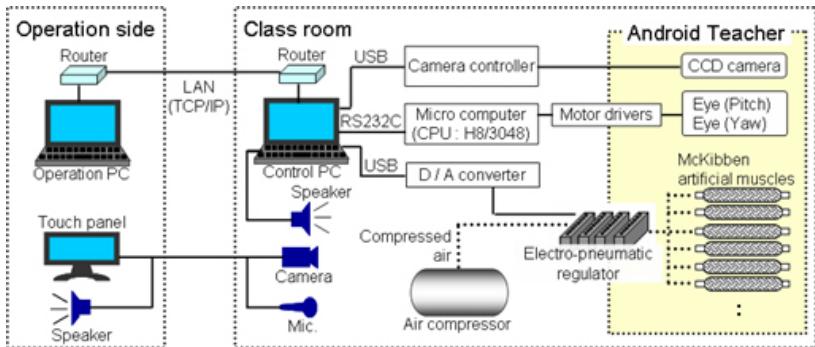
In the face robot, McKibben artificial muscles are used for actuating control points. Since it is small, lightweight and flexible, it can be distributed to curved surfaces of the skull like human muscles. And also, the facial skin is formed with soft urethane resin to realize the texture of human-like facial skin.

**Fig. 3.** Mechanical structure of SAYA**Fig. 4.** Actuator layout

## 2.2 Eyes' and Head's Movements

Fig. 1(c) and Fig. 3 show the internal structure of the face robot. The face robot has an 2 D.O.F. oculomotor mechanism that controls both pitch and yaw rotations of eyeballs by two DC motors. The two eyeballs move together since they are linked to each other. In addition, we mounted a CCD camera to the inside of the left-side eyeball. Thus, the eye direction is able to be controlled by recognizing the human skin color region from image of the CCD camera so that the face robot can pursue the visitor. In order to realize flexible head motions like a human cervical spine, we utilize a coil spring as the neck part for the head motion mechanism. Furthermore, the center of rotations for pitch rotation ("Pitch1") and yaw rotation were set in the base of the head. The forward and backward motions are realized by combining the head rotation ("Pitch1") and the neck bending ("Pitch2"). Movable positions and movable ranges were defined as shown in Fig. 3 by referring to anatomical knowledge. The roll-rotation, both pitch-rotations ("Pitch1" and "Pitch2") and the yaw-rotation are also actuated by McKibben artificial muscles as shown in Fig. 4.

We use electro-pneumatic regulators for controlling contractions of McKibben artificial muscles.



**Fig. 5.** System configuration

### 3 Remote Class Support System with the Android Robot SAYA

#### 3.1 System Structure

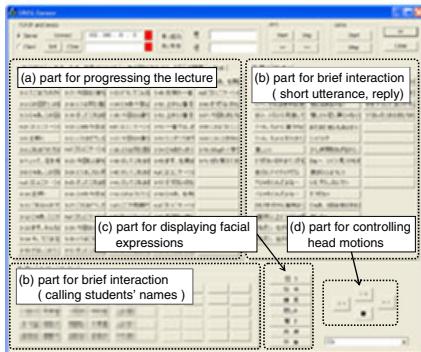
We developed the remote class support system with the android robot SAYA. In this system, we use SAYA as a role of a teacher. The detailed system configuration is shown in Fig. 5.

Captured images from CCD camera that is mounted in the eye ball of SAYA are transmitted to “Operation PC”. Thus, an operator can observe a classroom from the robot’s view point. Moreover, a video camera and a microphone are arranged at a rearward to obtain overview and sound of a classroom. Since the operator can move the view point of SAYA by controlling its eyes’ and head’s direction using these visual and sound information, SAYA is able to not only look around a classroom but also look at a certain child.

Commands from “Operation PC” are delivered to “Control PC” through the Internet. “Control PC” execute robot’s behaviors based on received commands.

#### 3.2 Interactive Behaviors

In this system, there are two operation modes that are “lecture mode” and “interaction mode”. In “lecture mode”, SAYA gives students explanations about slides that are displayed at the front of a classroom according to progress of the lecture. And also, around 95 utterances are prepared for this mode. On the other hand, in “interaction mode”, SAYA can perform interactive behaviors such as looking around, paying attention to a student, and talking to a student. SAYA can utter more than 80 sentences for talking to students such as “do your best”, “keep quiet”, “don’t look away” and so on. In addition, since SAYA has ability to express various facial expressions, it can express facial expressions according to its utterance. For example, when SAYA says “keep quiet” to a student, it expresses “anger”. We use women voices that

**Fig. 6.** Operation interface**Fig. 7.** Operation interface

were recorded beforehand as SAYA's voices. And also, names of students who participate in the class are registered. Therefore, SAYA is able to call student's name individually.

Since behaviors of SAYA consist of simple motions, an operator can execute these interactive behaviors easily. In addition, since SAYA is performed by tele-operation, it seems to apparently conduct the lecture and interact with students autonomously.

### 3.3 Operation Interface

Fig. 6 shows an example of the operation interface for an operator. As shown in the figure, there are many kinds of icons that are corresponding to robot's behaviors and utterances. It mainly consists of 4 parts : (a) part for progressing the lecture, (b) part for brief interaction, (c) part for displaying facial expressions, and (d) part for controlling head motions. An operator can easily control robot's behaviors by clicking these icons with a mouse or a touch panel.

Operating environment is shown in Fig. 7. Two kinds of images are displayed on observation monitor. One of them is image of CCD camera that is mounted on SAYA's eye ball. The other is image of video camera that is put on the rearward of the room. In addition, speakers output sound of the class room. The operator can monitor the class room, and he can control SAYA's behavior according to statuses of the lecture or students' reactions.

### 3.4 Contents of Science Class

The principle of leverage is adopted as a topic of science class for elementary school because it contains mathematical elements and experimental validations. In order to produce materials for the lecture, actual science textbooks of elementary schools were referred.

Table 1 shows the flow of the science class concerning leverage. First of all, in scene 1, SAYA gives self-introduction to students, and the class is opened. Main topics of the class consist three (scene 2 – 4). In scene 2, SAYA gives explanations

**Table 1.** The flow of science class about the leverage

Scene	Contents
1	Self introduction of SAYA and opening the class
2	Lecture on basic theory and mechanical advantages of leverage
3	Lecture on familiar examples of leverage in daily lives
4	Experiments for discovering the principle of leverage
5	Conclusion of the class

about the theory and mechanical advantages of leverage with some slides. In this scene, three important points of leverage are explained. That is, a pivot point, a point of effort, and a point of load. SAYA then shows familiar examples of leverages that are seen in our daily lives in scene 3. For example, scissors, bottle-opener, twitters, and so on. After that, in scene 4, SAYA lets students experiment to confirm the principle of leverage with an experimental kit, and they are able to experience the balancing theory of a lever through the experiment. In final scene 5, SAYA summarizes its talk and closes the class.

The class takes around thirty minutes. Moreover students can enjoy interacting with SAYA after the class for around twenty minutes.

## 4 Field Trial at Elementary Schools and Evaluation

### 4.1 Outline

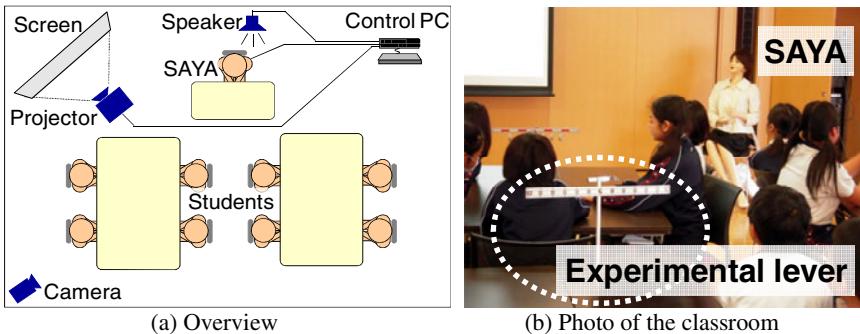
We carried out a field trial at an elementary school for verifying effectiveness of the system in educational fields. Our hypothesis is that elementary school students can attend a science class with interest by using the android robot SAYA as an educational system.

In the experiment, the class was conducted concerning the principle of leverage described in previous chapter, and students evaluated it.

### 4.2 Experimental Environment

Fig. 8 shows the experimental environment. We used a standard classroom of an elementary school for the experiment. As shown in Fig. 8 (a), android robot SAYA was put on the front of the class room. A video camera was placed on the back of the room to allow the operator to observe the room. A screen and a projector were also put on the front to show slides to students. In addition, there is an experimental kit of leverage on each desk every four people as shown in Fig. 8 (b).

We conducted experiments for three 5th grade classes. Each class has around twenty-five students, and 74 students participated in total.



**Fig. 8.** Experimental environment in each condition

#### 4.3 Experimental Procedure

First of all, an experimenter had students sit on the assigned position. He then explained the flow of the class and exited the room. After that, the class was begun soon. The class was conducted along the flow described in section 3.4. It was divided to three sections: scene 2, scene 3 and scene 4 in Table 1. During the class, the operator gave explanations about each topic of the lecture, and he also talked to students and asked some questions to students through SAYA.

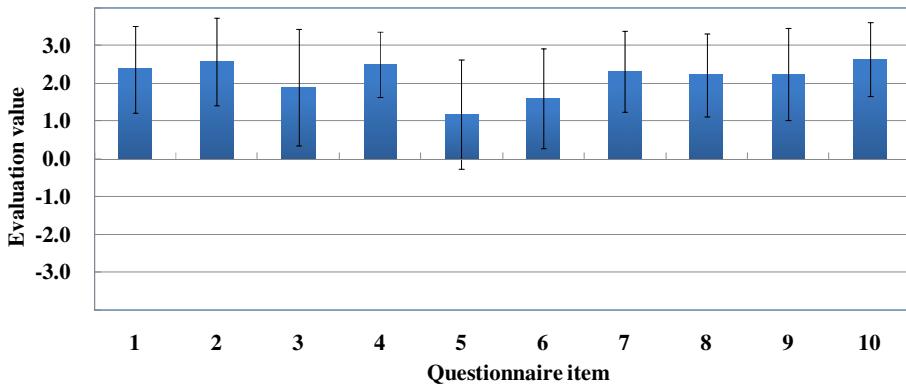
After the experiment, students answered questionnaires that were shown in the following section.

#### 4.4 Field Trials at Elementary Schools

#### 4.5 Evaluation

After the experiment, we asked students to evaluate the questionnaire that seemed to represent their impression in the class on a scale of -3 to 3, where 3 is most positive value. We prepared question items from the viewpoint of interest in the class, interest in the science, familiarity of SAYA as listed follows.

- Q.1 Were you able to concentrate on the class?
- Q.2 Did you have a good time in the class?
- Q.3 Did you get more nervous than usual classes? ( +3 : not getting nervous, -3 : getting nervous )
- Q.4 Were contents of the class easy to understand?
- Q.5 Were you easy to answer the teacher's questions?
- Q.6 Were you able to get familiar with the teacher?
- Q.7 Will you become to like science classes more?
- Q.8 Will you become to like learning more?
- Q.9 Will you be interested in science and technology?
- Q.10 Do you want to participate in this class again?



**Fig. 9.** Result of questionnaire

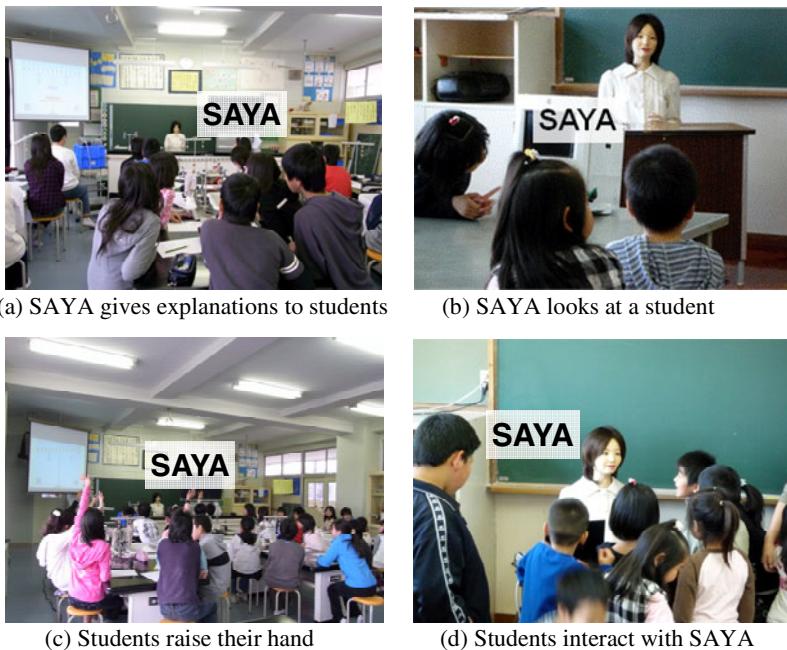
## 5 Experimental Results and Discussions

Fig. 9 shows the questionnaire result, and means and standard deviations of all question items are also shown. The value over zero means positive evaluation, while the value less than zero means negative. As a result, it is confirmed that positive evaluation was obtained in all questions. It can be said that students participated in SAYA's class interestedly and concentrated from these results, and our hypothesis is verified. However we also find that evaluation values of Q.5 and Q.6 relatively low. It seems that interactions between students and SAYA influenced these evaluations. For example, conversation interval and time delay of SAYA's reactions to the student's answer aren't smooth and natural. Therefore improvements of interaction manners and operability of SAYA for operator are required. Furthermore we will compare it to existing media such as video-conference system in order to investigate effectiveness of android robot SAYA as an interface of remote-class support systems. And also it will be evaluated with structured questionnaires and investigated from students' behaviors.

Fig. 10 shows photos of experiments in an elementary school. As shown in Fig. 10(a)(b), SAYA gives explanations about the principle of leverage according to the scenario, and students concentrated to hear SAYA's talk and pay attention to the screen. In Fig. 10(c), SAYA asks to students, and they raised their hand and tried to answer. As shown in Fig. 10(d), students enjoyed interacting with SAYA after the experiment. In all scenes, SAYA was controlled remotely by a hidden operator. According to these results, it was confirmed that students concentrated on SAYA's class and they reacted to SAYA's behaviors naturally as if they interacted with a real teacher.

We also asked for comments on the class to students as reference, and we obtained some comments as follows.

- Learning with the android robot SAYA was very amazing. I learned more pleasantly than usual classes.
- I was able to become to be interested in science and technology more than before.



**Fig. 10.** Field trial in elementary schools

- If the robot teacher was usually in a classroom, it would be very interesting and I would become to like science and technology.
- Explanations of SAYA were very clear. Therefore I was able to understand important topics of the class.
- I was able to concentrate on the class more than usual classes.
- I would like to participate in SAYA's class again.
- I was surprised at SAYA's abilities that called my name and looked at me.,
- I got nervous when SAYA called my name.

As results, we also confirmed that students attended the class actively and zestfully. And also, it is confirmed that android robot SAYA contributed to improve students' interest and motivation for science and technology.

## 6 Conclusion

This paper reported the remote-class support system with the android robot SAYA as a practical educational application. In this system, tele-operation system of SAYA was built. Moreover the principle of leverage was adopted as a topic of a science class, and field trials were conducted at elementary schools. In addition, effectiveness of the system in elementary school was investigated through field trials. As a result of the questionnaire, it is confirmed that students participated interestedly and concentrated on the class, and also using the android robot SAYA in educational fields contributes to improving students' interest and motivation for science and technology.

Our future work is to prepare various topics of the class and to investigate effectiveness of the developed system by conducting field trial for students who are at various ages and have various backgrounds. We will also improve interaction manners and operability of SAYA to realize smooth and natural interaction with students. In addition, since the greatest asset of android robots is that they give us a strong feeling of presence as if we communicate with real human even when we interact with robots, we will investigate its effectiveness by comparing other existing media such as only speaker and a video-conference system. Actually, Sakamoto et al. paid attention to an ability of android robots in realistic presence, and they used their android robot as a tele-communication medium [17]. Then, they confirmed that effectiveness of the android robot on conveying presence of a human who was in different place rather than existing media.

In addition, we will improve movements of SAYA's facial expressions, head motions and eye motions in order to give students highly realistic presence. Then we will add communicative behaviors such as body and arm movements for body languages. Operation environment and interface will be also improved to allow the operator to control SAYA's behaviors easily.

We believe that android-type tele-communication system also can be applied to other remote communication systems like a remote diagnosis.

## Acknowledgment

This research was partially supported by Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Young Scientists (Start-up), 21800058, 2009.

## References

1. Fujita, M.: AIBO towards the era of digital creatures. *International Journal of Robotics Research* 20(10), 781–794 (2001)
2. Wada, K., Shibata, T., Saito, T., Tanie, K.: Analysis of factors that bring mental effects to elderly people in robot assisted activety. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 2, pp. 1152–1157 (2002)
3. Hayashi, K., Sakamoto, D., Kanda, T., Shiomi, M., Koizumi, S., Ishiguro, H., Ogasawara, T., Hagita, N.: Humanoid robots as a passive-social medium - a field experiment at a train station. In: Proceedings of ACM/IEEE 2nd Annual Conference on Human-Robot Interaction, pp. 137–144 (2007)
4. Shiomi, M., Kanda, T., Ishiguro, H., Hagita, N.: Interactive Humanoid Robots for a Science Museum. *IEEE Intelligent Systems* 22(2), 25–32 (2007)
5. Watanabe, T., Okubo, M., Ogawa, H.: An Embodied Interaction Robots System Based on Speech. In: Proceedings of the 8th IEEE International Workshop on Robot and Human Communication (RO-MAN 1999), pp. 225–230 (1999)
6. Breazeal, C., Scassellati, B.: How to build robots that make friends and influence people. In: Proceedings of Intelligent Robots Systems (IROS 1999), pp. 858–863 (1999)
7. Kanda, T., Hirano, T., Eaton, D., Ishiguro, H.: Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Human Computer Interaction* 19(1-2), 61–84 (2004)

8. Kim, D., Kim, J.: Physical Learning Activities with a Teaching Assistant Robot in Elementary School Music Class. In: Proceedings of the 2009 Fifth International Joint Conference on INC, IMS and IDC, pp. 1406–1410 (2009)
9. Tanaka, F., Kimura, T.: The use of robots in early education: A scenario based on ethical consideration. In: Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication, pp. 558–560 (2009)
10. Kobayashi, H., Hara, F.: Study on face robot for active human interface-mechanisms of face robot and expression of 6 basic facial expressions. In: Proceedings of 2nd IEEE International Workshop on Robot and Human Communication, pp. 276–281 (1993)
11. Hashimoto, T., Hiramatsu, S., Tsuji, T., Kobayashi, H.: Development of the Face Robot SAYA for Rich Facial Expressions. In: Proceedings of SICE-ICASE International Joint Conference 2006, pp. 5423–5428 (2006)
12. Hashimoto, T., Kobayashi, H.: Development of the receptionist system with an anthropomorphism face. In: Proceedings of the 5th Asian Symposium on Applied Electromagnetics and Mechanics, pp. 190–196 (2005)
13. Ishiguro, H.: Android Science -Toward a new cross-interdisciplinary framework. In: Proceedings of International Symposium of Robotics Research (2005)
14. Oh, J., Hanson, D., Kim, W., Han, I., Kim, J., Park, I.: Design of Android Type Humanoid Robot Albert HUBO. In: Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1428–1433 (2006)
15. Mehrabian, A.: Communication without Words. *Psychology Today* 2(4), 53–55 (1968)
16. Ekman, P., Friesen, W.V.: The Facial Action Coding System. Consulting Psychologists Press (1978)
17. Sakamoto, D., Kanda, T., Ono, T., Ishiguro, H., Hagita, N.: Android as a telecommunication medium with a human-like presence. In: Proceedings of the ACM/IEEE International Conference on Human-robot Interaction, pp. 193–200 (2007)

# Essential Technology for Team Teaching Using Robot Partners

Akihiro Yorita and Naoyuki Kubota

Tokyo Metropolitan University, Graduate School of System Design,  
6-6, Asahigaoka, Hino, Tokyo, Japan  
yorita-akihiro@sd.tmu.ac.jp,  
kubota@tmu.ac.jp

**Abstract.** Recently, team teaching has been popular to realize high quality of education. This paper discusses the role of robots based on team teaching in remote education. Robot partners can help a teacher and make class more easily. There are three different aims of robot edutainment, i.e., Learning on Robots, Learning through Robots, and Learning with Robots. The last one is to apply human-friendly robots instead of personal computers for computer-assisted instruction. We propose a monitoring system of a class and tele-operating method of physical robot partners conducting the team teaching. The essence of the proposed method is in the user interface of remote education. The experimental results show the effectiveness of each system used in the remote education system.

**Keywords:** Robot Edutainment, Human-Robot Interactions, Remote Control and Monitoring, Robot Assisted Instruction.

## 1 Introduction

Recently, the number of students in a class is reduced to realize sophisticated education year by year. However, a problem is that the number of teachers is not enough for supporting all students. Computers and Internet are often used in the field of learning science. For example, Computer-supported collaborative learning (CSCL) is an emerging branch of the learning sciences concerned with studying how people can learn together with the help of computers [1]. The aim of the learning science is to discuss how to use computers to tie scientific understanding of knowledge obtained in schools with daily experience of students. A personal computer is useful to collect, access, edit, and store data, but agent-like communication capability is low in a personal computer. Therefore, a robot partner can be replaced with a personal computer.

On the other hand, team teaching has been popular to realize high quality of education. In general, two or more teachers possessing distinct sets of skills work together in a cooperative and coordinated fashion in the team teaching. A guest teacher can support a main teacher to play a specific role in lectures or experiments, but such a guest teacher might be asked to join the classes in different areas at the same time. Remote education through the Internet can be used to solve this problem, but an audio-visual system is not enough for communication and interaction with students in

the education. Because the system enables the teacher in the distance to have a class, but the style of the class doesn't change. For example, the English class needs assistant language teacher (ALT). Like this, it is necessary for another style of education. Therefore, we use robot partners for remote education in the team teaching in this paper.

First, we explain the background of robots used in the education. Next, we explain the robots used in remote education and the remote education system, and explain computational intelligence technologies used in the remote education. Furthermore, we discuss how to support students and teachers using robot partners. Finally, we discuss the future vision toward the realization of educational partner robotics.

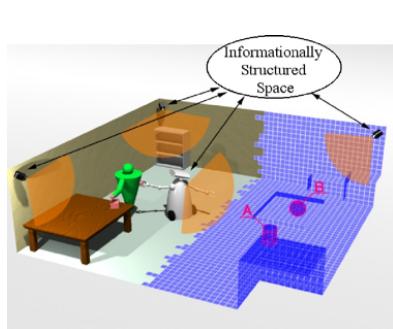
## 2 Robots for Remote Education

### 2.1 Roles of Robots in Education

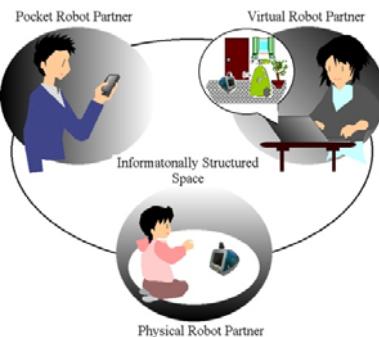
Various types of robots have been applied to the fields of education with entertainment (Edutainment) [6,7]. Basically, there are three different aims in robot edutainment. One is to develop knowledge and skill of students through the project-based learning by the development of robots (Learning on Robots). Students can learn basic knowledge on robotics itself by the development of a robot. The next one is to learn the interdisciplinary knowledge on mechanics, electronics, dynamics, biology, and informatics by using robots (Learning through Robots). The last is to apply human-friendly robots instead of personal computers for computer assisted instruction (Learning with Robots). A student learns (together) with a robot. In addition to this, such a robot can be used for supporting teachers by the teaching to students like a teaching assistant (robot TA) and the monitoring of the learning states of students [18]. An educational robot partner can teach something through interaction with students in daily situation. Furthermore, the robot can observe the state of friendship among students. This is very useful information for teachers, because it is very difficult for a teacher to extract such information from the daily communication with students. We showed the effectiveness of the learning with robots in the previous works [3-5]. A robot partner in educational fields cannot be the replacement of a human teacher, but the replacement of a personal computer. A student seldom shows physical reactions to a personal computer in the computer-assisted instruction (CAI), because the student is immersed into 2-dimensional world inside of the monitor. However, a student aggressively tries physical interactions to a robot, because the robot can express its intention through physical reactions. Of course, the robot also should play the role of a personal computer. Therefore, we propose the concept of robot-assisted instruction (RAI) to realize the style of education based on the learning with robots. A robot can be not only an assistant, but also a partner or collaborator in RAI.

### 2.2 Robot Partners for RAI

RAI requires the emerging synthesis of information technology (IT), network technology (NT), and robot technology (RT). Actually, the environment surrounding teachers, students, and robots should have a structured platform for gathering, storing, sharing, transforming, visualizing, and providing information. We have discussed the



**Fig. 1.** Informationally structured space

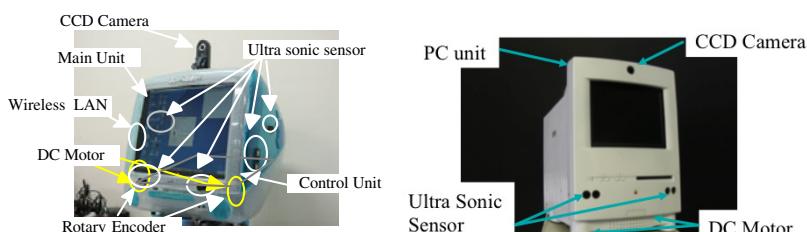


**Fig. 2.** Interaction with robot partners

concept of informationally structured space (Fig.1) to realize such a structured platform [16]. Furthermore, the structuralization of informationally structured space realizes the quick update and access of valuable and useful information for both people and robots. Furthermore, if the robot can share the environmental information with people, the communication with people might become very smooth and natural.

We can have three different types of robot partners from the interactive point of view in the informationally structured space (Fig.2). One is a physical robot partner. We can interact with the physical robot partner by using multi-modal communication like a human. The next one is a pocket robot partner. The pocket robot partner has no mobile mechanism, but we can easily bring it everywhere and can interact with the robot partner by touch and physical interface. The last one is a virtual robot partner. The virtual robot partner is in the virtual space in the computer, but we can interact with it through the virtual person or robot in the virtual space.

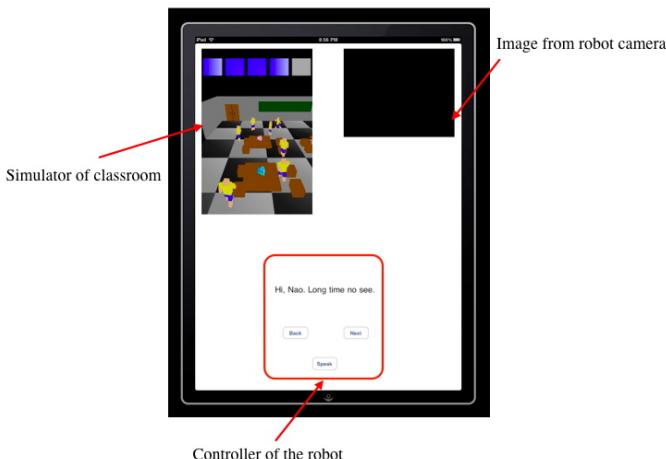
We developed PC-type of physical robot partners called MOBiMac (Fig. 3) in order to realize human-friendly communication and interaction. This robot has two CPUs and many sensors such as CCD camera, microphone, and ultrasonic sensors. Furthermore, the information perceived by a robot is shared with other robot by the wireless communication. Therefore, the robots can easily perform formation behaviors. We have applied steady-state genetic algorithm (SSGA), spiking neural networks (SNN), self-organizing map (SOM), and others for human detection, motion extraction, gesture recognition, and shape recognition based on image processing [2,8-10].



**Fig. 3.** Human-friendly Partner robots; MOBiMac

Furthermore, the robot can learn the relationship between the numerical information as a result of image processing and the symbolic information as a result of voice recognition [11]. MOBiMac can be also used as a standard personal computer.

We have used apple iPhone, iPod touch and iPad as pocket robot partners, because we can easily use the touch interface and accelerometer in the program development. In this paper, we use iPad as a remote controller of the robot and a display for interacting with students. Figure 4 shows the overview of interfaces used in iPad. A simulator mode is shown in the upper left. This display is used for selecting robot and monitoring students in a class. The information of students is extracted from the environmental camera in the classroom. A camera image sent from a physical robot partner is shown in the upper right. The operator can observe a state of the student, and can teach the student through the physical robot partner. Buttons for the control of physical robot partners are shown in the bottom of the display. The operator can choose a sentence of utterance from the list, and the robot speaks the chosen sentence.



**Fig. 4.** iPad as a remote controller and display

### 2.3 Team Teaching Using Physical Robot Partners

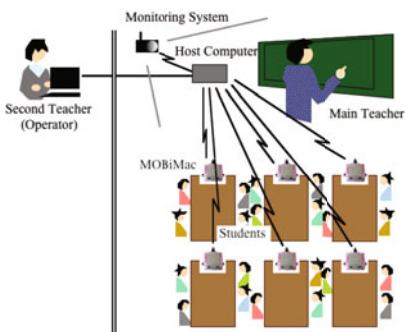
We assume team teaching by a main teacher and physical robot partners in a classroom, and a second teacher in a remote area. The second teacher controls the physical robot partner to teach a student individually. In the case of more than two robots, the second teacher selects a physical robot partner in the iPad simulator, and operates the robot manually by choosing the operating commands. The other physical robot partners autonomously speak according to the operating command by the second teacher. Figure 5 shows the total architecture of the remote control system for robots based on team teaching in the remote education. Basically, this system is composed of two rooms of a classroom and operation room.

There are a host computer, robots, and monitoring system in the classroom. The host computer is connected with a monitoring system of the classroom and robots by wireless communication, and has the database on the personal information of

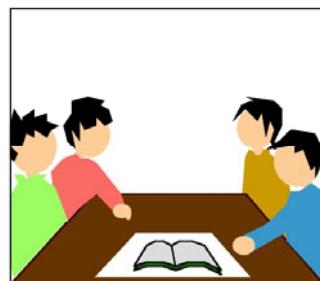
students, and educational environment. Furthermore, the host computer can send the state of students obtained from the robots to the teacher in the remote operation room. The camera equipped in the monitoring system takes an overall image of the classroom, while the camera equipped with the robots takes a local image on the table and the state of students.

There are a remote control computer, a main monitor, and monitoring system in the operation room. The teacher gives a talk to the student. Here, if the teacher calls a student A, then the operator shows the state of the student A to the main monitor. Furthermore, the operator sends the basic control commands to the robots in the classroom, and performs the monitoring of students in the classroom (Fig.6, the view from MOBiMac). The robots are controlled by semi-autonomous control.

We use English for elementary school students as a content of lecture. Furthermore, since the opportunity for the students to interact in the class is few, the second teacher can help students to study English conversation. The effect of using robots can reduce the degree of nervousness compared with foreign teachers. MOBiMac has a screen to display texts, images, and movies without using a blackboard in the classroom.



**Fig. 5.** A total architecture of remote



**Fig. 6.** The view from education system MOBiMac on the table

### 3 Monitoring System for Remote Education

#### 3.1 Evolutionary Robot Vision for Tracking Students

A monitoring system of students is very important to support the main and second teachers. It can confirm the learning state of students. Therefore, we use a people tracking system based on a local genetic algorithm based on clustering (LGAC). This method is a distributed search method based on local hill-climbing in a subset of clustered individuals. Furthermore, we use fuzzy evaluation for people tracking [15,17].

Basically, each individual is composed of diploid, i.e., the self-best solution and candidate solution. If the fitness value of the candidate solution is larger than that of the self-best solution, the candidate solution is replaced with the self-best solution. In this way, each individual performs the elitist selection. We use the term of personal best or self-best that inspired from particle swarm optimization (PSO) invented by

Eberhart and Kennedy [14]. Furthermore, we use the elitist crossover in the following updating rule;

$$\begin{aligned} g_{i,j} \leftarrow & g_{i,j} + \alpha_1 r_1 (g_{i,j}^S - g_{i,j}) \\ & + \alpha_2 r_2 (g_{L,j}^S - g_{i,j}) + \alpha_N N(0,1) \end{aligned} \quad (1)$$

where  $g_{i,j}^S$  is the self-best solution;  $g_{L,j}^S$  is the locally best solution in a cluster;  $r_1$  and  $r_2$ , are uniform random value between 0 and 1.0;  $N(0,1)$  is a normal random value with average of 0 and 1.0, and  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_N$  are coefficients. Furthermore, we can use adaptive mutation as follows;

$$\alpha_N = \beta_1 \cdot \frac{f_{\max} - f_i}{f_{\max} - f_{\min}} + \beta_2 \quad (2)$$

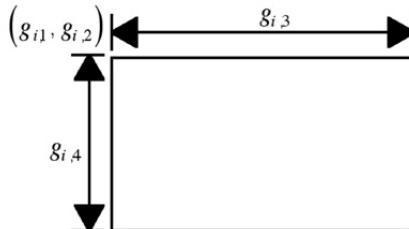
where  $f_{\max}$  and  $f_{\min}$  are the maximal value and minimal value of fitness values in a local cluster or the population; and  $\beta_1$  and  $\beta_2$  are coefficient and offset, respectively. In the adaptive mutation, the variance of the normal random number is relatively changed according to the fitness values of the population in case of maximization problems. The proposed method is similar to PSO, but in this paper, we explicitly use a mutation factor in order to trace local minima in a dynamic environment.

The human face candidate positions based on human skin and hair colors are extracted by LGAC with template matching. The fitness value is calculated by the following equation,

$$f_{Pixel} = C_{Skin} + C_{Hair} + \eta_1 \cdot C_{Skin} \cdot C_{Hair} - \eta_2 \cdot C_{Other} \quad (3)$$

where  $C_{Skin}$ ,  $C_{Hair}$  and  $C_{Other}$  indicate the numbers of pixels of the colors corresponding to human skin, human hair, and other colors, respectively;  $\eta_1$  and  $\eta_2$  are the coefficients ( $\eta_1, \eta_2 > 0$ ).

The robot must recognize a human face from complex background speedily. The human face candidate positions based on the colors of facial landmarks are extracted by LGAC with template matching. Figure 7 shows a candidate solution of a template used for detecting a human face. A template is composed of numerical parameters of  $g_{i,1}$ ,  $g_{i,2}$ ,  $g_{i,3}$ , and  $g_{i,4}$ . The number of individuals is  $G$ . The iteration of LGAC is repeated until the termination condition is satisfied.



**Fig. 7.** A template used for human detection in LGAC

### 3.2 Hierarchical Growing Neural Gas Algorithm for Monitoring Lecture Rooms

When we have a class, we need to observe the state of students. We proposed Hierarchical Growing Neural Gas (HGNG) in order to extract a human and objects from the background and to extract the spatial relationship among them.

We explain the learning algorithm of GNG [12,13]. The notation used in GNG is shown as follows;

$w_i$  : the  $n$ th dimensional vector of a node ( $w_i \in \mathbf{R}^n$ )

$A$  : A set of nodes

$N_i$  : A set of nodes connected to the  $i$ th node

$c$  : A set of edges

$a_{i,j}$  : Age of the edge between  $i$ th and  $j$ th node

Step 0. Generate two units at random position,  $w_{c1}, w_{c2}$  in  $\mathbf{R}^n$ . Initialize the connection set.

Step 1. Generate at random an input data  $v$  according to  $p(v)$  which is probability density function of data  $v$ .

Step 2. Select the nearest unit (winner)  $s_1$  and the second-nearest unit  $s_2$  by

$$s_1 = \arg \min_{i \in A} \|v - w_i\| \quad (4)$$

$$s_2 = \arg \min_{i \in A \setminus \{s_1\}} \|v - w_i\| \quad (5)$$

Step 3. If a connection between  $s_1$  and  $s_2$  does not exist already, create the connection. Set the age of the connection between  $s_1$  and  $s_2$  to zero;

$$a_{s_1, s_2} = 0 \quad (6)$$

Step 4. Add the squared distance between the input data and the winner to a local error variable;

$$E_{s_1} \leftarrow E_{s_1} + \|v - w_{s_1}\|^2 \quad (7)$$

Step 5. Update the reference vectors of the winner and its direct topological neighbors by the learning rate  $\eta_1$  and  $\eta_2$ , respectively, of the total distance to the input data.

$$w_{s_1} \leftarrow w_{s_1} + \eta_1 \cdot (v - w_{s_1}), \quad (8)$$

$$w_j \leftarrow w_j + \eta_2 \cdot (v - w_j) \text{ if } c_{s_1, j} = 1 \quad (9)$$

Step 6. Increment the age of all edges emanating from  $s_1$ .

$$a_{s_1, j} \leftarrow a_{s_1, j} + 1 \text{ if } c_{s_1, j} = 1 \quad (10)$$

Step 7. Remove edges with an age larger than  $a_{\max}$ . If this results in units having no more emanating edges, remove those units as well.

Step 8. If the number of input data generated so far is an integer multiple of a parameter  $\lambda$ , insert a new unit as follows.

- Select the unit  $q$  with the maximum accumulated error.

$$q = \arg \max_{i \in A} E_i \quad (11)$$

- Select the unit  $f$  with the maximum accumulated error among the neighbors of  $q$ .

- Add a new unit  $r$  to the network and interpolate its reference vector from  $q$  and  $f$ .

$$w_r = 0.5 \cdot (w_q + w_f) \quad (12)$$

- Insert edges connecting the new unit  $r$  with units  $q$  and  $f$ , and remove the original edge between  $q$  and  $f$ .

- Decrease the error variables of  $q$  and  $f$  by a fraction  $\alpha$

$$E_q \leftarrow E_q - \alpha E_q \quad (13)$$

$$E_f \leftarrow E_f - \alpha E_f \quad (14)$$

- Interpolate the error variable of  $r$  from  $q$  and  $f$

$$E_r = 0.5 \cdot (E_q + E_f) \quad (15)$$

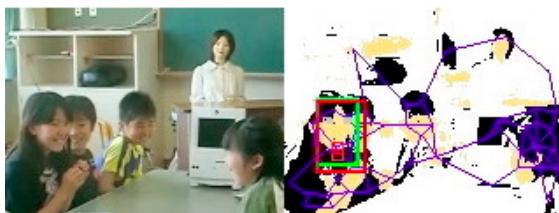
Step 9. Decrease the error variables of all units

$$E_i \leftarrow E_i - \beta E_i \quad (\forall i \in A) \quad (16)$$

Step 10. Continue with step 2 if a stopping criterion (e.g., net size or some performance measure) is not yet fulfilled.

Next, we explain the Hierarchical GNG (HGNG). HGNG consists of  $l$  layers of GNG. The lowest layer of GNG directly uses the input data and learns the structure of data distribution. The  $(i+1)$ th layer of GNG uses the reference vector of the selected node in the  $i$ th layer of GNG. Therefore, the higher layer of GNG learns the relationship among nodes in the lower layer of GNG.

Figure 8 shows the result of face detection with GNG. Original image is in the left and face detection is in the right. The snapshot can detect a face correctly.



**Fig. 8.** The result of face detection with GNG

## 4 Experimental Results

This section shows experimental results of the proposed method for the remote education. First, we show a remote communication between a teacher and a student through robot partners in Fig 9. A physical robot partner and pocket robot partner are used in a classroom and a remote office, respectively. Figure 9 (a) shows a person tries to show the notebook to the teacher through the camera equipped with the physical robot partner. The image taken by the robot is sent to the screen of iPad used as the pocket robot partner, and the teacher can confirm it. Figure 9 (b) and (c) show the teacher can watch the person and check several screenshot images of the notebook sent from the physical robot partner. The physical robot partner can move toward the person, and control the position of pan-tilt camera according to the operation on the screen of iPad by the teacher.

It is possible for the second teacher to identify a human face and the picture through the camera equipped with the robot. In the case of showing object shown in Fig.10, we can recognize the object as an orange. In this way, the performance of transfer speed and resolution of images enough to recognize words, pictures, and objects in human-robot interaction is required. as the specification of a remote education system to conduct an English class,



(a) The person shows a notebook to the robot



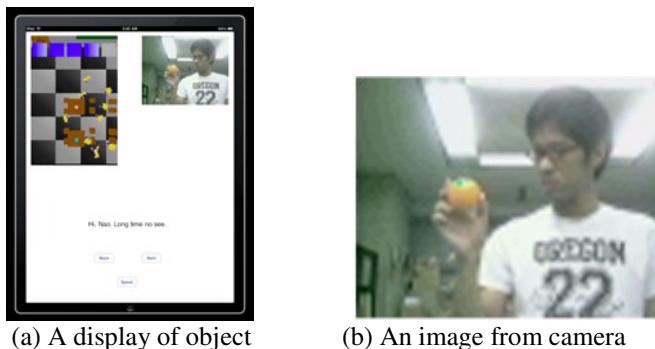
(b) A display of the person



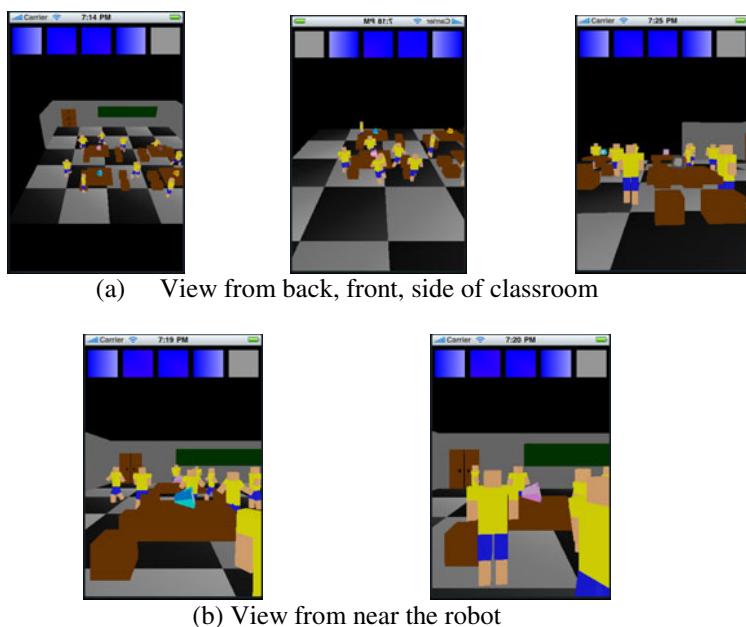
(c) A display of the notebook

**Fig. 9.** Transfer image to iPad

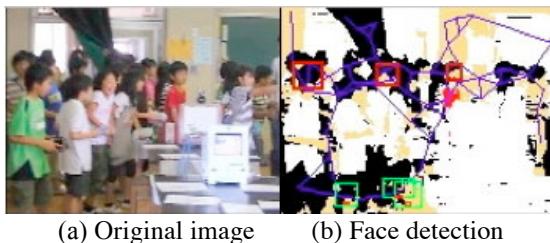
Next, we show a simulation result of visualizing the overview of a classroom. In the simulator of iPhone, the abstract state of students is monitored from the different viewpoints. This system can display the state of a desk by expanding according to a pinch operation using the touch interface on the screen, and can change the viewpoint

**Fig. 10.** In the case of showing object

by a scroll motion using the touch interface on the edge of screen. Basically, physical robot partners respond to students and move semi-autonomously, because the teacher cannot receive all of questions from students in the class at the same time. According to questions from students, the teacher can choose a target student by touching the robot belonging to a group of the target student. Figure 11 shows snapshots of screen in the selection process of the robot. The teacher selected the target robot by expanding the screen size and by changing the viewpoint of the classroom.

**Fig. 11.** The simulator in iPhone

Finally, Fig.12 shows image processing results using an image from environmental camera where a red rectangle shows tracking person. Accordingly, the number of rectangles in the simulator is reflected the number of recognized students in a class. The experimental results show that the proposed method can detect faces of students correctly.



**Fig. 12.** The result of face detection with GNG from environmental camera

## 5 Summary

In this paper, we discussed the applicability of robots in the remote education for team teaching. First, we explained the robot partners used in remote education. Next, we discussed how to interact and communicate with students based on the team teaching composed of the main and second teachers and several robot partners in a remote education. We proposed a monitoring system and tele-operating method of physical robot partners. The essence of the proposed method is in the user interface of remote education. A physical robot partner is human-friendly and interesting for students, while the human interface of iPad is easy and intuitive for teachers. These features are realized by the platform of informationally structured space. The experimental results show the effectiveness of each component used in remote education system.

As future work, we will integrate the components used in the remote education system, and develop the monitoring system of the learning state of students. Furthermore, we intend to improve the performance of people tracking in the monitoring system for watching the overview of a classroom.

## Acknowledgments

A part of this research and development was done in cooperation with Prof. Hiroshi Kobayashi and Dr. Takuya Hashimoto of Tokyo University of Science.

## References

1. Stahl, G., Koschmann, T., Suthers, D.: Computer-supported collaborative learning: An historical perspective. In: Sawyer, R.K. (ed.) Cambridge handbook of the learning sciences. Cambridge University Press, Cambridge (2006)
2. Kubota, N., Nishida, K.: Cooperative Perceptual Systems for Partner Robots Based on Sensor Network. International Journal of Computer Science and Network Security (IJCSNS) 6(11), 19–28 (2006)

3. Kubota, N., Ozawa, S.: Tele-operated Robots for Monitoring Based on Sensor Networks. In: Proc. of SICE Annual Conference 2008, pp. 3355–3360 (2008)
4. Kubota, N., Tomioka, Y., Ozawa, S.: Intelligent Systems for Robot Edutainment. In: Proc. of 4th International Symposium on Autonomous Minirobots for Research and Edutainment, pp. 37–46 (2007)
5. <http://www.robotis.com/zbxe/main>
6. Veltman, K.H.: Edutainment, Technotainment and Culture, Civita Annual Report (2003)
7. Ito, M.: Engineering play: Children's software and the cultural politics of edutainment. Discourse 27(2), 139–160 (2004)
8. Kubota, N.: Visual Perception and Reproduction for Imitative Learning of A Partner Robot. WSEAS Transaction on Signal Processing 2(5), 726–731 (2006)
9. Kubota, N.: Computational Intelligence for Structured Learning of A Partner Robot Based on Imitation. Information Science (171), 403–429 (2005)
10. Kubota, N., Nojima, Y., Kojima, F., Fukuda, T.: Multiple Fuzzy State-Value Functions for Human Evaluation through Interactive Trajectory Planning of a Partner Robot. Soft Computing 10(10), 891–901 (2006)
11. Kubota, N., Yorita, A.: Structured Learning for Partner Robots based on Natural Communication. In: Proc (CD-ROM) of 2008 IEEE Conference on Soft Computing in Industrial Applications (SMCia), pp. 303–308 (2008)
12. Fritzke, B.: A growing neural gas network learns topologies. In: Tesauro, G., Touretzky, D.S., Leen, T.K. (eds.) Advances in Neural Information Processing Systems, vol. 7, pp. 625–632. MIT Press, Cambridge (1995)
13. Fritzke, B.: Growing cell structures - a self-organizing network for unsupervised and supervised learning. Technical Report ICSTR-93-026, International Computer Science Institute, Berkeley (May 1993)
14. Kennedy, J., Eberhart, R.: Particle Swarm Optimization. In: Proc. IEEE Int. Conf. Neural Netw., Perth, Australia, pp. 1942–1945 (1995)
15. Yorita, A., Kubota, N.: Fuzzy-Based Evolutionary Robot Vision for People Tracking. International Journal of Intelligent Computing in Medical Sciences and Image Processing (IC-MED) 3(2), 119–129 (2009)
16. Satomi, M., Masuta, H., Kubota, N.: Hierarchical Growing Neural Gas for Information Structured Space. In: Proc. of IEEE Symposium Series on Computational Intelligence 2009, pp. 54–59 (2009)
17. Yorita, A., Kubota, N.: Multi-stage fuzzy evaluation in evolutionary robot vision for face detection. Evolutionary Intelligence 3(2), 67–78 (2010)
18. Yorita, A., Hashimoto, T., Kobayashi, H., Kubota, N.: Remote Education based on Robot Edutainment in Elementary School. International Journal of Social Robotics (2010) (submitted)

# The Style of Information Service by Robot Partners

Aiko Yaguchi and Naoyuki Kubota

Graduate School of System Design, Tokyo Metropolitan University,  
6-6 Asahigaoka, Hino-shi, Tokyo, 191-0065, Japan  
[yaguchi-aiko@sd.tmu.ac.jp](mailto:yaguchi-aiko@sd.tmu.ac.jp), [kubota@tmu.ac.jp](mailto:kubota@tmu.ac.jp)

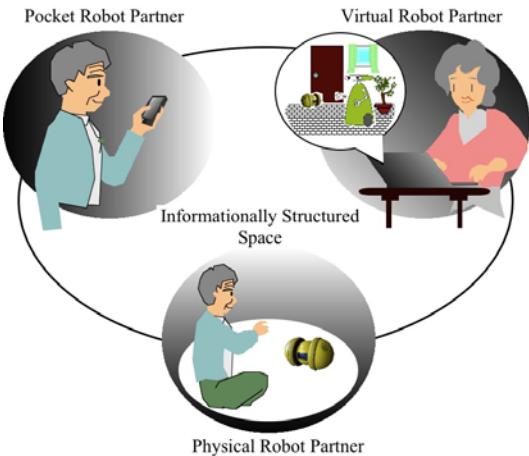
**Abstract.** This paper discusses the style of information service by robot partners. In order to perform information service suitable to each person, the robot partner should extract the preference and daily behavior patterns of the person. Therefore, we propose a learning method of the relationship between human interaction and its corresponding behavior of the robot partner. In this paper, we use a robot music player; miuro, and we focus on the music selection for providing the comfortable sound field for the person. The experimental results show that the proposed method can learn the relationship between human interaction and its corresponding behavior, and that the proposed method can provide the person with the preferable song as the comfortable sound field. Furthermore, we show the seamless sharing of information between different types of robot partners.

**Keywords:** Robot Partners, Information Service, Sound Field, Behavioral Learning, Multi-Objective Behavior Coordination.

## 1 Introduction

The attention to the aging society and declining birthrate has been heightened in advanced countries including Japan and Korean. The communication and interaction with others is so important for healthy long life of aged people. Therefore, robot partners will play the important role in keeping social connection among people each other in the aging society. Robot partners are excellent as human interface to exchange information among people. Moreover, robot partners can share various types of information with other robots and external information systems through a network. The necessary information of individual person is different owing to daily schedule, health condition, and personal preference. To maintain social connection, it is necessary for robot partners to do such personalized information services.

A robot partner will be a concierge in the information service. We will have three different types of robot partners from the interactive point of view (Fig.1). One is a physical robot partner. We can interact with the physical robot partner by using multimodal communication like a human. The next one is a pocket robot partner. The pocket robot partner has no mobile mechanism, but we can easily bring it everywhere and can interact with the robot partner by touch and physical interface. The last one is a virtual robot partner. The virtual robot partner is in the virtual space in the computer, but we can interact with it through the virtual person or robot in the virtual space. The



**Fig. 1.** Interaction with robot partners

interaction style of these three robots is different, but they share the same personal database and interaction logs, and can interact with the person based on the same interaction rules independent from the style of interfaces. The person can accept the information service ubiquitously from a change of the style of robot partners according to the situation.

People have to communicate their intentions with robots to accept information services through interaction. In general, the basic behaviors for robots are designed beforehand, but the learning and adaptation methods for basic behaviors such as collision avoidance and target tracing have been proposed until now [1-4]. Furthermore, various types of human-friendly robots have been developed, but the aim of most previous works is in the interaction between a robot and person, not the behavioral learning. Robovie was developed to realize natural communication, and the architecture of behavior modules was proposed to generate episode chains [5-7]. In the previous studies of multiple interactive systems to support a person dependent on the situation, ITACO system has been proposed [8]. ITACO system was designed to build emotional relationships between human and robot. However, the detail of support and personal adaptation is not explained. Kidd et al. proposed to use mental commitment robot; Paro to promote social interaction of aged people [9], but the detail of use in a home is not explained. In general, it is difficult to learn the human-like interactive behaviors through the interaction and communication with people. Therefore, we propose a learning method of the relationship between human interaction and its corresponding behaviors of the robot partner according to the shared situation.

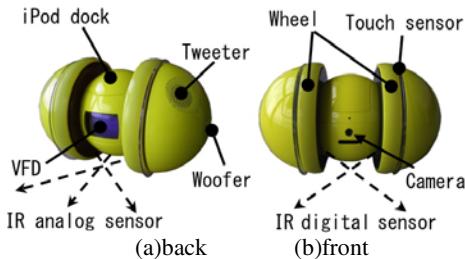
The style of multi-modal communication is diverse and specialized according to an individual person, and the information service can be also done by various methods such as visual, audio, and text information. Therefore, the style of information service should be specialized according to an individual person. In this paper, we propose a learning method to provide information service suitable to an individual person. We can obtain the user information such as preferences and daily behavioral patterns by the relationship learning. Furthermore, we discuss seamless sharing of information among different types of robot partners.

In section 2, we discuss information service by robot partners. In section 3, we describe behavior control for a robot partner miuro. In section 4, we show experimental results of communication learning with a pocket robot partner and a physical robot partner, and in section 5, the paper is summarized, and future work is mentioned.

## 2 Information Service by Robot Partners

### 2.1 A Robot Partner miuro

We use a physical robot partner called miuro [10]. This physical robot partner is developed as a robot music player by ZMP INC. in Japan, and can act in autonomous motion with the big size of wheels attached with both sides (Fig.2). The robot has many sensors such as camera, 10 infrared distance sensors, touch sensors, and accelerometer. Furthermore, miuro can be controlled by infrared remote controller. The user can play music contents registered beforehand in miuro by remote controller, and iPod can be connected and built in the body of miuro. The robot can play music contents in a PC server and programs of Net Radio connected through wireless LAN (802.11b/g) by its high quality speakers.



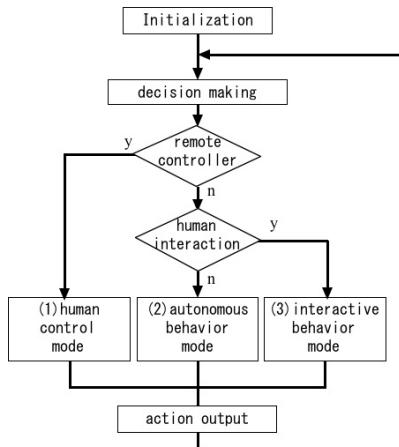
**Fig. 2.** Robot partner miuro

### 2.2 Control Architecture of miuro

Basically, miuro has three modes of human control, autonomous behavior and interactive behavior. These decision modes are changed according to the remote controller or human interaction. In the human control mode, the user can directly control miuro, choose the predefined behaviors, and play the music contents by the remote controller. The autonomous behavior mode is composed of random walk, target tracing, information gathering, and map building, and the action output of miuro is decided by the behavior coordination according to the internal state and local sensing information from external environments. The interactive behavior mode is composed of music play, human tracking, and amusing play.

Figure 3 shows the detail of decision making of miuro. After the sensing, if the remote controller is activated, the behavior mode is transited to the human control mode (Fig.3 (1)). In this mode, miuro is directly controlled and used as a music player. On the other hand, if the human interaction is extracted or recognized, the behavior mode is transited to the interactive behavior mode (Fig.3 (3)). In this mode, miuro

also aggressively tries to contact with the person, to collect and extract personal information such as preference and custom, and to provide various types of information suitable to the situation.



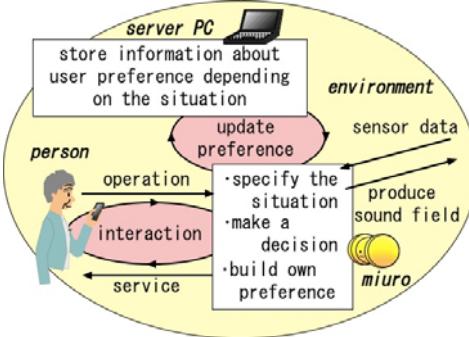
**Fig. 3.** The total procedure of decision making of miuro

### 2.3 Information Service by miuro

Basically, information is defined as meaningful or valuable data from the viewpoint of information science. On the other hand, sensory information includes visual information, audio information, tactile information, and olfactory information, from psychology [11]. This means we can perceive the meaning or value from such sensory information. In this study, we try to realize information service by music, sound and voice. Especially, we discuss the method of providing the comfortable mood for user by the sound field using music suitable to the situation.

Figure 4 shows the role of miuro in information service. The most important advantage of miuro is in the high quality of music play with autonomous sensing and moving capability. Therefore, miuro can provide a person with the sound field according to the situation extracted as a result of sensing. The second advantage of miuro is in wireless communication with PC server and Internet. The extracted information on preference and custom is stored in the PC server and referred from the PC server. Furthermore, miuro can download music contents from web sites according to the extracted preference of the person, and play them. However, the downloaded music contents might be different from the preference, but this can be a new finding of music for the person. This means the person does not perform the search of music contents, but miuro can be a concierge of information service.

The roles of music plays can be divided into music appreciation and background music. A person directly enjoys listening music in the music appreciation, while the person enjoys other tasks with listening music in the mode of background music. In the latter, the background music plays the important role in entraining the person into the target task. Based on the portability and mobility of miuro, we focus on the music selection as the background music in the following.



**Fig. 4.** The roles of miuro in information service

### 3 Computational Intelligence for Behavior Control

#### 3.1 Autonomous Behavior Mode

Various intelligent methods for autonomous behavior control have been proposed such as production rules, Bayesian networks, neural networks, fuzzy inference systems, and classifier systems [12-16]. We have applied fuzzy inference systems to represent behavior rules of robots, because the behavioral rules can be designed easily and intuitively by human behavioral representations. We apply simplified fuzzy inference to the motion control of miuro. In general, a fuzzy if-then rule is described as follows,

If  $x_1$  is  $A_{i,1}$  and ... and  $x_M$  is  $A_{i,M}$   
 Then  $y_1$  is  $w_{i,1}$  and ... and  $y_N$  is  $w_{i,N}$

where  $A_{i,j}$  and  $w_{i,k}$  are the Gaussian membership function for the  $j$ th input and the singleton for the  $k$ th output of the  $i$ th rule;  $M$  and  $N$  are the numbers of inputs and outputs, respectively. Fuzzy inference is performed by,

$$\mu_{A_{i,j}}(x_j) = \exp\left(-\frac{(x_j - a_{i,j})^2}{b_{i,j}^2}\right) \quad (1)$$

$$\mu_i = \prod_{j=1}^M \mu_{A_{i,j}}(x_j) \quad (2)$$

$$y_k = \frac{\sum_{i=1}^R \mu_i w_{i,k}}{\sum_{i=1}^R \mu_i} \quad (3)$$

where  $a_{i,j}$  and  $b_{i,j}$  are the central value and the width of the membership function  $A_{i,j}$ ;  $R$  is the number of rules. Moreover, the  $k$ th output  $y_k$  is motor output. In general, a mobile robot has a set of behaviors for achieving various objectives, and must integrate these behaviors according to the environmental conditions. Therefore, we proposed the method for multi-objective behavior coordination [17,18]. The multi-objective behavior coordination can integrate outputs of several behaviors according to the time-series of perceptual information. A behavior weight is assigned to each behavior. miuro can achieve a smooth behavior by dynamic update of the behavior weights along perceptual information. According to the behavior weights, the final output is calculated by,

$$y_k = \frac{\sum_{j=1}^K wgt_j(t) \cdot y_{j,k}}{\sum_{j=1}^K wgt_j(t)} \quad (4)$$

where  $K$  is the number of behaviors;  $wgt_j(t)$  is a behavior weight of the  $j$ th behavior over the discrete time step  $t$ , and output  $y_{j,k}$  is the calculation result from simplified fuzzy inference used for the  $j$ th behavior. The update amount of each behavior weight is calculated as follows,

$$\begin{bmatrix} \Delta wgt_1 \\ \Delta wgt_2 \\ \vdots \\ \Delta wgt_K \end{bmatrix} = \begin{bmatrix} dw_{1,1} & dw_{1,2} & \cdots & dw_{1,L} \\ dw_{2,1} & & & \\ \vdots & & \ddots & \vdots \\ dw_{K,1} & & \cdots & dw_{K,L} \end{bmatrix} \begin{bmatrix} ci_1 \\ ci_2 \\ \vdots \\ ci_L \end{bmatrix} \quad (5)$$

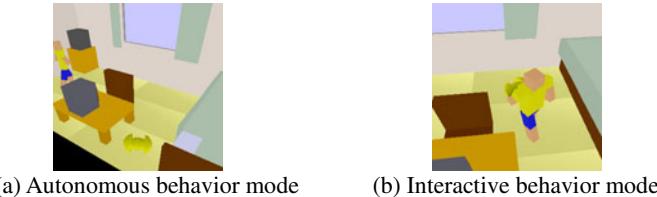
where  $ci_i$  is the parameter on the perceptual information;  $L$  is the number of perceptual inputs.

First, the multi-objective behavior coordination can integrate collision avoidance and target tracing behaviors [18-20]. The inputs to the fuzzy controller for collision avoidance are the measured distance to the obstacle by infrared distance sensors equipped in front of the robot ( $M_c=3$ ). The inputs to the fuzzy controller for target tracing are the estimated distance to the target point and the relative angle to the target point from the moving direction ( $M_t=2$ ). Outputs are a moving velocity and a steering angle. Final outputs are calculated by the weighted average of outputs form these two behaviors based on the multi-objective behavior coordination according to the degree of danger. The degree of danger is calculated utilizing the minimum value of distances to obstacles by measurement as follows,

$$d = \left( S - \min_j x_j \right) / S \quad (6)$$

where  $S$  is the sensing range of infrared distance sensors.

Figure 5 shows the snapshot of autonomous behavior mode and interactive behavior mode in the pocket robot partner. The robot takes random walk in the autonomous behavior mode in Fig.5 (a), and the robot takes the human tracking behavior in the interactive behavior mode in Fig.5 (b).



**Fig. 5.** Behavior modes in the pocket robot partner

### 3.2 Interactive Behavior Mode

The interactive behavior mode is composed of human tracking, music play, and amusing play. As mentioned before, miuro is designed as a music player, and this paper focuses on the method of providing the comfortable mood for a person by the sound field using music suitable to the situation. The individual preferable sound field corresponding to situations such as cooking, reading, studying, and sleeping is different among people. The robot music player, miuro, should provide the individual person with the comfortable mood, but it is difficult to design the relationship between a situation and its corresponding sound field. Furthermore, such relationship should be extracted and learned through communication and interaction with the person. In this paper, we assume a specific situation is corresponding to a specific human interaction.

We propose a learning method of the relationship between the human interaction and its corresponding music play through the interaction with the person. The robot reacts to a human interaction. If the person prefers the reaction of robot as the result of human interaction, the person tries to repeat the same interaction with the robot because of expecting the same output. Otherwise, the person tries to change to other interactions. Therefore, the learning results of the robot can reflect human preferences. Based on this assumption, we propose a learning method based on linear reward-inaction algorithm to obtain the relationship between the human interaction and its corresponding reaction of the robot. The selection strength  $s_{i,x}$  of the  $i$ th reaction of a robot with a human interaction  $x(t)$  is updated as follows,

$$s_{i,x}(t) = \begin{cases} s_{i,x}(t) + R & \text{if } x(t) = x(t-1) \\ s_{i,x}(t) & \text{otherwise} \end{cases} \quad (7)$$

where  $R$  is a reward value defined as the constant; the number of human interaction is  $m$ ; the number of reactions is  $n$ . The selection probability  $p_i(t)$  of  $i$ th reaction is calculated by using a Boltzmann selection scheme as follows,

$$p_i(t) = \frac{\exp(s_{i,x}(t)/T)}{\sum_j \exp(s_{j,x}(t)/T)} \quad (8)$$

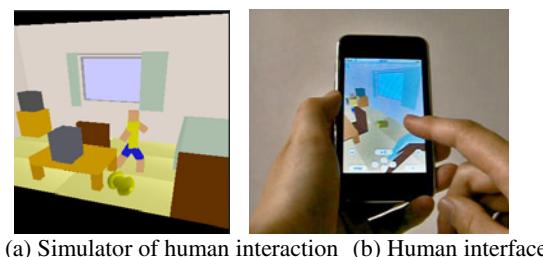
Where  $T$  is the temperature ( $T > 0$ ). When the temperature is high, the robot randomly selects a reaction. As the temperature decreases, the robot deterministically selects the reaction with the high selection strength even if the difference among selection strengths is very small.

## 4 Experimental Results

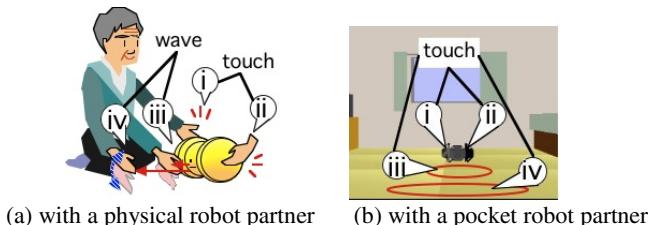
### 4.1 Communication Learning on a Pocket Robot Partner and a Physical Robot Partner

This section shows several experimental results of the proposed method. In order to realize the efficient and quick learning, we use iPhone as the pocket robot partner. The touch interface provides a user with simple and repeatable interaction on the touch screen. After the learning on the pocket robot partner, we can transfer the learning result to the physical robot partner. The human interaction is done with miuro on the simulator of iPhone by the touch interface (Fig.6). The point touched on the screen is calculated approximately according to the sensing range of miuro in order to reflect the physical interaction with miuro seamlessly. For example, the physical approach of a hand to the front of miuro is corresponding to the touch of the position below miuro on the simulator, and the physical contact to the right side of miuro is corresponding directly to the right side of miuro on the simulator. Basically, the behavior mode of miuro on the simulator is the autonomous behavior mode, and the mode is transited to the interactive behavior mode by the touch to miuro.

We conduct experiments on the learning of the relationship between human interaction and its corresponding preferable music selection as the generation of comfortable sound field. Figure 7 shows human interaction used in the experiments; (i) touch from the left side, (ii) touch from the right side, (iii) hand gesture near with miuro, and (iv) hand gesture away from miuro. The number of songs is 5, and the feature of these songs is different.



**Fig. 6.** Interaction with a pocket robot partner

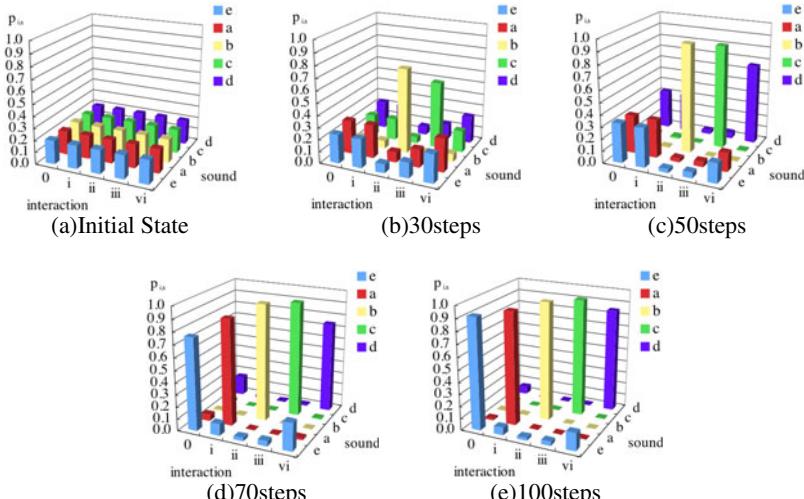


**Fig. 7.** Human interaction with miuro

We assume the person repeats the same interaction if the preferable song is selected according to the prediction of the person. Therefore, the selection strength is updated according to eq.(7) if the preferable song is selected according to the prediction of the person. The number of trials is 100 corresponding to the times of human interactions in an experiment. The parameters used in the experiment, temperature  $T$  is 0.1 and reward  $R$  is 0.05.

Figure 8 shows an experimental result of the learning through interaction with miuro on the simulator. The initial selection strength is set at the same value (Fig.8 (a)). The selection strength of two songs is increasing as the repeat of human interactions (Fig.8 (b)). The person tried to make miuro learn the relationship between these two interactions and these two songs, respectively. The rest of relationship learning was done through the repeated interaction (Fig.8 (c)). As a result, miuro learned the relationship between human interaction and its corresponding song as the final values of selection strength shown in the Fig.8 (d). However, the song (e) is not selected in the relationship learning since the number of songs is more than that of human interactions.

The learning result is shared with the decision making rules of miuro as a physical robot partner. Figure 9 shows an experimental result of the human interaction after the learning result is reflected to miuro. The person tried to four human interactions in the different order three times, *i.e.*, 12 times of human interactions in total. As a result, the predicted preferable songs are selected correctly according to their corresponding human interactions, respectively. Furthermore, the result of relationship learning in the interactive behavior mode in the physical robot partner is reflected to the interaction of miuro on the simulator in the pocket robot partner. In this way, the experimental results show that the proposed method can learn the relationship between human interaction and its corresponding behavior, and that the proposed method can provide the person with the preferable song as the comfortable sound field.



**Fig. 8.** The change of selection strength of miuro in an experiment



(a) Touch interaction from left side (b) Gesture interaction with miuro

**Fig. 9.** Interaction with a physical robot partner

## 5 Conclusion

This paper discussed the style of information service by robot partners. In order to perform information service suitable to the individual person, the robot partner should extract the preference and custom. Therefore, we propose a learning method of the relationship between human interaction and its corresponding behavior, and we focus on the music selection for providing the comfortable sound field for the person. The experimental results show that the proposed method can learn the relationship between human interaction and its corresponding behavior, and that the proposed method can provide the person with the preferable song as the comfortable sound field. Furthermore, we show the seamless sharing of information between different types of robot partners. This means the proposed method does not choose the communication style beforehand, and the user can choose suitable communication style according to the situation. However, we did not discuss the sharing of the situation between people and robot partners.

As future works, we intend to improve the performance of human gesture and interaction recognition, and discuss the architecture of informationally structured space [21] for the robot partners communicating with aged people. Furthermore, we will develop a map building method based on the location-dependent memory and long-term memory of human interaction in order to extract human intention correctly.

## References

1. Russell, S.J., Norvig, P.: *Artificial Intelligence*. Prentice-Hall, Inc., Englewood Cliffs (1995)
2. Brady, M., Paul, R.: *Robotics Research: The First International Symposium*. The MIT Press, Massachusetts (1984)
3. Brooks, R.A.: Planning Collision Free Motions for Pick and Place Operation. In: *Robotics Research*, pp. 5–38. MIT Press, Cambridge (1983)
4. Thrun, S., Burgard, W., Fox, D.: *Probabilistic Robotics*. MIT Press, Cambridge (2005)
5. Ishiguro, H., Ono, T., Imai, M., Maeda, T., Kanda, T., Nakatsu, R.: Robovie: A Robot Generates Episode Chains in Our Daily Life. In: *Proceedings of 32nd International Symposium on Robotics (ISR 2001)*, pp. 1356–1361 (2001)
6. Kanda, T., Ishiguro, H., Imai, M., Ono, T., Mase, K.: A Constructive Approach for Developing Interactive Humanoid Robots. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2002)*, pp. 1265–1270 (2002)

7. Kanda, T., Ishiguro, H., Imai, M., Ono, T.: Body Movement Analysis of Human-Robot Interaction. In: International Joint Conference on Artificial Intelligence (IJCAI 2003), pp. 177–182 (2003)
8. Ogawa, K., Ono, T.: ITACO: Constructing an Emotional Relationship between Human and Robot. In: The 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008), Munich, Germany, August 1-3 (2008)
9. Kidd, C.D., Taggart, W., Turkle, S.: A Sociable Robot to Encourage Social Interaction among the Elderly. In: IEEE International Conference on Robotics and Automation (ICRA 2006), Orlando, FL, USA, May 15-19 (2006)
10. ZMP Inc., Miuro, C.: The Robot that Play Your Favorite Music at Your Favorite Place. Journal of Robotics Society of Japan 26(1), 34–35 (2008) (in Japanese)
11. Eysenck, M.W.: Psychology an Integrated Approach. Prentice-Hall, Pearson Education, Ltd. (1998)
12. Jang, J.-S.R., Sun, C.-T., Mizutani, E.: Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice-Hall, Inc., Englewood Cliffs (1997)
13. Kubota, N., Nishida, K.: Perceptual Control Based on Prediction for Natural Communication of A Partner Robot. IEEE Transactions on Industrial Electronics 54(2), 866–877 (2007)
14. Kubota, N., Pei, J., Yorita, A., Kojima, H.: Human Tracking and Natural Communication of Partner Robots. In: Proc. (CD-ROM) of International Symposium on Advanced Intelligent Systems (ISIS 2007), Sokcho, Korea, September 5-8, pp. 205–210 (2007)
15. Liu, H.: A Fuzzy Qualitative Framework for Connecting Robot Qualitative and Quantitative Representations. IEEE Transactions on Fuzzy Systems 16(6), 1522–1530 (2008)
16. Kubota, N., Yorita, A.: Structured Learning for Partner Robots based on Natural Communication. In: Proc. (CD-ROM) of 2008 IEEE Conference on Soft Computing in Industrial Applications (SMCia 2008), Muroran, Japan, June 25-27, pp. 303–308 (2008)
17. Kubota, N.: Multi-Objective Design of Neuro-Fuzzy Controllers for Robot Behavior Coordination. In: Jin, Y. (ed.) Multi-Objective Machine Learning, pp. 557–584 (March 2006)
18. Kubota, N., Aizawa, N.: Self-Adaptation in Intelligent Formation Behaviors of Multiple Robots based on Fuzzy Control. In: 2009 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2009), Jeju Island, Korea, August 20-24 (2009)
19. Kubota, N., Wakisaka, S.: An Emotional Model Based on Location-Dependent Memory for Partner Robots. Journal of Robotics and Mechatronics 21(3) (2009)
20. Kubota, N., Nishida, K.: Cooperative Perceptual Systems for Partner Robots Based on Sensor Network. International Journal of Computer Science and Network Security 6(11), 19–28 (2006)
21. Kubota, N., Yorita, A.: Topological Environment Reconstruction in Informationally Structured Space for Pocket Robot Partners. In: 2009 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA2009), Daejeon, Korea, December 15-18 (2009)
22. Gregory, R.L.: The Oxford Companion to The Mind. Oxford University Press, Oxford (1998)
23. Gergen, K.J., Gergen, M.M.: Social Psychology, 2nd edn. Springer, New York (1986)
24. Brdiczka, Crowley, J.L., Reignier, P.: Learning Situation Models in A Smart Home. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 39(1), 56–63 (2009)
25. Pfeifer, R., Scheier, C.: Understanding Intelligence. The MIT Press, Hillsdale (1999)

26. Arkin, R.C., Fujita, M., Takagi, T., Hasegawa, R.: An Ethological and Emotional Basis for Human-Robot Interaction. *Robotics and Autonomous Systems* 4(3-4), 191–201 (2003)
27. Kubota, N., Liu, H., Mori, Y.: Visual Perception for Emotion-Based Communication of A Partner Robot. *Kansei Engineering International* 8(1), 41–50 (2009)
28. Khemapech, I., Duncan, I., Miller, A.: A Survey of Wireless Sensor Networks Technology. In: PGNET, Proc. the 6th Annual PostGraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting, EPSRC (2005)

# Improved Full-Discretization Method for Milling Chatter Stability Prediction with Multiple Delays

XiaoJian Zhang<sup>1</sup>, CaiHua Xiong<sup>1,\*</sup>, and Ye Ding<sup>2</sup>

<sup>1</sup> State Key Laboratory of Digital Manufacturing Equipment and Technology,  
School of Mechanical Science and Engineering, Huazhong University of Science and  
Technology, Wuhan 430074, China  
chxiong@mail.hust.edu.cn

<sup>2</sup> State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering,  
Shanghai Jiao Tong University, Shanghai 200240, China

**Abstract.** This paper presents an analytical stability prediction method in milling chatter analysis with multiple delays. The improved full-discretization method is used to determine the critical boundary of chatter stability. The proposed method can take many practical factors into consideration, such as helix angle, non-constant pitch, and cutter runout, etc. The influence of the multiple delays is explored in detail and the simulation results show that the cutter runout has the great influence on the elevation of the stability boundary. The finding proves the validity of the proposed method by experiment verification.

**Keywords:** Full-discretization; prediction; milling; chatter stability; multiple delays; runout.

## 1 Introduction

The regenerative effect is the general explanation for machine tool chatter, which can be modeled as delayed differential equations (DDEs) with infinite dimensional state spaces [1]. In the case of milling, the cutting is highly interrupted as each tooth enters and leaves the workpiece which can be described as a DDE with a time periodic coefficient. For the purpose of machining parameter optimization to improve productivity, different methods have been developed to determine the stability boundaries for milling, such as frequency domain techniques [2-3], semi-discretization method [4-5], temporal finite element analysis (TFEA) method [6], D-partition method [7], and numerical determined methods [8-11]. Generally, the semi-analytical predictions of stability take much less time than numerical methods to give stability regions over the process parameter space [12-14].

The complexity of the DDE increases when many practical factors, e.g., helix angle, non-constant pitch cutter, and cutter teeth runout, are considered. However, the stability prediction is more suitable for the actual situation with these practical factors being considered. The helix angle of the mill can have an important role on the areas of added flip lobes [12-13]. Zatarain *et al.* [15] analyzed the effect by multi-frequency solution

---

\* Corresponding author.

while Insperger *et al.* [16] using the semi-discretization method. Stability islands are separated by the lines where the axial depth of cut is equal to the multiples of the helix pitch, which differ from the parametrically induced islands reported in [14-15]. The non-constant pitch effect on chatter was demonstrated by frequency solution [16], which has enabled the optimization of tool geometry [17-18]. The variable pitch/helix tool case has also been analyzed with the semi-discretization method [22-23]. Radial runout is also an important factor for stability prediction since it changes the cutting geometry for different tooth, Schmitz *et al.* [19] examine the effect of cutter runout on surface topography, surface location error and stability in end milling based on time-domain simulation. Recently, Insperger *et al.* [20] showed that the effect of the tool runout can sometimes prevent the proper determination of stability, and the analysis of the vibration signal and the corresponding Poincare section were used to determine if chatter vibrations happened. Radial runout can be considered as one of the multiple delays problems which have been widely studied [16-18].

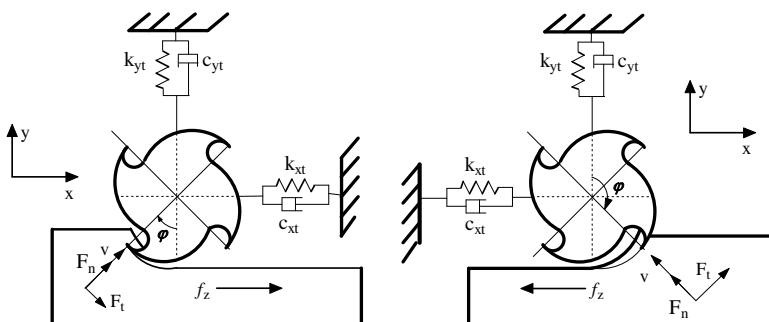
In this paper, an analytical prediction method of milling chatter stability is presented based on the full-discretization method (FDM) using the direct integration scheme first proposed by Ding *et al.* [21], and the chatter stability limits will be investigated by considering the effects of the helix angle, the non-constant pitch and the cutter runout via the updated FDM. The remainder of this paper is organized as follows. In Section 2, the mechanical model is introduced and different factors are included. In Section 3, the updated FDM is given for prediction of the milling stability with the many factors, and multiple delays, especially. In Section 4, the effects of different factors for the milling chatter stability limits are described in detail and the experimental results prove the efficiency of the proposed method. Conclusions are presented in the last section.

## 2 Mechanical Model

The mechanical model of 2 DOF milling process is shown in Fig.1. The tool is assumed to flexible relative to the rigid workpiece. The equation of the two degrees of freedom is:

$$\begin{aligned} m_x \ddot{x} + c_x \dot{x} + k_x x &= F_x(t), \\ m_y \ddot{y} + c_y \dot{y} + k_y y &= F_y(t). \end{aligned} \quad (1)$$

The cutting forces  $F_x(t)$  and  $F_y(t)$  are time-dependent due to the rotation of the tool.



**Fig. 1.** Mechanical model (up-milling and down-milling)

The cutter can be modeled as a stack of differential disk elements along the depth of cut, as shown in Fig. 2. The tangential and the normal forces acting on this differential element are:

$$\begin{aligned} dF_{jt}(t, z) &= g(\varphi_j(t, z)) [K_{tc}(f(t, z))dz + K_{te}dz], \\ dF_{jn}(t, z) &= g(\varphi_j(t, z)) [K_{nc}(f(t, z))dz + K_{ne}dz]. \end{aligned} \quad (2)$$

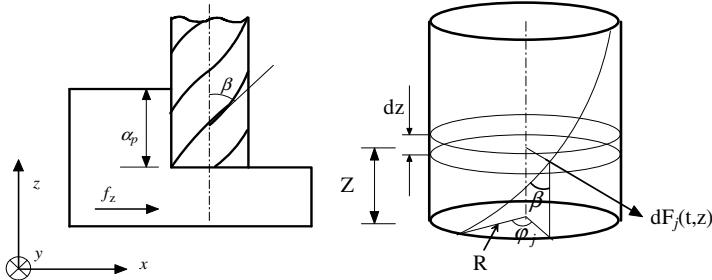
where  $K_t$  and  $K_n$  are the tangential and the normal cutting coefficients, respectively,  $f(t, z)$  is the instantaneous chip thickness. The angular position of the  $j$ -th cutting edge corresponding to the investigated differential element of the tool is

$$\varphi_j(t, z) = \frac{2\pi\Omega}{60}t - \frac{z \tan \beta}{R} + j \frac{2\pi}{N}, \quad (3)$$

where  $\Omega$  is the spindle speed in [rpm],  $z$  denotes the axial direction,  $\beta$  is the helix angle,  $R$  is the radius of the tool and  $N$  is the number of teeth. The function  $g_j(t)$  is the screen function, it is equal to 1, if the  $j$ -th tooth is in cut, and it is 0, if not:

$$g(\varphi_j(t, z)) = \begin{cases} 1 & \varphi_{in} < \varphi_j(t, z) < \varphi_{ex} \\ 0 & \text{otherwise} \end{cases}, \quad (4)$$

where  $\varphi_{in}$  and  $\varphi_{ex}$  are the angles where the teeth enter and exit the cut, respectively.



**Fig. 2.** Cutting force model with helix angle

Generally, the feed per tooth  $f_{jz} \ll R$  ( $j=1, \dots, N$ ), and the path of the cutter flute is approximated by a circular path. The  $x$  and  $y$  components of the differential cutting force are given as (see Fig. 1):

$$\begin{aligned} dF_{jx}(t, z) &= dF_{jt}(t, z) \cos \varphi_j(t, z) + dF_{jn}(t, z) \sin \varphi_j(t, z), \\ dF_{jy}(t, z) &= -dF_{jt}(t, z) \sin \varphi_j(t, z) + dF_{jn}(t, z) \cos \varphi_j(t, z). \end{aligned} \quad (5)$$

And the chip thickness can be expressed as

$$f(\varphi_j(t, z)) = (f_{jz} + x(t - T_j) - x(t)) \sin(\varphi_j(t, z)) + (y(t - T_j) - y(t)) \cos(\varphi_j(t, z)) \quad (6)$$

where  $f_{jz}$  is the nominal feed per tooth and  $T_0 = 60/N\Omega$  is the nominal tooth passing period. It is known that the enter and exit angles are different for each teeth [22]. The

value of delay corresponding to every disk element can be determined using the methods in Ref. [23].

The cutting force acting on tooth  $j$  is obtained by integration of Eq.(5) along the axial direction:

$$\begin{aligned} F_{jx}(t) &= \int_0^{a_p} dF_{jx}(t, z) = \int_0^{a_p} g(\varphi_j(t, z))(f_j(t, z))(K_{tc} \cos \varphi_j(t, z) + K_{nc} \sin \varphi_j(t, z))dz \\ &\quad + \int_0^{a_p} g(\varphi_j(t, z))(K_{te} \cos \varphi_j(t, z) + K_{ne} \sin \varphi_j(t, z))dz, \\ F_{jy}(t) &= \int_0^{a_p} dF_{jy}(t, z) = \int_0^{a_p} g(\varphi_j(t, z))(f_j(t, z))(-K_{tc} \sin \varphi_j(t, z) + K_{nc} \cos \varphi_j(t, z))dz \\ &\quad + \int_0^{a_p} g(\varphi_j(t, z))(-K_{te} \sin \varphi_j(t, z) + K_{ne} \cos \varphi_j(t, z))dz \end{aligned} \quad (7)$$

where  $a_p$  is the axial depth of cut. The resultant forces are the sum of the forces acting on the teeth:

$$F_x = \sum_{j=1}^N F_{jx}(t), \quad \text{and} \quad F_y = \sum_{j=1}^N F_{jy}(t). \quad (8)$$

By substituting Eq.(7) into Eq.(1), the equations of motion can be written as follows:

$$\begin{aligned} m_x \ddot{x} + c_x \dot{x} + k_x &= \sum_{j=1}^N \int_0^{a_p} g(\varphi_j(t, z))(f_j(t, z))(K_{tc} \cos \varphi_j(t, z) + K_{nc} \sin \varphi_j(t, z))dz + f_{x0} \\ m_y \ddot{y} + c_y \dot{y} + k_y &= \sum_{j=1}^N \int_0^{a_p} g(\varphi_j(t, z))(f_j(t, z))(-K_{tc} \sin \varphi_j(t, z) + K_{nc} \cos \varphi_j(t, z))dz + f_{y0}. \end{aligned} \quad (9)$$

where

$$\begin{aligned} f_{x0} &= \sum_{j=1}^N \int_0^{a_p} g(\varphi_j(t, z))(K_{te} \cos \varphi_j(t, z) + K_{ne} \sin \varphi_j(t, z))dz \\ f_{y0} &= \sum_{j=1}^N \int_0^{a_p} g(\varphi_j(t, z))(-K_{te} \sin \varphi_j(t, z) + K_{ne} \cos \varphi_j(t, z))dz \end{aligned} \quad (10)$$

### 3 Improved Full-Discretization Method

It is obviously that Eq.(9) has multiple delays. When the cutter runout is not considered, Eq.(9) is degenerated to the single delay problem, by using some simple algebraic transformation, the degenerated form can be written as

$$\dot{\mathbf{x}}(t) = \mathbf{A}_0 \mathbf{x}(t) + \mathbf{A}(t) \mathbf{x}(t) - \mathbf{A}(t) \mathbf{x}(t-T) + \mathbf{B} \mathbf{f}(t) \quad (11)$$

#### 3.1 Signal Delay Problem

The signal delay problem is introduced by the FDM presented by Ding et al which has high computational efficiency and high numerical precision [21].

By the transformation as described in [24], each term of Eq.(11) can be determined as follows:

$$\mathbf{A}_0 = \begin{bmatrix} -\mathbf{M}^{-1} \mathbf{C}/2 & \mathbf{M}^{-1} \\ \mathbf{C}\mathbf{M}^{-1} \mathbf{C}/4 - \mathbf{K} & -\mathbf{C}\mathbf{M}^{-1}/2 \end{bmatrix}, \mathbf{x}(t) = \begin{bmatrix} \mathbf{q}(t) \\ \mathbf{p}(t) \end{bmatrix}, \mathbf{q}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \mathbf{p} = \mathbf{M}\dot{\mathbf{q}} + \mathbf{C}\mathbf{q}/2$$

$$\mathbf{A}(t) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -m_x^{-1} \sum_{j=1}^N \int_0^{a_p} g \cdot (K_{tc} \cos \varphi_j + K_{nc} \sin \varphi_j) \sin \varphi_j dz & -m_x^{-1} \sum_{j=1}^N \int_0^{a_p} g \cdot (K_{tc} \cos \varphi_j + K_{nc} \sin \varphi_j) \cos \varphi_j dz & 0 & 0 \\ -m_y^{-1} \sum_{j=1}^N \int_0^{a_p} g \cdot (-K_{tc} \sin \varphi_j + K_{nc} \cos \varphi_j) \sin \varphi_j dz & -m_y^{-1} \sum_{j=1}^N \int_0^{a_p} g \cdot (-K_{tc} \sin \varphi_j + K_{nc} \cos \varphi_j) \cos \varphi_j dz & 0 & 0 \end{bmatrix}$$

and

$$\mathbf{Bf}(t) = updown * \begin{bmatrix} 0 \\ 0 \\ m_x^{-1} \sum_{j=1}^N \int_0^{a_p} g \cdot (K_{tc} \cos \varphi_j + K_{nc} \sin \varphi_j) dz + m_x^{-1} f_{x0} \\ m_y^{-1} \sum_{j=1}^N \int_0^{a_p} g \cdot (-K_{tc} \sin \varphi_j + K_{nc} \cos \varphi_j) dz + m_y^{-1} f_{y0} \end{bmatrix}$$

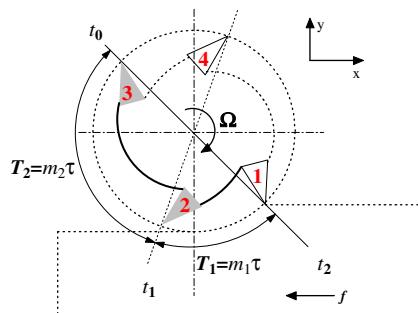
$$updown = \begin{cases} 1, & \text{for upmilling} \\ -1, & \text{for downmilling} \end{cases}$$

### 3.2 Multiple Delays Problem

When the teeth runout or non-constant pitch is considered, Eq. (9) can be put in the following vector form:

$$\dot{\mathbf{x}}(t) = \mathbf{A}_0 \mathbf{x}(t) + \sum_{j=1}^N [\mathbf{A}_j(t)(\mathbf{x}(t) - \mathbf{x}(t - T_j))] + \mathbf{Bf}(t) \quad (12)$$

The multiple delay form in milling process can describe several practical cases: constant pitch cutter with runout [22], non-constant pitch cutter without runout [18, 20, 25-27] and unevenly pitched cutter with runout. Firstly, a problem with  $N$  delays are given to illustrate the updated method of the FDM [21].



**Fig. 3.** Model with  $N$  delays

A  $N$ -tooth cutter with non-constant pitch is shown as in Fig.3. The least period of the dynamic milling system equals the spindle rotation period ( $T$ ). First, equally divide  $T=T_1+T_2+\dots+T_N$  into  $m=m_1+m_2+\dots+m_N$  small time intervals and  $T=(m_1+m_2+\dots+m_N)\tau$ , where  $m_1=\text{int}(T_1/\tau)$ ,  $m_2=\text{int}((T_1+T_2)/\tau)-m_1, \dots$ ,  $m_N=\text{int}(T/\tau)-\sum_{j=1}^{N-1}m_j$ , and ‘int’ indicates the operation that rounds positive number towards zero. The instant ‘ $t_2$ ’ is designated as the start of the time, i.e. the instant of zero.

Similar to the single delay problem, on each time interval  $k\tau \leq t \leq (k+1)\tau$ , ( $k=0, \dots, m$ ), the response of Eq.(12) with the initial condition  $\mathbf{x}_k = \mathbf{x}(k\tau)$  can be transformed as

$$\begin{aligned}\mathbf{x}_{k+1} = & e^{\mathbf{A}_0\tau} \mathbf{x}(k\tau) + \int_0^\tau \left\{ e^{\mathbf{A}_0\xi} \sum_{j=1}^N \mathbf{A}_j(k\tau+\tau-\xi) [\mathbf{x}(k\tau+\tau-\xi) - \mathbf{x}(k\tau+\tau-\xi-T_j)] \right\} d\xi \\ & + \int_0^\tau \left[ e^{\mathbf{A}_0\xi} \sum_{j=1}^N \mathbf{Bf}_j(k\tau+\tau-\xi) \right] d\xi\end{aligned}\quad (13)$$

where

$$\mathbf{A}_j(t) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \mathbf{A}_j(3,1) & \mathbf{A}_j(3,2) & 0 & 0 \\ \mathbf{A}_j(4,1) & \mathbf{A}_j(4,2) & 0 & 0 \end{bmatrix}, \quad \begin{aligned}\mathbf{A}_j(3,1) = & -m_x^{-1} \int_0^{a_p} g \cdot (K_{ic} \cos \varphi_j + K_{nc} \sin \varphi_j) \sin \varphi_j dz \\ \mathbf{A}_j(3,2) = & -m_x^{-1} \int_0^{a_p} g \cdot (K_{ic} \cos \varphi_j + K_{nc} \sin \varphi_j) \cos \varphi_j dz \\ \mathbf{A}_j(4,1) = & -m_y^{-1} \int_0^{a_p} g \cdot (-K_{ic} \sin \varphi_j + K_{nc} \cos \varphi_j) \sin \varphi_j dz \\ \mathbf{A}_j(4,2) = & -m_y^{-1} \int_0^{a_p} g \cdot (-K_{ic} \sin \varphi_j + K_{nc} \cos \varphi_j) \cos \varphi_j dz\end{aligned}$$

and  $\varphi_j = 2\pi \frac{m_j}{m}$ , ( $j=1, 2, \dots, N$ ) is the pitch angle for each tooth.

The state item, the time-periodic terms and the time-delay term are approximately linearized like in [21] and substituted into Eq.(13) to produce:

$$\mathbf{x}_{k+1} = \left( \mathbf{F}_0 + \sum_{j=1}^N \mathbf{F}_{0,1}^{<j>} \right) \mathbf{x}_k + \sum_{j=1}^N \mathbf{F}_{k+1}^{<j>} \mathbf{x}_{k+1} + \sum_{j=1}^N \mathbf{F}_{m-1}^{<j>} \mathbf{x}_{k+1-m} + \sum_{j=1}^N \mathbf{F}_m^{<j>} \mathbf{x}_{k-m} + \sum_{j=1}^N \mathbf{C}_k^{<j>} \quad (14)$$

where

$$\mathbf{F}_{0,1}^{<j>} = \left[ \frac{\Phi_2}{\tau} \mathbf{A}_{j,k+1} + \frac{\Phi_3}{\tau} \frac{\mathbf{A}_{j,k} - \mathbf{A}_{j,k+1}}{\tau} \right],$$

$$\mathbf{F}_{k+1}^{<j>} = (\Phi_1 - \Phi_2 / \tau) \mathbf{A}_{j,k+1} + (\Phi_2 - \Phi_3 / \tau) \frac{\mathbf{A}_{j,k} - \mathbf{A}_{j,k+1}}{\tau}$$

$$\mathbf{F}_{m-1}^{<j>} = -\mathbf{F}_{k+1}^{<j>}$$

$$\mathbf{F}_m^{<j>} = -\mathbf{F}_{0,1}^{<j>}$$

$$\mathbf{C}_k^{<j>} = \Phi_1 \mathbf{Bf}_{j,k+1} + \Phi_2 (\mathbf{Bf}_{j,k} - \mathbf{Bf}_{j,k+1}) / \tau$$

and

$$\Phi_0 = e^{\mathbf{A}_0\tau}, \Phi_1 = \int_0^\tau e^{\mathbf{A}_0\xi} d\xi, \Phi_2 = \int_0^\tau \xi e^{\mathbf{A}_0\xi} d\xi, \Phi_3 = \int_0^\tau \xi^2 e^{\mathbf{A}_0\xi} d\xi$$

The discrete map form is

$$\mathbf{y}_{k+1} = \mathbf{D}_k \mathbf{y}_k + \mathbf{E}_k \quad (15)$$

Where the  $4(m+1)$  dimensional vector  $\mathbf{y}_k$  is

$$\mathbf{y}_k = \text{col}(\mathbf{x}_k \ \mathbf{x}_{k-1} \dots \mathbf{x}_{k+1-m_i} \ \mathbf{x}_{k-m_i} \dots \mathbf{x}_{k+1-m_j} \ \mathbf{x}_{k-m_j} \dots \mathbf{x}_{k+1-m} \ \mathbf{x}_{k-m}) \quad (16)$$

However, the coefficient matrix  $\mathbf{D}_k$  is different

$$\mathbf{D}_k = \begin{bmatrix} \text{inv} \cdot \left( \mathbf{F}_0 + \sum_{j=1}^N \mathbf{F}_{0,j}^{<j>} \right) & 0 & 0 & \dots & 0 & \text{inv} \cdot \mathbf{F}_{k+1}^{<j>} & \text{inv} \cdot \mathbf{F}_{k+1}^{<j>} & \dots & \text{inv} \cdot \mathbf{F}_{k+1}^{<j>} & \text{inv} \cdot \mathbf{F}_{k+1}^{<j>} & \dots & 0 & 0 & 0 \\ \mathbf{I} & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \mathbf{I} & 0 & 0 \end{bmatrix} \quad (17)$$

( $i, j \in \{m_1, m_2, \dots, m_N\}$ )

$$\mathbf{E}_k = \text{col}(\left[ \mathbf{I} - \mathbf{F}_{j,k+1} \right]^{-1} \cdot \mathbf{C}_k^{<j>} \cdot \mathbf{0}, \dots, \mathbf{0}), \text{ where } \text{inv} = \left[ \mathbf{I} - \sum_{j=1}^N \mathbf{F}_{k+1}^{<j>} \right]^{-1}$$

If  $m_j=m_i$  ( $i \neq j$ ), the corresponding coefficients of  $m_i$  and  $m_j$  in Eqs.14-17 must be added.

## 4 Verification

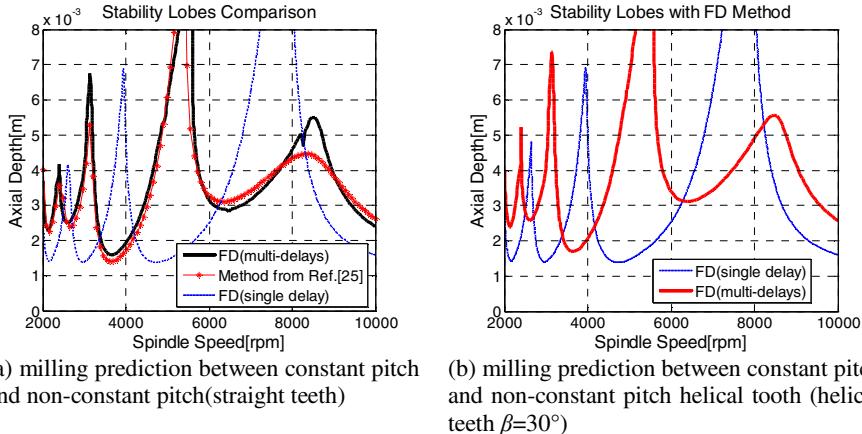
The proposed method for milling prediction is compared with the prior methods with several cases.

### Case1: non-constant pitch tool without runout

A four-fluted variable pitch cutter with a diameter of 19.05mm and a helix angle of 30° in [25] is used for verification of the method. It has varying pitch angles 70 deg-110 deg-70 deg-110 deg, so the model has 4 delays:  $T_1=7/36T$ ,  $T_2=11/36T$ ,  $T_3=7/36T$ ,  $T_4=11/36T$ . The modal parameters and the cutting force coefficients are same as ref [25], the discretization number is 36 and the stability charts are plotted as in Fig.4 for the material Al 356. As shown in Fig.4, the constant pitch case is also added for comparison.

It is known that the milling stability has been predicted by Altintas et al. by time domain simulation and frequency domain method in Ref. [25]. By the proposed method, the milling predictions are same with the results in Ref. [25], as shown in Fig.4. The computation time is longer than the frequency domain method, but this method can take helix angle into account compared with the frequency domain method.

It is shown that the helix angle has little influence of stability in large radial cutting, and the method proposed can predict the milling stability well. However, the updated method can't be used for geometry design of tool as in Refs. [20-21].

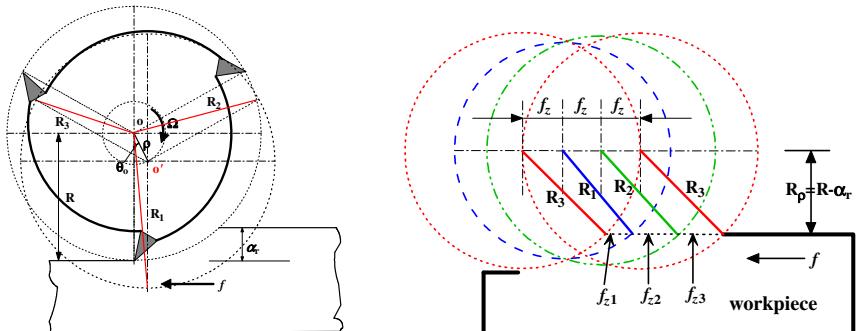


**Fig. 4.** Stability lobes for regular and non-constant pitch cutter for half immersion down milling of work material Al356;  $K_t = 697$  MPa,  $K_r = K_n / K_t = 0.367$ .

### Case2: constant pitch tool with runout

Generally, radial runout exists because of spindle imperfections, misalignments, thermal deformation and wear of components in spindle assembly, etc.

As shown in Fig.5, only the tool radial runout is considered, and the ‘axial runout’ is neglected[28]. Here, a three-flute end milling model is adopted for illustration, ‘o’ is the center of geometric axis and ‘o” is the center of rotating axis.



**Fig. 5.** Cutting force model with radial runout

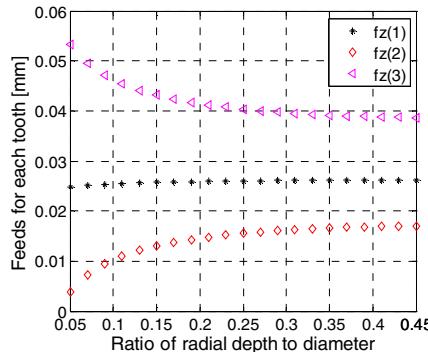
Since the existence of the radial runout ( $\rho$ ), the feed is different for each tooth. Suppose the offset is a constant vector ( $\rho$ ,  $\theta_0$  is independent of time) and according to the geometrical relationship in Fig.5 to get the actual radius and feed for each tooth:

$$R_1 = \sqrt{R^2 + \rho^2 - 2R\rho \cos(\pi - \theta_0)}, \quad f_{z1} = f_z + \sqrt{R_1^2 - R_\rho^2} - \sqrt{R_3^2 - R_\rho^2}$$

$$R_2 = \sqrt{R^2 + \rho^2 - 2R\rho \cos(\pi/3 - \theta_0)}, \quad f_{z2} = f_z + \sqrt{R_2^2 - R_\rho^2} - \sqrt{R_1^2 - R_\rho^2}$$

$$R_3 = \sqrt{R^2 + \rho^2 - 2R\rho \cos(\pi/3 + \theta_0)}, f_{z3} = f_z + \sqrt{R_3^2 - R_\rho^2} - \sqrt{R_2^2 - R_\rho^2}$$

If  $f_{zj} < 0$  ( $j=1, \dots, N$ ), the  $j^{\text{th}}$  tooth never cut. For example,  $f_z=0.0273\text{mm}$ ,  $\rho=7.2\text{um}$ ,  $\theta_0=65.09^\circ$ ,  $R=4\text{mm}$ . The relationship between feeds for each teeth and the ratio of radial depth to tool diameter ( $aD$ ) is given as shown in Fig.6.



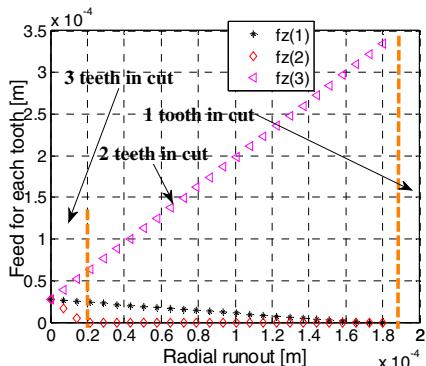
**Fig. 6.** Relationship between  $f_{zj}$  ( $j=1, 2, \dots, N$ ) and  $aD$

When the nominal feed, tool diameter and runout are fixed, the small of radial depth, the larger differences of feeds for each teeth. For a fixed radial depth, when radial runout grows, the differences of feeds for each tooth grow, as shown in Fig.7, and the milling stability lobes are changing correspondingly.

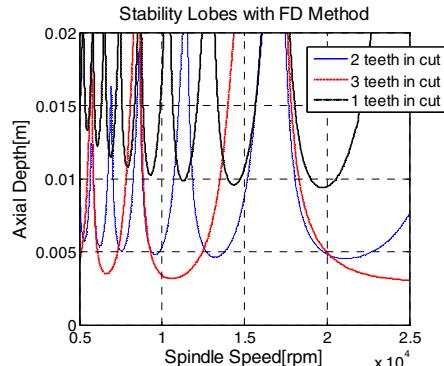
Here the parameters in Ref. [22] are adopted unchanged just for illustration, to show the effect of runout on stability limits. Down milling tests are analyzed with a diameter of 16mm and a helix angle of  $30^\circ$ . The other parameters are listed in Table 1. When runout occurred and the number of teeth changed, the corresponding milling stability limits are predicted as in Fig.8. The number of discretization adopted is 36, and the accuracy analysis can be seen in Ref.[29].

**Table 1.** Cutting parameters of the milling system [22]

Direction	Mode	Natural frequencies (Hz)	Modal effective masses (kg)	Damping ratios
X	1	898.22	1.576	0.040041
	2	1135.3	44.259	0.00535
Y	1	858.51	0.852	0.036768
	2	1185.1	22.589	0.00777
Cutting coefficients	$K_t = 1209.355\text{MPa}$ , $K_n = 501.095\text{MPa}$			



**Fig. 7.** Relationship of feeds and radial runout for different tooth,  $aD=5/16$



**Fig. 8.** Stability lobes for different number of tooth in cut due to radial runout(3 teeth in cut is the single delay case)

As can be seen from Fig.8, the occurrence of cutter runout can lead to high stability boundary. It is due to the effect of the changed delays. The prediction results are in accordance with the analysis in Ref. [22] that if radial runout is large enough the milling process of a three-fluted cutter becomes a single tooth cutter cutting with the increased feed.

## 5 Conclusion

In this work, the FDM is updated for milling stability prediction with multiple delays and radial runout. The formulations are deduced by taking into account some practical factors, such as helix angle of tool, radial runout and non-constant pitch of tool, which have different influences on milling lobes. The experimental results in the literature verified the validity of the proposed method.

## Acknowledgements

This work was partially supported by the National Key Basic Research Program (grant 2005CB724103), and the National Natural Science Foundation of China (Grant 50835004). The authors would also like to thank Prof. E. Budak for his help in implementing the program with variable pitch cutters by the frequency domain method.

## References

- [3] Merdol, S.D., Altintas, Y.: Multi Frequency Solution of Chatter Stability for Low Immersion Milling. *Journal of Manufacturing Science and Engineering, Transactions of the ASME* 126(3), 459–466 (2004)
- [4] Insperger, T., Stépán, G.: Semi-discretization method for delayed systems. *International Journal for Numerical Methods in Engineering* 55(5), 503–518 (2002)
- [5] Insperger, T., Stépán, G.: Updated semi-discretization method for periodic delay-differential equations with discrete delay. *International Journal for Numerical Methods in Engineering* 61(1), 117–141 (2004)
- [6] Bayly, P.V., Halley, J.E., Mann, B.P., Davies, M.A.: Stability of Interrupted Cutting by Temporal Finite Element Analysis. *Journal of Manufacturing Science and Engineering, Transactions of the ASME* 125(2), 220–225 (2003)
- [7] Faassen, R., Van de Wouw, N., Oosterling, J., Nijmeijer, H.: Prediction of regenerative chatter by modelling and analysis of high-speed milling. *International Journal of Machine Tools and Manufacture* 43(14), 1437–1446 (2003)
- [8] Smith, S., Tlusty, J.: An overview of modeling and simulation of the milling process. *Journal of Engineering for Industry (Transactions of the ASME)* 113(2), 169–175 (1991)
- [9] Balachandran, B.: Nonlinear dynamics of milling processes. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 793–819 (2001)
- [10] Zhao, M., Balachandran, B.: Dynamics and stability of milling process. *International Journal of Solids and Structures* 38(10-13), 2233–2248 (2001)
- [11] Schmitz, T., Davies, M., Medicus, K., Snyder, J.: Improving high-speed machining material removal rates by rapid dynamic analysis. *CIRP Annals-Manufacturing Technology* 50(1), 263–268 (2001)
- [12] Zatarain, M., Munoa, J., Peigne, G., Insperger, T.: Analysis of the influence of mill helix angle on chatter stability. *CIRP Annals - Manufacturing Technology* 55(1), 365–368 (2006)
- [13] Insperger, T., Munoa, J., Zatarain, M., Peigne, G.: Unstable islands in the stability chart of milling processes due to the helix angle. In: *CIRP 2nd Int. Conference High Performance Cutting*, Vancouver, British Columbia, Canada (2006)
- [14] Szalai, R., Stépán, G.: Stability boundaries of high-speed milling corresponding to period doubling are essentially closed curves. In: *Proceedings of IMECE 2003 ASME International Mechanical Engineering Congress and R&D Expo*, Washington D.C, pp. 63–68 (2003)
- [15] Szalai, R., Stépán, G.: Lobes and lenses in the stability chart of interrupted turning. *Journal of Computational and Nonlinear Dynamics* 1(3), 205–211 (2006)
- [16] Altintas, Y., Engin, S., Budak, E.: Analytical stability prediction and design of variable pitch cutters. *Journal of Manufacturing Science and Engineering* 121(2), 173–178 (1999)
- [17] Budak, E.: An analytical design method for milling cutters with nonconstant pitch to increase stability, Part 2: Application. *Journal of Manufacturing Science and Engineering, Transactions of the ASME* 125(1), 35–38 (2003)
- [18] Budak, E.: An analytical design method for milling cutters with nonconstant pitch to increase stability, Part I: Theory. *Journal of Manufacturing Science and Engineering, Transactions of the ASME* 125(1), 29–34 (2003)
- [19] Schmitz, T.L., Couey, J., Marsh, E., Mauntler, N., Hughes, D.: Runout effects in milling: Surface finish, surface location error, and stability. *International Journal of Machine Tools and Manufacture*. 47(5), 841–851 (2007) (SPEC. ISS.)
- [20] Insperger, T., Mann, B.P., Surmann, T., Stépán, G.: On the chatter frequencies of milling processes with runout. *International Journal of Machine Tools and Manufacture* 48(10), 1081–1089 (2008)

- [21] Ding, Y., Zhu, L., Zhang, X., Ding, H.: A full-discretization method for prediction of milling stability. *International Journal of Machine Tools and Manufacture* 50(5), 502–509 (2010)
- [22] Wan, M., Zhang, W., Dang, J., Yang, Y.: A unified stability prediction method for milling process with multiple delays. *International Journal of Machine Tools and Manufacture* 50(1), 29–41 (2010)
- [23] Kline, W., DeVor, R.: The effect of runout on cutting geometry and forces in end milling. *International journal of machine tool design & research* 23(2-3), 123–140 (1983)
- [24] Gu, Y., Chen, B., Zhang, H., Guan, Z.: Precise time-integration method with dimensional expanding for structural dynamic equations. *AIAA journal* 39(12), 2394–2399 (2001)
- [25] Altintas, Y., Engin, S., Budak, E.: Analytical stability prediction and design of variable pitch cutters. *Journal of Manufacturing Science and Engineering* 121(2), 173–178 (1999)
- [26] Gouskov, A., Voronov, S., Paris, H., Batzer, S.: Nonlinear dynamics of a machining system with two interdependent delays. *Communications in Nonlinear Science and Numerical Simulation* 7(4), 207–221 (2002)
- [27] Sims, N.D., Mann, B., Huyanan, S.: Analytical prediction of chatter stability for variable pitch and variable helix milling tools. *Journal of Sound and Vibration* 317(3-5), 664–686 (2008)
- [28] Liang, S., Wang, J.: Milling force convolution modeling for identification of cutter axis offset. *International Journal of Machine Tools and Manufacture* 34(8), 1177–1190 (1994)
- [29] Insperger, T.: Full-discretization and semi-discretization for milling stability prediction: Some comments. *International Journal of Machine Tools and Manufacture* 50(7), 658–662 (2010)

# An Algorithm to Generate Compact Dual NURBS Tool Path with Equal Distance for 5-Axis NC Machining

Qingzhen Bi, Yuhang Wang, Limin Zhu, and Han Ding

Shanghai Jiaotong University, Shanghai 200240, P.R. China

{biqz,yhwang,zhulm,hding}@sjtu.edu.cn,

<http://www.sjtu.edu.cn>

**Abstract.** A new algorithm is proposed to generate compact dual NURBS tool paths with equal distance (DNTPED) for 5-axis NC machining. The DNTPED has significant advantages over conventional linear tool path in the NC machining of free-form surface since it reduces the tangency discontinuities along the tool path. The discontinuities, the inherent character of the linear interpolation, are the important sources of practical feed-speed fluctuation of the machine tool. The new algorithm is proposed to generate DNTPED based on the theories of rational motion. Here the rational rigid movement of a cutter is represented by a B-spline curve in dual quaternion space. DNTPED is obtained directly from the B-spline dual quaternion curve with the help of the blossom form of the B-spline curve. Comparing with the existing method, the DNTPED consists of less data and each NURBS can be treated as one curve, not as a collection of bézier segments.

## 1 Introduction

5-axis NC machine has been widely used in the sculpture machining due to the ability to manufacture complicated shapes, reduce the number of set ups and allow the use of shorter cutter. High demands on surface quality and cutting efficiency require the 5-axis process must be adapted to the requirements of high speed machining (HSM). The smoothness of feed velocity are obviously desirable in HSM. However, in practical 5-axis machining, the linear tool path is mostly adopted and it suffers from a number of shortcomings in the smoothness of feed velocity. The important source of feed fluctuation is the tangency and curvature discontinuities at the segment junctions of linear tool path since the discontinuities usually cause frequent decelerations and re-accelerations. Though some lookahead algorithm is helpful to achieve a higher feed rate, the feed fluctuation at segment cannot be completely avoided because of the discontinuous nature of the linear tool path. Non-uniform rational B-splines (NURBS) tool path can reduces the discontinuities along the tool path and has been proven to be superior over linear tool path.

There are two essential challenges for the application of NURBS tool path: NURBS tool path generation in CAM software and interpolator in NC unit. The proposed algorithm is located at the CAM stage dedicated to the tool path generation for 5-axis machining.

Existing researches focus on the NURBS tool path generation for 3-axis NC machining [1] while the NURBS tool path generation method for 5-axis NC machining is seldom studied. The NURBS tool path for 5-axis machining is more difficult because the trajectories of both the tool axis and the tool control point must be simultaneously considered. Fleisig [2] propose three-spline method to describe the 5-axis tool path. The position curve and the orientation curve are represented by two near arc length parameterized quintic polynomial splines respectively. The parameter synchronization of the two splines are realized by a third spline. Ueda [3] develop a NURBS tool path interpolation algorithm in post-process of 5-axis NC machining. Two spline curve are used. The tool tip spline is interpolated in the translational coordinate system. The tool axis spline is interpolated in the rotational coordinate system which is defined by the structure of the 5-axis machine tool. Koch [4] adopts two B-spline curves to guide tool path for 5-axis machining. The tool is guided to move along a first guide path while a further point on the axis of the tool, which is located at a distance from the work portion, is guided along a prescribed second path. The DNTPED for 5-axis machining is proposed in Ref. [5] and supported by some commercial NC system such as Siemens 840D [6]. The DNTPED is defined by the two NURBS curves of the tool tip point and a second point on the tool axis. The NURBS tool trajectory of both tool tip point and tool direction are guaranteed.

Two algorithms are proposed to generate dual NURBS tool path in existing researches. One is proposed in Ref. [5]. By combining the tool tip location and the second point on the tool axis, a discrete point in  $\mathbb{R}^6$  is first obtained. The two B-spline curves are generated by interpolating the discrete points in  $\mathbb{R}^6$ . The method ensures the synchronization of parameters of the two B-spline curves, but the equal distance cannot be guaranteed. Zhang [7] and Bi [8] study the 5-axis NC tool path described by a Bézier motion based on the theories of line geometry and kinematics. The trajectories of all points in the cutter undergoing the Bézier motion is a rational Bézier curve. The DNTPED is then easily obtained and the parameters of the two curves are synchronized naturally.

The proposed algorithm generates a DNTPED based on a B-spline motion. The obtained tool path is more compact than the one interpolated by the Bézier motion. The remainder of the paper is organized as follows. In Section 2, the preliminary knowledge of dual quaternion are briefly introduced. Section 3 constructs the cutter's B-spline motion. Based on the B-spline motion, Section 4 presents the algorithm to generate DNTPED. The advantages and validation of the algorithm are showed by a computational example in Section 5. Section 6 draws a conclusion.

## 2 Representation of a Spatial Displacement with a Dual Quaternion

This section introduces the dual quaternion for describing rational B-spline motion of a cutter, as noted in Ref. [8] and reported in detail in Ref. [9].

The dual quaternion is the combination of a dual number and two quaternions, which is represented as

$$\hat{\mathbf{Q}} = \mathbf{Q} + \varepsilon \mathbf{R} \quad (1)$$

where the symbol  $\varepsilon$  is the dual unit such that  $\varepsilon^2 = 0$ . The real part,  $\mathbf{Q} = Q_1 i + Q_2 j + Q_3 k + Q_4$ , is the quaternion defined by the Euler parameters of a rotation where  $i, j, k$  are the quaternion units satisfying the conditions  $i^2 = j^2 = k^2 = -1$ , and  $ij = -ji = k$ . The dual part is also a quaternion given by:

$$\mathbf{R} = \frac{1}{2} \mathbf{DQ} \quad (2)$$

where  $D = d_1 i + d_2 j + d_3 k$  is the quaternion formed from the translation vector  $\mathbf{d} = (d_1, d_2, d_3)$ . The dual quaternion representation is homogenous. The multiplication of two dual quaternions using quaternion algebra is the dual quaternion that composes two spatial transformations.

Rigid body transformations of a point can be easily represented by the dual quaternion. Assume that the homogenous coordinates of a point is  $\mathbf{p} = (p_1, p_2, p_3, p_4)$ , the dual quaternion of the point is defined by

$$\hat{\mathbf{p}} = \varepsilon (p_1 i + p_2 j + p_3 k) + p_4 \quad (3)$$

The spatial displacements  $\hat{\mathbf{p}}'$  of the point  $\mathbf{p}$  will be given by

$$\hat{\mathbf{p}}' = \bar{\hat{\mathbf{Q}}} \hat{\mathbf{p}} \hat{\mathbf{Q}}^* = \mathbf{Q} \hat{\mathbf{p}} \mathbf{Q}^* + p_4 (\mathbf{R} \mathbf{Q}^* - \mathbf{Q} \mathbf{R}^*) \varepsilon \quad (4)$$

where  $\bar{\hat{\mathbf{Q}}}$  and  $\hat{\mathbf{Q}}^*$  are the dual conjugate and the conjugate of  $\hat{\mathbf{Q}}$ , respectively. They are defined by

$$\bar{\hat{\mathbf{Q}}} = \mathbf{Q} - \varepsilon \mathbf{R} = (Q_1 - R_1) i + (Q_2 - R_2) j + (Q_3 - R_3) k + (Q_4 - R_4)$$

$$\hat{\mathbf{Q}}^* = \mathbf{Q}^* + \varepsilon \mathbf{R}^* = -(Q_1 + R_1) i - (Q_2 + R_2) j - (Q_3 + R_3) k + (Q_4 + R_4)$$

$\hat{\mathbf{p}}'$  can be transform into the matrix form as the follow equation

$$\hat{\mathbf{p}}' = ([\mathbf{H}^+] [\mathbf{H}^-] + [\mathbf{H}^-] [\mathbf{H}^{0+}] - [\mathbf{H}^+] [\mathbf{H}^{0-}]) \mathbf{p} \quad (5)$$

where  $[\mathbf{H}]$  is the matrix form according to Eq. (4). The construction of the matrix form is referred to in the Appendix.

## 3 B-Spline Motion Interpolation

### 3.1 Rational B-Spline Motion

A motion is called rational, if there is a parameter  $t \in \mathbf{R}$ , such that all point paths under the motion are rationally parameterized by the parameter t. Since

the spline with monomial basis can be converted to the spline with corresponding spline basis function, these point paths equation can be converted into a rational spline form and the motion is referred to as a rational spline motion. It has been proved that any rational motion may be represented as a rational dual quaternion curve [9]. The Rational B-spline Motion is given by

$$\hat{\mathbf{Q}}(t) = \sum_{i=0}^n N_{i,d}^K(t) \hat{\mathbf{Q}}_i \quad (6)$$

where  $\hat{\mathbf{Q}}_i$  is the the control dual quaternion and  $N_{i,d}^K$  is the normalized B-spline basic function of degree  $d$  over a knot sequence  $K$ . The knot sequence  $K$  is a non-decreasing values of knot  $k_i$ :

$$k_1, \underbrace{k_1, \dots, k_1}_{n_1=d}, \underbrace{k_2, \dots, k_2}_{0 < n_2 \leq d+1}, \dots, \underbrace{k_{r-1}, \dots, k_{r-1}}_{0 < n_{r-1} \leq d+1}, \underbrace{k_r, \dots, k_r}_{n_r=d}, k_r \quad (7)$$

where  $n_i$  is the multiplicity of  $k_i$  in  $K$  and  $n_i < d + 1$  ( $i = i, \dots, r - 1$ ).  $n$  in Eq. (6) is given by

$$n = \sum_{i=1}^r n_i - d \quad (8)$$

By substituting Eq. (6) into Eq. (4) or Eq. (5), it can be found that a B-spline dual quaternion curve of degree  $d$  describes a rational motion of degree  $2d$ .

### 3.2 Interpolation of Discrete Cutter Locations

The discrete cutter locations  $cl_j$  ( $j = 1, \dots, l$ ) can be represented by the group of tool tip point  $p_{1,j}$  and a second point on tool axis  $p_{2,j}$ . According to Eq. (1), the corresponding discrete dual quaternions  $\hat{\mathbf{q}}_j$  is determined by the tool location  $p_{1,j}$  and tool direction  $p_{2,j} - p_{1,j}$ . By interpolating the discrete dual quaternions  $\hat{\mathbf{q}}_j$ , the B-spline motion of the cutter is then obtained. The interpolation method is referred to Ref. [10].

## 4 Dual NURBS Tool Path with Equal Distance

This section addresses the rational trajectory of the point on the cutter that is undergoing a B-spline motion. Ref. [10] and [8] study the method to transform the B-spline motion into NURBS tool path. However, the obtained tool path usually consists of too many piecewise Bézier segments because a B-spline motion must be first partitioned into piecewise rational Bézier representation. Unlike existing work, the compact DNTPED is generated directly from the B-spline motion.

#### 4.1 Trajectory of a Point under a B-Spline Motion

For simplicity, the B-spline basis function in Eq. (6) is replaced with the symbol  $\Lambda$  and  $\Lambda = \sum_{i=0}^n N_{i,d}^K(t)$ . By substituting Eq. (6) into Eq. (4), the trajectory of the point  $\mathbf{p}$  is formulated as

$$\begin{aligned}\hat{\mathbf{p}}(t)' &= (\Lambda \bar{\mathbf{Q}}_i) \hat{\mathbf{p}} (\Lambda \hat{\mathbf{Q}}_i^*) \\ &= (\Lambda \mathbf{Q}_i) \hat{\mathbf{p}} (\Lambda \mathbf{Q}_i^*) + \\ &\quad p_4 ((\Lambda \mathbf{R}_i) (\Lambda \mathbf{Q}_i^*) + (\Lambda \mathbf{Q}_i) (\Lambda \mathbf{R}_i^*)) \varepsilon\end{aligned}\quad (9)$$

According to Eq. (5), Eq. (9) can also be translated into matrix form as

$$\begin{aligned}\mathbf{p}'(t) &= ((\Lambda [\mathbf{H}_i^+]) (\Lambda [\mathbf{H}_j^-]) + (\Lambda [\mathbf{H}_j^-]) (\Lambda [\mathbf{H}_i^{0+}]) \\ &\quad + (\Lambda [\mathbf{H}_i^+]) (\Lambda [\mathbf{H}_j^{0-}])) \mathbf{p}\end{aligned}\quad (10)$$

The key problem to obtain the trajectory of a point under the B-spline motion is then transformed into computing the multiplication of the two B-splines.

#### 4.2 Direct Multiplication of B-Splines

Most of existing algorithms to perform B-spline multiplication are indirect methods, such as some sampling and knot insertion strategy. Only Refs. [11] and [12] study the direct approaches for B-spline multiplication based on blossom representation of B-splines. The rational trajectory in Eq. (10) is also computed with the help of blossom approach. The outstanding feature of the blossom representation is that it provides a uniform and simple approach to computing values of a polynomial using a variety of representations from Bézier curves to NURBS.

The blossom of a polynomial  $P(u)$  of degree  $d$  is a multivariable function  $p(u_1, u_2, \dots, u_d)$ , which exists uniquely for the polynomial  $P(u)$ . The blossom  $p(u_1, u_2, \dots, u_d)$  has the symmetry, multi-affine and diagonal properties that are respectively expressed as:

$$p(u_1, \dots, v, \dots, w, \dots, u_d) = p(u_1, \dots, w, \dots, v, \dots, u_d), \quad (11)$$

$$p(u_1, \dots, \alpha v + \beta w, \dots, u_d) =$$

$$\alpha p(u_1, \dots, v, \dots, u_d) + \beta p(u_1, \dots, w, \dots, u_d), \quad (\alpha + \beta = 1), \quad (12)$$

$$P(u) = p(u, u, \dots, u). \quad (13)$$

The blossom of a B-spline function has another important property, called dual functional property:

$$P(u) = \sum_{i=0}^n N_{i,d}^K(u) p_i \quad p_j = p(k_{j+1}, k_{j+2}, \dots, k_{j+d}). \quad (14)$$

This means that the control points of a B-spline can be obtained by evaluating the corresponding blossom segment at the approximate  $d$  consecutive knots.

The two B-splines expressed in Eq. (10) have the same knot vector and the degree. Consider the two B-splines function  $E, F$  (with blossom  $e$  and  $f$ ) with the knot vector  $K$  and the degree  $d$ , the product of  $E$  and  $F$  is another B-spline function  $G$  of degree  $D = 2d$ , with the knot vector

$$R = k_1, \underbrace{k_1, \dots, k_1}_{m_1=D}, \underbrace{k_2, \dots, k_2}_{0 < m_2 \leq D+1}, \dots, \underbrace{k_{r-1}, \dots, k_{r-1}}_{0 < m_{r-1} \leq D+1}, \underbrace{k_r, \dots, k_r}_{m_r=D}, k_r \quad (15)$$

where  $m_i = n_i + d$ . By the basic properties of B-spline function, the number of control points  $m + 1$  can be calculated as

$$m + 1 = \sum_{i=1}^r (n_i + d) - 2d + 1 = n + (r - 1)d + 1 \quad (16)$$

With given B-spline control points of  $e_0, e_1, \dots, e_n$  and  $f_0, f_1, \dots, f_n$ , the control points  $g_0, g_1, \dots, g_m$  is required to computed. The B-spline function  $G$  is then obtained as a result:

$$G(t) = \sum_{i=0}^m N_{i,D}^R(t) g_i = E(t)F(t). \quad (17)$$

By the dual functional property, the control points of the production B-spline are the blossoms  $g$  of  $G$  evaluated on the  $D$  continuous knot values of Eq. (15). Now the B-spline control points  $g_i$  of B-spline function  $G(t)$  are computed as the values of the following blossoms:

$$\begin{aligned} g_0 &= g(k_1, k_2, \dots, k_D), \\ g_1 &= g(k_2, k_3, \dots, k_{D+1}), \\ &\dots \\ g_m &= g(k_{m+1}, k_{m+2}, \dots, k_{m+D}). \end{aligned} \quad (18)$$

The piecewise blossom  $g_i$  blossom of the product B-spline  $G$  is determined by the blossoms of its tow factor B-splines  $E$  and  $F$  by [11],

$$g(k_1, k_2, \dots, k_D) = \frac{\sum e(k_{i_1}, k_{i_2}, \dots, k_{i_d}) f(k_{j_1}, k_{j_2}, \dots, k_{j_d})}{\binom{D}{d}} \quad (19)$$

where the summation is taken for all index sequences  $(i_1, i_2, \dots, i_d)$  of the B-spline function  $E$  and complementary index sequences  $(j_1, j_2, \dots, j_d)$  of the B-spline function  $F$  of the set  $\{1, 2, \dots, D\}$ . The direct computation algorithm of this combinational problem lacks efficiency. The sliding windows technique in Ref. [12] can be applied to improve the computational speed.

The symmetric and multi-affine properties of blossom provide the way to compute  $e$  and  $f$  of any set of arguments. If the knot value  $u_{j+1}^*$  is expressed by the knots  $k_i$  and  $k_{i+j+1}$ , which are two neighboring elements in  $K$ :

$$u_{j+1}^* = \frac{k_{i+j+1} - u_{j+1}^*}{k_{i+j+1} - k_i} k_i + \frac{u_{j+1}^* - k_i}{k_{i+j+1} - k_i} k_{i+j+1}. \quad (20)$$

The blossom of  $f$  is transformed to

$$\begin{aligned} & f(k_{i+1}, \dots, k_{i+j}, u_{j+1}^*, u_{j+2}^*, \dots, u_d^*) \\ &= \frac{k_{i+j+1} - u_{j+1}^*}{k_{i+j+1} - k_i} f(k_i, k_{i+1}, \dots, k_{i+j}, u_{j+2}^*, \dots, u_d^*) \\ &+ \frac{u_{j+1}^* - k_i}{k_{i+j+1} - k_i} f(k_{i+1}, \dots, k_{i+j}, k_{i+j+1}, u_{j+2}^*, \dots, u_d^*) \end{aligned} \quad (21)$$

The transformation must be applied recursively. The blossom of  $f$  is evaluated when all the knot sequences in the transformed blossoms become the knot sequences of the original B-spline blossoms  $f_i$  [11].

### 4.3 Dual NURBS Tool Path with Equal Distance

With the help of direct multiplication of two B-splines, the DNTPED is directly generated from B-spline motion. The DNTPED consists of two NURBS splines that are the trajectory of the tool tip point  $p_1$  and the second point on tool axis  $p_2$ .

The trajectory of the point  $\mathbf{p}$  under the B-spline motion  $\hat{\mathbf{Q}}(t)$  can be formulated by Eqs. (10) and (17) as

$$\begin{aligned} \mathbf{p}'(t) &= ((\Lambda[\mathbf{H}_i^+]) (\Lambda[\mathbf{H}_j^-]) + (\Lambda[\mathbf{H}_j^-]) (\Lambda[\mathbf{H}_i^{0+}])) \\ &\quad + (\Lambda[\mathbf{H}_i^+]) (\Lambda[\mathbf{H}_j^{0-}]) \mathbf{p} \\ &= (\Psi[\mathbf{Hg}_i^a] + \Psi[\mathbf{Hg}_i^b] - \Psi[\mathbf{Hg}_i^c]) \mathbf{p} \\ &= \Psi([\mathbf{Hg}_i^a] + [\mathbf{Hg}_i^b] - [\mathbf{Hg}_i^c]) \mathbf{p} \end{aligned} \quad (22)$$

where

$$\begin{aligned} \Psi &= \sum_{i=0}^m N_{i,D}^R(t), \\ \Psi[\mathbf{Hg}_i^a] &= \Psi[\mathbf{H}_i^+] \Psi[\mathbf{H}_i^-], \\ \Psi[\mathbf{Hg}_i^b] &= \Psi[\mathbf{H}_i^-] \Psi[\mathbf{H}_i^{0+}], \\ \Psi[\mathbf{Hg}_i^c] &= \Psi[\mathbf{H}_i^+] \Psi[\mathbf{H}_i^{0-}]. \end{aligned}$$

The trajectory  $\mathbf{p}'$  is then represented as

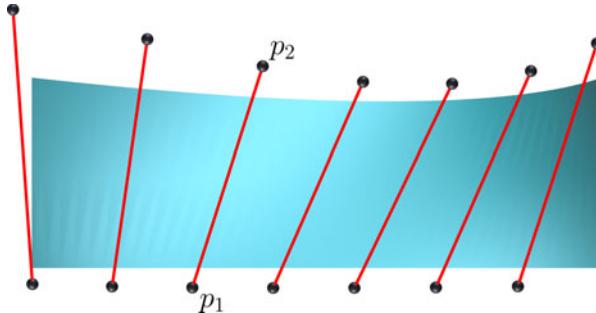
$$\mathbf{p}' = \sum_{i=0}^m N_{i,D}^R(t) ([\mathbf{Hg}_i] \mathbf{p}) \quad (23)$$

where  $[\mathbf{Hg}_i] = [\mathbf{Hg}_i^a] + [\mathbf{Hg}_i^b] + [\mathbf{Hg}_i^c]$ . Obviously, the trajectory  $\mathbf{p}'$  of the point  $\mathbf{b}$  is the form of B-spline with the knot vector  $R$ , the degree  $D$  and the control points  $[\mathbf{Hg}_i] \mathbf{p}$ .

## 5 Computational Example

Two computational examples are given to show the advantages of the proposed algorithm. The B-spline rational motion is first interpolated from a group of initial discrete cutter locations. In the interpolation, the knots is determined by the sum of near arc length of the tool tip point and the second point on the tool axis. The multiplicities of the knot values  $n_2, \dots, n_{r-1}$  are all equal to 1 except those of the first and last knot values. The existing method based on Bézier motion and  $\mathbb{R}^6$  interpolation are compared with the algorithm.

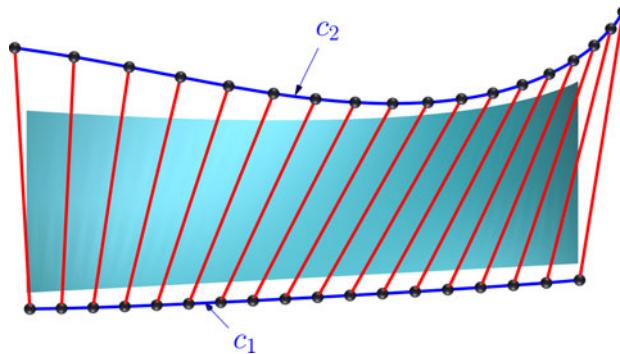
The DNTPED can also be generated by Bézier motion [4]. Like the Bézier curve, the high degree and the scarcity of local support make Bézier motion less useful than B-spline motion in tool path generation. The other method is to convert a B-spline motion into piecewise Bézier motions and transform the Bézier motions to a sequence of tool path [10]. The tool path is composed of  $r-1$  bérizer segments if the B-spline motion is expressed by Eq. (6).  $2(r-1)(2d+1)$  Bézier control points are required to describe the tool path while the number of control points of the DNTPED from B-spline motion is only  $2[r(d+1)-1]$ .  $2d(r-2)$  control points are reduced. Therefore the proposed algorithm can generate more compact tool path DNTPED. The compact tool path is convenient to treat it as one curve, not as a collection of bérizer segments.



**Fig. 1.** Initial discrete cutters

**Table 1.** Control dual quaternion points of the B-spline motion

$Q_1$	$Q_2$	$Q_3$	$Q_4$	$R_1$	$R_2$	$R_3$	$R_4$
-0.0088	0.0543	0	0.9985	25.6727	0.7333	0.9044	0.1853
-0.1091	0.0527	0	0.9967	23.7161	2.6021	0.7180	2.6440
-0.2082	0.0493	0	0.9779	20.4469	5.3360	1.2768	4.1972
-0.2621	0.0398	0	0.9643	16.5019	8.9029	2.2589	4.1353
-0.2712	0.0294	0	0.9621	13.7239	11.6231	2.8833	3.4948
-0.2416	0.0098	0	0.9708	10.1887	15.4068	3.2342	2.2609
-0.1614	-0.0161	0	0.9902	7.5967	18.5547	2.4039	1.3297
-0.0649	-0.0366	0	0.9972	5.7411	20.6486	0.6374	1.1322



**Fig. 2.** Dual NURBS tool path with equal distance

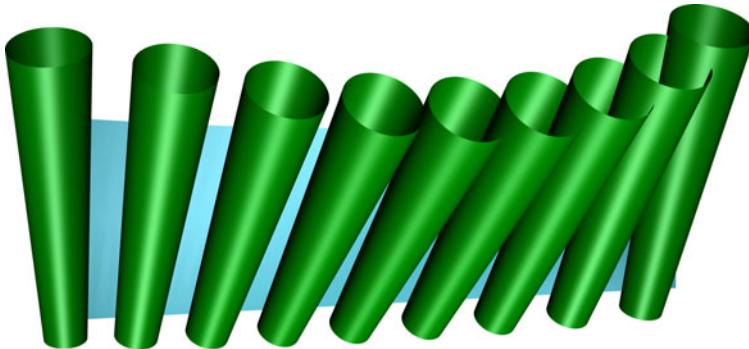
**Table 2.** Control points of the DNTPED

$c_1$				$c_2$			
$x_1$	$y_1$	$z_1$	$w_1$	$x_2$	$y_2$	$z_3$	$w_2$
51.3691	1.4601	-0.9939	1.0000	54.6213	1.9847	28.8246	1.0000
49.3982	3.2808	-1.1242	0.9990	52.5992	6.8117	28.6170	0.9990
66.8006	6.7545	-1.5974	1.3968	71.1993	14.6804	39.5792	1.3968
101.5800	14.0708	-2.8302	2.2001	108.3440	30.8922	60.9088	2.2001
178.6570	31.5962	-5.6824	4.0061	190.6700	68.7073	108.1950	4.0061
165.6370	43.7587	-5.5884	3.9976	176.9310	90.6031	104.4760	3.9976
218.0480	74.8057	-8.2844	5.5988	232.9180	148.1620	142.1270	5.5988
197.5180	74.3264	-7.6129	5.1986	210.9390	144.5950	131.0440	5.1986
207.4410	85.2895	-8.5234	5.6008	221.4940	163.4060	139.5820	5.6008
138.2330	70.6943	-5.9080	3.9991	147.4160	129.4320	98.3469	3.9991
180.7520	111.6500	-8.5398	5.5997	192.4150	196.5940	135.9310	5.5997
162.9280	108.6290	-7.9241	5.2000	173.2110	187.8660	126.0570	5.2000
170.1420	122.2910	-8.6829	5.6001	180.6450	208.1550	135.3420	5.6001
112.4490	96.5315	-6.1015	3.9998	118.8280	157.5000	97.0661	3.9998
143.6610	149.0920	-8.3694	5.6004	150.7220	232.7820	137.1460	5.6004
128.4820	143.4540	-7.7969	5.1994	134.2250	219.7960	128.1340	5.1994
133.0680	159.8740	-8.3319	5.5992	138.3970	240.7540	138.8230	5.5992
85.2494	124.2750	-5.8113	3.9987	87.2617	178.3530	101.4030	3.9987
72.9416	137.4120	-5.1881	4.0060	72.1423	184.2840	105.6110	4.0060
36.6836	79.0519	-2.8275	2.2004	35.3364	101.8690	59.3182	2.2004
21.0273	52.5875	-1.7766	1.3969	19.5175	64.6905	38.5499	1.3969
13.4313	39.2758	-1.1378	0.9985	11.8621	46.0322	28.1546	0.9985
11.5506	41.3481	-0.9889	1.0000	9.3589	45.2320	28.6778	1.0000

*Example.* The proposed interpolation algorithm is applied to generate the DNTPED for the 5-axis flank milling of a ruled surface. The ruled surface, shown in Fig. 1, is defined by a set of straight lines based on two directrices. The conical cutter with 30 mm height is adopted to machine the ruled surface. The bottom diameter and top diameter of the conical cutter are 4 mm and 10 mm, respectively. The "standard" positioning method presented in Ref. [13] is used to obtain some initial discrete cutter locations at the rulings, as shown in Fig. 2. The homogenous coordinates of the initial tool tip point and the second point on the tool axis are set as  $p_{10} = (0, 0, 0, 1)$  and  $p_{20} = (0, 0, 30, 1)$ . By interpolating the 8 cutter locations, a B-spline motion of degree 3 is obtained. The B-spline motion is represented by the knots

$$0, \underbrace{0}_{n_1=3}, \underbrace{0.285714}_{n_2=1}, \underbrace{0.428571}_{n_3=1}, \underbrace{0.571429}_{n_4=1}, \underbrace{0.714286}_{n_5=1}, \underbrace{1}_{n_6=3} \quad (24)$$

and the 8 control dual quaternion points that are listed in Table 1.



**Fig. 3.** Discrete cutters in the dual NURBS tool path with equal distance

The DNTPED is then generated by the proposed method. According to Eq. (15), the knot sequences of the two curves  $c_1$  and  $c_2$  are both given by

$$0, \underbrace{0}_{n_1=6}, \underbrace{0.285714}_{n_2=4}, \underbrace{0.428571}_{n_3=4}, \underbrace{0.571429}_{n_4=4}, \underbrace{0.714286}_{n_5=4}, \underbrace{1}_{n_6=6} \quad (25)$$

$2[r(d+1)-1] = 2[6(3+1)-1] = 46$  controls points are required to describe the DNTPED. The control points are listed in Table 2. The DNTEPD is shown in Fig. 2 and Fig. 3 illustrates the trajectory of the cutter. Siemens has developed the Oricurve function to treat the DNTPED [5]. The two NURBS curves and the fixed distance can be directly imported into the NC system. If the B-spline motion is first divided into five Bézier motions, 70 control points will be necessary to represent the two curves and 24 control points are added. Furthermore, each curve of the DNTPED will consist of 5 Bézier segments and not a whole NURBS curve.

## 6 Conclusion

The proposed interpolation algorithm is helpful to generate DNTPED for 5-axis NC machining. The rational rigid movement of a cutter is first interpolated by a B-spline curve in dual quaternion space. DNTPED is obtained directly from the B-spline dual quaternion curve with the help of the blossom form of the B-spline curve. The computational example show that the algorithm can generate more compact tool path than existing algorithms.

**Acknowledgements.** The authors acknowledge the financial support of the National Basic Research Program of China (No. 2005CB724103), the Science & Technology Commission of Shanghai Municipality (No. 10JC1408000) and the National Natural Science Foundation of China (No. 50875171).

## References

1. Shih, J., Chuang, F., et al.: One-sided offset approximation of freeform curves for interference-free NURBS machining. Computer Aided Design 40(9), 931–937 (2008)
2. Fleisig, R., Spence, A.: A constant feed and reduced angular acceleration interpolation algorithm for multi-axis machining. Computer Aided Design 33(1), 1–16 (2001)
3. Ueda, Y., Ishida, T., Takeuchi, Y.: Development of CAM System for 5-Axis NURBS Interpolated Machining(NC Data Generation by Generalized Post-Processor). Transactions of the Japan Society of Mechanical Engineers Part C 75(749), 216–222 (2009)
4. Koch, J.: Method of directing the movement of a tool as part of a process to remove material from a block of materia. US6632053 (2001)
5. Langeron, J., Duc, E., Lartigue, C., Bourdet, P.: A new format for 5-axis tool path computation, using Bspline curves. Computer Aided Design 36(12), 1219–1229 (2004)
6. Siemens: Milling with SINUMERIK 5-axis machining. Edition 05/2009 (2009)
7. Zhang, W., Zhang, Y., Ge, Q.: Interference-free tool path generation for 5-axis sculptured surface machining using rational Bezier motions of a flat-end cutter. International Journal of Production Research 43(19), 4103–4124 (2005)
8. Bi, Q., Zhu, L., Wang, Y., Ding, H.: Analytical envelope surface representation of a conical cutter undergoing rational motion. The International Journal of Advanced Manufacturing Technology 47(5), 719–730 (2010)
9. Röschel, O.: Rational motion design - a survey. Computer Aided Design 30(3), 169–178 (1998)
10. Xia, J., Ge, Q.: Kinematic approximation of ruled surfaces using NURBS motions of a cylindrical cutter. In: Proc. 2000 ASME Design Automation Conference, Baltimore, MD, Paper No. DETC2000/DAC-14280 (2000)
11. Ueda, K.: Multiplication as a general operation for splines. Curves and Surfaces in Geometric Design, 475–482 (1994)
12. Chen, X., Riesenfeld, R., Cohen, E.: An Algorithm for Direct Multiplication of B-splines. IEEE transactions on automation science and engineering 6(3), 433–442 (2009)

13. Monies, F., Rubio, W., Redonnet, J., Lagarrigue, P.: Comparative study of interference caused by different position settings of a conical milling cutter on a ruled surface. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 215(9), 1305–1317 (2001)
14. Xia, J.: Motion based geometric modeling. PhD thesis, State University of New York at Stony Brook (2001)

## Appendix A

Assuming  $\hat{\mathbf{Q}} = \mathbf{Q} + \varepsilon \mathbf{R}$ , the spatial displacement  $\mathbf{p}'$  of the point  $p$  is given by

$$\mathbf{p}' = ([\mathbf{H}^+] [\mathbf{H}^-] + [\mathbf{H}^-] [\mathbf{H}^{0+}] - [\mathbf{H}^+] [\mathbf{H}^{0-}]) \mathbf{p} \quad (26)$$

where the matrix forms of the dual quaternion are represented as follows [14].

$$\begin{aligned} [\mathbf{H}^+] &= \begin{bmatrix} Q_4 & -Q_3 & Q_2 & Q_1 \\ Q_3 & Q_4 & -Q_1 & Q_2 \\ -Q_2 & Q_1 & Q_4 & Q_3 \\ -Q_1 & -Q_2 & -Q_3 & Q_4 \end{bmatrix} \\ [\mathbf{H}^-] &= \begin{bmatrix} Q_4 & -Q_3 & Q_2 & -Q_1 \\ Q_3 & Q_4 & -Q_1 & -Q_2 \\ -Q_2 & Q_1 & Q_4 & -Q_3 \\ Q_1 & Q_2 & Q_3 & Q_4 \end{bmatrix} \\ [\mathbf{H}^{0+}] &= \begin{bmatrix} 0 & 0 & 0 & R_1 \\ 0 & 0 & 0 & R_2 \\ 0 & 0 & 0 & R_3 \\ 0 & 0 & 0 & R_4 \end{bmatrix} \\ [\mathbf{H}^{0-}] &= \begin{bmatrix} 0 & 0 & 0 & -R_1 \\ 0 & 0 & 0 & -R_2 \\ 0 & 0 & 0 & -R_3 \\ 0 & 0 & 0 & R_4 \end{bmatrix} \end{aligned}$$

# Prediction of Cutting Forces Integrated Run-Out Effect for Five-Axis Peripheral Milling with a Cylindrical Cutter

Qiang Guo, Yuwen Sun, Feifei Xu, and Zhenyuan Jia

School of Mechanical Engineering,  
Dalian University of Technology, Dalian, 116024, China  
xiands@dlut.edu.cn

**Abstract.** This paper developed a mechanistic cutting force model integrated run-out effect for five-axis peripheral milling process with a cylindrical cutter. At each cutting element along the cutter axis, a unique cutting element coordinate system is established and then cutting forces introducing cutter run-out parameters are computed in cutting element coordinate system. Subsequently, differential cutting forces are transformed into the coordinate system of workpiece. In order to obtain cutting force coefficients, peripheral milling experiments are executed, and the run-out effect is represented by two parameters that could be obtained through experimental data. At last, five-axis milling forces are predicted and experimental verification. The results show that the predicted cutting forces using the presented model with cutter run-out effect are great consistent with experimental data.

**Keywords:** Five-Axis, Run-Out Effect, Cutting Forces, Peripheral milling.

## 1 Introduction

In the process of five-axis peripheral milling ruled surfaces, attentions have been drawn onto the global optimization of cutter positions for flank milling ruled surfaces. Especially, the Minimax optimization mode of tool paths, which is proposed by Ding and Zhu [1], approximates the tool envelope surface to data points on design surface following minimum zone criterion recommended by ANSI and ISO standards based on the developed interchangeability principle. It is another significant advance after gong's LSQ optimization model [2]. On the other hand, the prediction and control of milling forces in five-axis flank milling have seldom been aroused the researchers' attentions although they counts for much in ensuring milling precision and satisfying the requirement of tight tolerances.

Up to now, the mechanistic cutting force model is one of most widely used prediction method of cutting forces, which takes cutting forces as a function of instantaneous uncut chip thickness (IUCT) and cutting force coefficients. Notable researches [3-4] in cutting force modeling mainly focus on three-axis milling processes. Only a few literatures involve five-axis cutting forces. Using Z-buffer methods, Zhu et al. [5] and

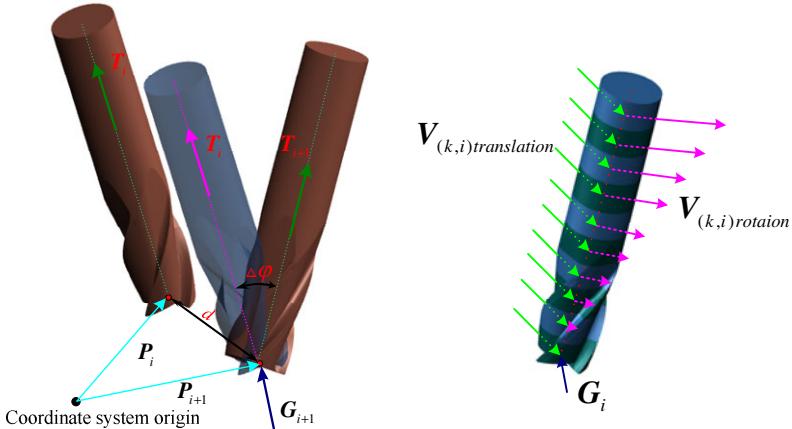
Fussell et al. [6] suggested process simulation algorithms for five-axis milling with a ball-end cutter. W. B. Ferry and Y. Altintas [7] presented a valuable and easy-computing mathematical model for predicting cutting forces in Five-Axis milling process using tapered-helical-ball-end cutters with variable pitch and serrated flutes, and it forms one of the fundamental bases of calculating five-axis milling forces. T. Bailey and M. A. [8] developed a cutting force model which took the cutting edge profile as a NURBS curve, and this method offers the advantages of representing any arbitrary cutting edge design in a genetic way. Meanwhile, the existence of cutter run-out in milling process may result in cutting forces' fluctuation of different cutting edges, and some researchers have considered this phenomenon in three-axis milling process [9-10].

This paper presents a prediction method of cutting forces integrated run-out effect for five-axis peripheral milling process, and the undeformed chip thickness is calculated considering the real moving trace of a point on the cutting edge. The predicted results are also verified by experimental data.

## 2 Instantaneous Uncut Chip Thickness Considering Runout Effect

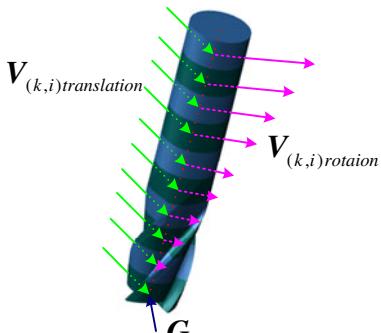
### 2.1 Feedrate in Cutting Element Coordinate System

Generally, tool paths are given in form of a series of cutting position points and corresponding vectors of cutter axis in workpiece coordinate system between two successive cutting position points ( $P_i, P_{(i+1)}$ ), the velocity of translation and rotation are considered as constants, if the feedrate is  $F$ , then the time interval between two sequential cutter locations  $P_i$  and  $P_{(i+1)}$  is known as shown in Fig. 1 That is



**Fig. 1.** The cutter motion between two positions

**Fig. 2.** The velocity of  $i_{th}$  cutting position



$$\Delta t = |\mathbf{P}_{(i+1)} - \mathbf{P}_i| / F \quad (1)$$

The rotational velocity can be further expressed as [11]

$$\omega = \Delta\varphi / \Delta t, \Delta\varphi = \arccos(\mathbf{T}_{(i+1)} \cdot \mathbf{T}_i) / |\mathbf{T}_{(i+1)}| |\mathbf{T}_i| \quad (2)$$

where the included angle between two successive cutter axes vectors is  $\Delta\varphi$ , and  $\omega$  is the rotational velocity of cutter axis. The rotational center vector  $\mathbf{G}_i$  can be evaluated using the following equation.

$$\mathbf{G}_i = (\mathbf{T}_i \times \mathbf{T}_{(i+1)}) / |\mathbf{T}_i \times \mathbf{T}_{(i+1)}| \quad (3)$$

When predicting cutting forces, the cutting location point is chosen as the rotational center. As shown in Fig. 2, at  $i_{th}$  element along the cutter axis, the linear velocity is computed as

$$\begin{cases} V_{(k,i)magrotation} = \omega(k-1/2)dz \\ \mathbf{V}_{(k,i)vectorrotation} = (\mathbf{G}_i \times \mathbf{T}_i) / |\mathbf{G}_i \times \mathbf{T}_i| \end{cases} \quad (4)$$

where  $V_{(k,i)magrotation}$  and  $\mathbf{V}_{(k,i)vectorrotation}$  are the magnitude and the direction of the liner velocity respectively.  $dz$  is the element length along the cutter axis. Then the total velocity at  $k_{th}$  cutting element of  $i_{th}$  cutting position can be evaluated as

$$\mathbf{V}_{(k,i)total} = V_{(k,i)magtrans} \mathbf{V}_{(k,i)vectortrans} + V_{(k,i)magrotation} \mathbf{V}_{(k,i)vectorrotation} \quad (5)$$

with

$$\begin{cases} V_{(k,i)magtrans} = F \\ \mathbf{V}_{(k,i)vectortrans} = (\mathbf{P}_{(i+1)} - \mathbf{P}_i) / |(\mathbf{P}_{(i+1)} - \mathbf{P}_i)| \end{cases} \quad (6)$$

In calculating the mechanics of five-axis cutting process, three-axis cutting principle is utilized. Hence, each cutting element along cutter axis should have an identical coordinate system, which is proposed in literature [7]. Take  $i_{th}$  cutting element as an instance, then the cutter element coordinate system is defined as

$$\begin{cases} \mathbf{Y}(P_i, k) = \mathbf{T}_i \times \mathbf{V}_{(k,i)total} / |\mathbf{T}_i \times \mathbf{V}_{(k,i)total}| \\ \mathbf{X}(P_i, k) = \mathbf{Y}(P_i, k) \times \mathbf{T}_i / |\mathbf{Y}(P_i, k) \times \mathbf{T}_i| \\ \mathbf{Z}(P_i, k) = \mathbf{T}_i \end{cases} \quad (7)$$

In this case, the total velocity vector can be divided into three velocities along coordinate axes, and the included angles between the total velocity and the element coordinate axes (X, Y, Z axis) can be expressed as

$$\begin{cases} \alpha = \cos^{-1}(\mathbf{V}_{(k,i)total} \cdot \mathbf{X}(P_i, k) / |\mathbf{V}_{(k,i)total}| \|\mathbf{X}(P_i, k)\|) \\ \beta = \cos^{-1}(\mathbf{V}_{(k,i)total} \cdot \mathbf{Y}(P_i, k) / |\mathbf{V}_{(k,i)total}| \|\mathbf{Y}(P_i, k)\|) \\ \gamma = \cos^{-1}(\mathbf{V}_{(k,i)total} \cdot \mathbf{Z}(P_i, k) / |\mathbf{V}_{(k,i)total}| \|\mathbf{Z}(P_i, k)\|) \end{cases} \quad (8)$$

Thus, the feedrate per second and per tooth in cutting element coordinate system can be calculated as

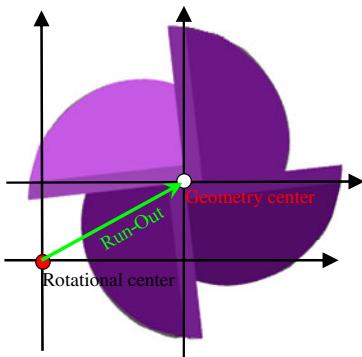
$$\begin{cases} F_{X(k,i)} = \mathbf{V}_{(k,i)total} \cos(\alpha) \\ F_{Y(k,i)} = \mathbf{V}_{(k,i)total} \cos(\beta) \text{ and} \\ F_{Z(k,i)} = \mathbf{V}_{(k,i)total} \cos(\gamma) \end{cases} \quad \begin{cases} f_{x(k,i)} = F_{X(k,i)} \phi_j / 2\pi n \\ f_{y(k,i)} = F_{Y(k,i)} \phi_j / 2\pi n = 0 \\ f_{z(k,i)} = F_{Z(k,i)} \phi_j / 2\pi n \end{cases} \quad (9)$$

The coordinate system of each cutting element along the cutter axis has the same Z axis but different x and y axes, so differential cutting forces in each element coordinate system should be converted to a united coordinate system, the cutter coordinate system is selected as the united coordinate system. The origin of cutter coordinate system is at the end of the cutter, the unit vectors can be expressed as

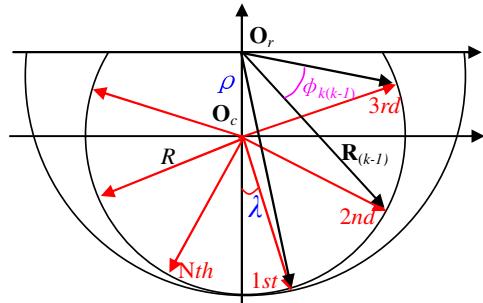
$$\begin{cases} \mathbf{Y}(P_i) = \mathbf{T}_i \times \mathbf{V}_{(0,i)translation} / |\mathbf{T}_i \times \mathbf{V}_{(0,i)translation}| \\ \mathbf{X}(P_i) = \mathbf{Y}(P_i) \times \mathbf{T}_i / |\mathbf{Y}(P_i) \times \mathbf{T}_i| \\ \mathbf{Z}(P_i) = \mathbf{T}_i \end{cases} \quad (10)$$

For the cutter coordinate system has the same Z unit vector at the  $k_{th}$  cutting element along the cutter but different X and Y unit vectors, there should have an include angle between X vector of each cutting element and X vector of the cutter coordinate system, and this include angle can be written as

$$\psi_{k,i} = \cos^{-1}(\mathbf{X}(P_i) \cdot \mathbf{X}(P_i, k) / |\mathbf{X}(P_i)| \|\mathbf{X}(P_i, k)\|) \quad (11)$$



**Fig. 3.** Departure of two centers



**Fig. 4.** Definition of Run-out parameter

## 2.2 Instantaneous Uncut Chip Thickness

Run-out is the difference between the geometrical center of cutter and the rotational center of principal axis as seen in Fig. 3 and Fig. 4. There are two parameters which can define the Run-out effect, an offset and a location angle (LA). For the existence of offset, each cutting edge's rotational radius and pitch angle are different from their standard values. Hence, the IUCT at immersion angle  $\theta$  can not be computed using the classic equation. And for the location angle, the magnitude and direction of the largest radius are various. Then assumed the cutter has N cutting edges illustrated as Fig. 4, the radius of the cutting edge can be computed and expressed as follows

$$\mathbf{R}_k = \boldsymbol{\rho} + \mathbf{R}(\lambda + (k-1) \times 2\pi/N) \quad (12)$$

where  $\mathbf{R}_k$  is the rotational radius vector of  $k_{th}$  cutting edge of cutter,  $\boldsymbol{\rho}$  is the offset vector,  $R$  is the nominal cutter radius and  $\lambda$  is the location angle, which is a constant in each cutting element along the cutter axis. LA changes with the helix angle of cutter. So for  $k_{th}$  cutting element along the cutter axis, the LA is

$$\lambda_k = \lambda_0 + ((k - 0.5) \tan idz)/R \quad (13)$$

where  $i$  is the helix angle of the cutter, and  $\lambda_0$  is the LA at the end of the cutter or the first cutting element along the cutter axis.

Run-out parameters not only affect the cutting edge's radius but also affect the pitch angles, which are different from their original values and can be evaluated as

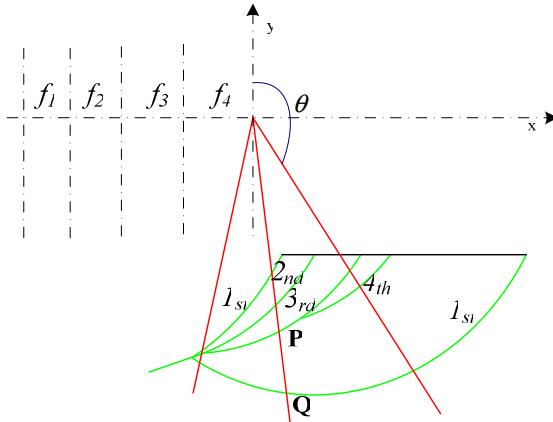
$$\phi_{(k-1)k} = \arccos \left( \mathbf{R}_{(k-1)} \cdot \mathbf{R}_k / (\|\mathbf{R}_{(k-1)}\| \cdot \|\mathbf{R}_k\|) \right) \quad (14)$$

where  $\phi_{(k-1)k}$  is the pitch angles between  $k_{th}$  and  $(k-1)_{th}$  cutting edge, and the first pitch angle is the angle between the first cutting edges and the last cutting edge. Thus, the actual feedrate per tooth of  $k_{th}$  cutting edge can be evaluated as.

$$f_k = fN\phi_{(k-1)k}/2\pi \quad (15)$$

Since IUCT is determined by actual feedrate per tooth and immersion angle, it is different from the classical one due to the existence of cutter run-out effect. Fig. 5 takes the cutter with four cutter edges as an instance. For the existence of the run-out effect,  $f_1, f_2, f_3, f_4$  are different from each other and rotational radiiuses are also changeable, the IUCT between current cutting trace and previous trace is:

$$h = \|\mathbf{Q} - \mathbf{P}\| = \sqrt{(q_x - p_x)^2 + (q_y - p_y)^2 + (q_z - p_z)^2} \quad (16)$$



**Fig. 5.** IUCT determination with cutter run-out

The above formula is usually solved by numerical algorithm. From this it can be seen that three-axis milling force model can be extended into five-axis milling process.

## 2.3 Entering Angle and Exiting Angle

In Five-axis milling process, the cutter axis vector has two tilt angles in workpiece coordinate system. Due to the existence of these two angles, different cutting element has different radial cutting depth, and thus different cutting element has different entering and existing angles.

### 2.3.1 Entering Angle Determination

In workpiece coordinate system, the surface before machining could be considered as a set of points  $\mathbf{PT}_i$  composing of a series of point vectors. The cutter axis in workpiece coordinate system is expressed as

$$\mathbf{L}_i(z) = \mathbf{P}_i + z\mathbf{T}_i \quad (17)$$

The vector on the surface of cylinder whose radius is  $R$  can be expressed in the cutting element coordinate system using the following equation.

$$\mathbf{Arc}(\eta) = R \cos(\eta)\mathbf{i} + R \sin(\eta)\mathbf{j} \quad (18)$$

Here,  $\eta$  is the including angle between the vector and the X-axis of cutting elements coordinate system. A bounding box that embraces the cutter is defined and this box also contains some points of the workpiece surface without machining. At the height of  $z$  along the cutter axis, the plane paralleling to the XOY plane of cutter coordinate system is found out. A curve belonged to this plane and the blank surface of workpiece

simultaneously is forming. A part of the curve is enclosed in the bounding box. Then the equation below is adopted to determine the entering vector.

$$\mathbf{V}_i(z) = \mathbf{P}\mathbf{T}_i - \mathbf{L}_i(z) \quad (19)$$

And the norm of this vector is taken as target to determine the entering vector which is fixed when the following equation is satisfied.

$$|\mathbf{V}_i(z) - \mathbf{R}_k| \leq \varepsilon \quad (20)$$

where  $\varepsilon$  is the preset value of accuracy control in iteration. Then the entering angle can be evaluated taking advantage of the entering vector and the Y-axis of cutter coordinate system.

$$\theta_{\text{entry}} = \cos^{-1} (\mathbf{V}_i(z) \cdot \mathbf{Y}(P_i, k) / |\mathbf{V}_i(z)| |\mathbf{Y}(P_i, k)|) \quad (21)$$

Then this angle could be transformed from cutting coordinate system into cutting element. That is

$$\theta_{\text{entry}(k,i)} = \theta_{\text{entry}} - \psi_{k,i} \quad (22)$$

Apparently, this equation is adequate for down-milling while in up-milling process this equation is changed into exiting angle.

### 2.3.2 Exiting Angle Determination

The exiting angle is decided in the coordinate system of cutting element. At two successive cutting locations the distance is evaluated as

$$d = |\mathbf{P}_i - \mathbf{P}_{(i-1)}| \quad (23)$$

Further, the interval angle between two continuous cutting edges is expressed as

$$\delta\varphi = \Delta\varphi f_{x(k,i)} / d \quad (24)$$

Then, the existing angle is evaluated using following equations.

$$a^2 (1 - \frac{R^2}{b^2} \sin^2 \alpha_k) = (f - R \cos \alpha_k)^2 \quad (25)$$

With  $\begin{cases} a = R_k / \cos(\delta\varphi \cos \vartheta) \\ b = R_k / \cos(\delta\varphi \sin \vartheta) \end{cases}$ ,  $\vartheta = \cos^{-1} \left( \frac{\mathbf{G}_i \cdot \mathbf{Y}(P_i, k)}{|\mathbf{G}_i| |\mathbf{Y}(P_i, k)|} \right)$ .

$$\theta_{\text{exit}(k,i)} = \min\{\alpha_k\} + \pi + \psi_{k,i} \quad (26)$$

### 3 The Cutting Force Model

For the  $j_{th}$  flute and  $k_{th}$  cutting element along the cutter, at  $t$  moment  $\phi$  can be calculated using the following equation

$$\phi(j, k, t) = \phi_r + \sum_{l=1}^j \psi_l - k \times dz \times \tan i / r + 2\pi nt \quad (27)$$

where  $\phi_r$  is the reference immersion angle at time  $t=0$ ,  $\psi_l$  is the flute angle of  $l_{th}$  cutter edge of cutter. Then, the cutting forces in three directions are computed adopting classical mechanical-numerical approach which considers cutting forces as the function of cutting force coefficients, IUCT and cutting element's thickness.

$$\begin{cases} df_{(j,k,T)}(\phi) = K_T(h)h_{(j,k)}(\phi)dz \\ df_{(j,k,R)}(\phi) = K_R(h)h_{(j,k)}(\phi)dz \\ df_{(j,k,A)}(\phi) = K_A(h)h_{(j,k)}(\phi)dz \end{cases} \quad (28)$$

$h_{(j,k)}(\phi)$  is IUCT of  $j_{th}$  flute and  $k_{th}$  cutting element along the cutter axis,  $df_{(j,k,x)}(\phi)$  ( $x=T, R$  and  $A$  represent the tangential, radial and axial directions, respectively) is the cutting force in  $x$  direction at the angle of  $\phi$  and  $K_x(\phi)$  ( $x=T, R$  and  $A$ ) is the cutting force coefficient which can be defined as a function of IUCT.

$$K_x(h) = a_x + b_x e^{c_x h} \quad (x = T, R \text{ and } Z) \quad (29)$$

where  $a_x$ ,  $b_x$  and  $c_x$  are constants.

Further, these forces need to be transformed into the cutter coordinate system through from element coordinate system. That is

$$\begin{aligned} & \left[ df_{(j,k,X)}(\phi) \, df_{(j,k,Y)}(\phi) \, df_{(j,k,Z)}(\phi) \right]^T \\ &= \mathbf{T}_{r(j,k)} \mathbf{T}_{r(k,i)} \left[ df_{(j,k,T)}(\phi) \, df_{(j,k,R)}(\phi) \, df_{(j,k,A)}(\phi) \right]^T \end{aligned} \quad (30)$$

where  $\mathbf{T}_{r(j,k)}$  and  $\mathbf{T}_{r(k,i)}$  are rotational matrixes

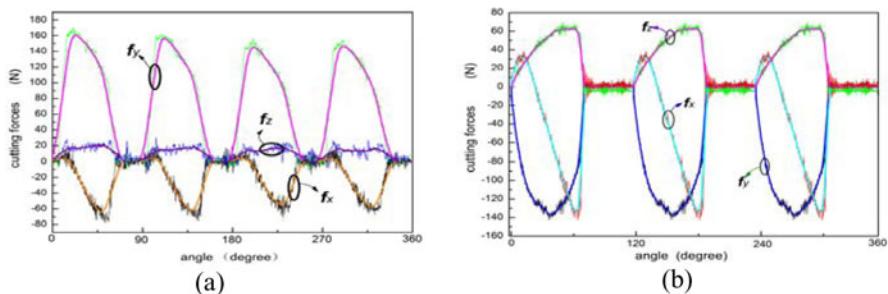
Subsequently, the total force components in three directions of cutter coordinate system can be computed by summing all cutting elements' force component.

$$f_l = \sum_{j=1}^N \sum_{k=1}^M df_{(j,k,l)}(\phi) \quad l = (X, Y, \text{and } Z) \quad (31)$$

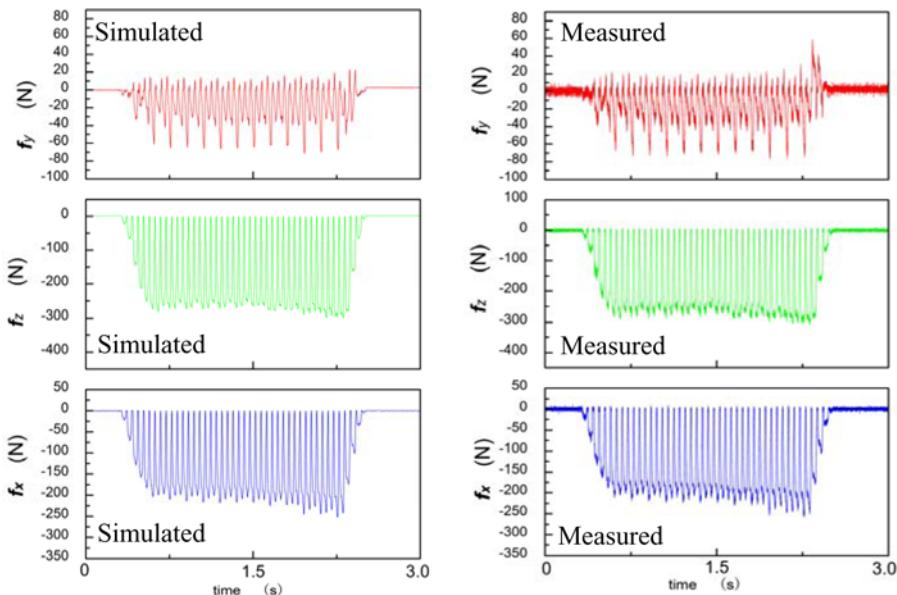
where  $M$  is the number of cutting elements along the cutter axis.

## 4 Simulations and Verifications

The proposed cutting force model integrated with cutter Run-Out effect has been experimentally verified. In experiments the flat-end milling cutter with  $30^\circ$  helix angle and the radius of 12 mm was used to mill Aluminum alloy 7075. All experiments were flank machining and the rotating speed of the principal axis was 450 RPM. The prediction result of cutting forces with cutter run-out effect is shown in Fig.6(a) with the run-out offset is 0.0062mm and the location angle is  $51.6^\circ$ . Although the measured values have some fluctuations with a low magnitude, it is normal in real milling process when the signs of cutting forces are recorded with a 30K sampling frequency. It can be seen that the simulated result has a good agreement with the measured one. Meanwhile, for further investigating the run-out effect on cutting forces, as illustrated in Fig.6(b)



**Fig. 6.** Measured and simulated cutting forces



**Fig. 7.** The magnitude and shape of measured and simulated cutting forces

the result of cutting experiment with negligible run-out is also given. The comparison results show that if the run-out exists the magnitude of cutting force for each cutting edge of the cutter is different at every rotation of cutter while the magnitudes of cutting forces without run-out keeps almost constant. Further, an inclined surface is continuously milled with a 35° tilt angle of cutter axis. The rotational speed of cutter is 450RPM and the feed rate per tooth is 0.13mm. Fig.7 shows the magnitude and shape of measured cutting forces and predicted cutting forces. From this we can see that the predicted magnitude and shape of cutting forces also have a good agreement with the measured magnitude and shape of measured cutting forces.

## 5 Conclusions

This paper presents a prediction model of cutting forces taking the cutter Run-Out parameters into account for five-axis flank milling with a cylindrical cutter. The predicted cutting forces illustrates good agreements with the experimental results both in the magnitude and shape of cutting forces. It verilates the proposed cutting force model has a good accuracy of prediction. Also, it shows that the cutter run-out has a significant effect on the magnitude of cutting force and has obvious difference of cutting forces impacting on each cutting edge of cutter. Since five-axis flank milling is very complex, more milling experiments are necessary to further improve the proposed prediction model of cutting force for five-axis milling process. Some details which the proposed model is applied on sculptured surface for five-axis milling needs further research.

## Acknowledgements

This work was supported by General Program of National Natural Science Foundation of China (Grant No. 50775023), Key Program of National Natural Science Foundation of China (Grant No. 50835001) and NCET under Grant No. NCET-8-0081.

## References

1. Ding, H., Zhu, L.M.: Global Optimization of Tool Path for Five-axis Flank Milling with a Cylindrical Cutter. *Science in China-Ser. E* 52(8), 2449–2459 (2009)
2. Gong, H., Wang, N.: Optimize tool paths of flank milling with generic cutters based on approximation using the tool envelope surface. *Computer-Aided Design* 41, 981–989 (2009)
3. Merchant, M.E.: Basic mechanics of the metal cutting process. *ASME Journal of Applied physics*, 168–175 (1944)
4. Budak, E., Altintas, Y., Armarego, E.J.A.: Prediction of Milling force coefficients form orthogonal cutting data. *ASME Journal of Engineering for industry* 118, 216–224 (1996)
5. Zhu, R., Kapoor, S.G., Devor, R.E.: Mechanistic Modeling of the Ball End Milling Process for Multi-Axis Machining of Free-Form Surfaces. *ASME Journal of Manufacture and Science Engineer* 123, 369–379 (2001)

6. Fussell, B.K., Jerard, R.B., Hemmett, J.G.: Modeling of cutting Geometry and forces for 5-Axis Sculptured Surface Machining. Computer-Aided Desig 35, 333–346 (2003)
7. Ferry, W.B., Altintas, Y.: Virtual Five-Axis Flank Milling of Jet Engine Impeller - Part I: Mechanics of Five-Axis Flank milling. ASME Journal of Manufacturing Science and Engineering 130 (2008)
8. Bailey, T., Generic, M.A.: Simulation Approach for Multi-Axis Machining, Part 1: Modeling Methodology. ASME Journal of Manufacturing Science and Engineering 124, 624–633 (2002)
9. Lee, K.Y., Kim, H.M., Park, S.S.: A run-out measuring method using modeling and simulation in four-fluted end milling. In: Journal of Materials Processing Technology, 187-188, pp. 207–211 (2007)
10. Ko, J.H., Cho, D.-W.: Determination of Cutting-Condition-Independent Coefficients and Runout parameters in Ball-End milling. International J. Adv. Manufacture Technolog. 26, 1211–1221 (2005)
11. Hu, G., Ning, W.: Analytical calculation of the envelope surface for generic milling tools directly from CL-data based on the moving frame method. Computer-Aided Desig 41, 848–855 (2009)

# Measurement Error Compensation Using Data Fusion Technique for Laser Scanner on AACMMs

Licheng Wang<sup>1</sup>, Han Ding<sup>1</sup>, and Shaojun Wang<sup>2</sup>

<sup>1</sup> State Key Lab of Digital Manufacturing Equipment and Technology of China,

Huazhong University of Science and Technology, Wuhan, 430074, P.R. China

<sup>2</sup> IET, Southeast Missouri State University, Cape Girardeau, MO, 63701, USA

wlccn@hotmail.com, dinghan@mail.hust.edu.cn, swang@semo.edu

**Abstract.** Non-contact measurement is widely used for free-form surface inspection due to its effectiveness and convenience. However, the non-contact measurement accuracy is one order less than tactile probe. In this paper, a method of error analysis and compensation for a laser scanner on Articulated Arm Coordinate Measuring Machines (AACMMs) is presented. Based on the constrained optimization algorithm, the coordinate transformation is obtained between low accuracy data and high accuracy data. The deviations between the reference points to the datum are calculated. The average deviation as an error compensation value is used to correct other point clouds for reducing the measurement system errors. Experimental results show that this method can improve the accuracy of laser scanning data on AACMMs.

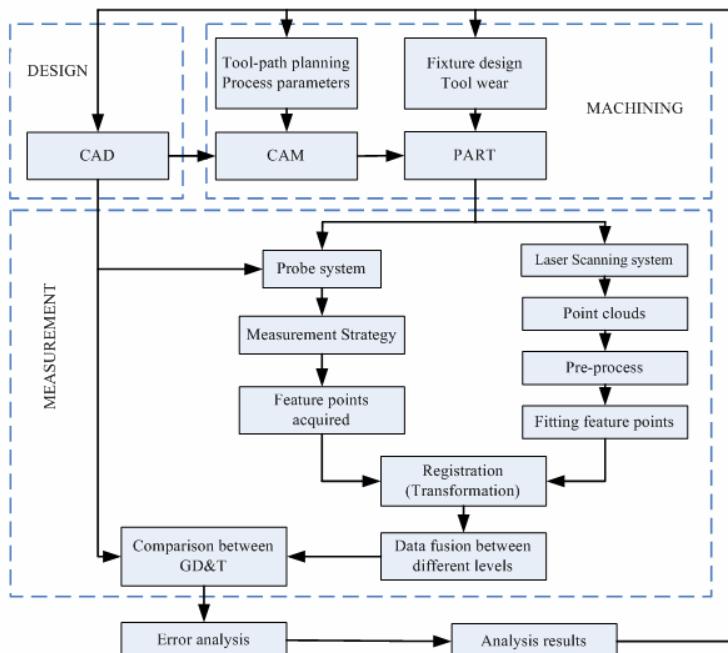
**Keywords:** Laser scanning; tactile probe; data fusion; error compensation.

## 1 Introduction

Nowadays, non-contact measurement systems for reverse engineering and inspection have been widely developed and applied. The laser scanner as a non-contact metrology device can acquire a high density of points from the target surface in relatively short time. However, laser scanners' accuracy is at least one order magnitude less compared with tactile probing, yielding many noise and redundancy data. Therefore, the combination of non-contact and contact measurements has been presented for different measurement purpose [1-3]. Sitnik [4] presented a hybrid Opto-Mechanical Measurement Machine that combines advantages of the high accuracy contact measurement by Coordinated Measuring Machine (CMM) and the high speed non-contact measurement by optical techniques. Shen [5] developed a multi-sensor coordinate measuring system with touch probe and vision sensors for rapid and high-precise inspection. The information obtained from optical sensors guide the touch probe to acquire accurate sampling data of critical surface area. Mears [6] also presented a state of art methodology by integrating CMM to the machine tool for inspection of machining operations. Research mentioned above indicates that the integration of

measurement system and machining system is gaining more and more attention in recent years.

Che and Ni [7] studied a multiple-axis laser-stripe sensor integrated Computer Numerical Control (CNC) system for 3-D measurement. To improve the scanning system accuracy, a constrained optimization calibration algorithm was developed. Contri [8] proposed a method to evaluate the uncertainty of 3D points in optical sensors. Since laser scanning system is based on the principle of optical triangulation, the laser scan depth and the projected angle have strong influence on the measurement accuracy. A standard calibration process of laser scanning system has been carried out by Feng and Xi [9]. Santolaria [10] addressed a one-step calibration method for laser line scanners, and an integrated mathematical model was presented for obtaining both intrinsic and extrinsic parameters. Gestel [11] introduced a performance evaluation test for laser line scanners on CMMs. A planar artifact was used to identify the influence of in-plane and out-of-plane angle, as well as scan depth on systematic and random errors of the laser scanner. After calibration, laser scanners on CMM systematic error can be controlled within 10 $\mu\text{m}$ . Although many high performance non-contact measurements have been developed, there are only a few researches concerning the use of non-contact sensors by integrating it to machine tools for inspection of machining operations.



**Fig. 1.** Framework of machining-measurement integration

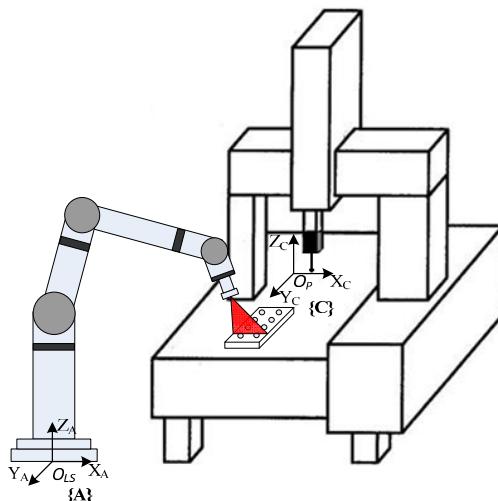
A machining-measurement integration system for manufacturing complex surface parts (Fig. 1) is under development by the research group of the authors. As shown in

Fig.1, the measured information from hybrid measurement module need to be gained rapidly and, as a feedback, be sent in time to machining process module. The focus of this paper is to study the hybrid measurement method within the machining-measurement integration system.

This paper is organized as follows: section 2 introduces a constrained optimization algorithm to obtain the transform matrix between conjugate pairs located in different frames. In section 3, experimental studies are given in details. Through the data fusion method and error analysis, the general accuracy of laser scanning data on AACMMs is improved.

## 2 Data Fusion Formulation

As shown in Fig. 2, a part is inspected by both laser scanner and tactile probe. To obtain the required measurement results, the data from different sensors need to be transformed to a common coordinate system. There are some related works of data fusion surveyed in ref.[12]. To transform points to another frame, a set of reference features to be inspected in different frames has to be indicated. These features are called conjugate pairs. Through these conjugate pairs as input values, a transformation matrix can be solved.



**Fig. 2.** Hybrid measurement using CMM and AACMM

### 2.1 A Constrained Least Square Formulation for Finding the Transformation Matrix

Assuming  $\{C\}$  is the CMM frame,  $P_C$  is points in  $\{C\}$ ;  $\{A\}$  is the AACMM frame,  $P_A$  is points in  $\{A\}$ , and transformation between  $\{C\}$  and  $\{A\}$  is  ${}^C T_A$ , the transformation from points in  $\{C\}$  to  $\{A\}$  can be given as:

$$P_C = {}^C T_A P_A \quad (1)$$

where the transformation  ${}^C T_A$  comprises two components, rotation matrix:  ${}^C R_A$  and translation vector:  ${}^C P_{A_0}$ , Eq.(1) can be written as:

$${}^C T_A = \begin{bmatrix} {}^C R_A & {}^C P_{A_0} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

or:

$$\begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix}_{CMM} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & p_1 \\ r_{21} & r_{22} & r_{23} & p_2 \\ r_{31} & r_{32} & r_{33} & p_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}_{AACMM} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad (3)$$

The problem of finding the solution of transform matrix  ${}^C T_A$  can be converted into solve a set of three simultaneous equations:

$$\begin{cases} x_{iC} = r_{11}x_{iA} + r_{12}y_{iA} + r_{13}z_{iA} + p_1 \\ y_{iC} = r_{21}x_{iA} + r_{22}y_{iA} + r_{23}z_{iA} + p_2 \\ z_{iC} = r_{31}x_{iA} + r_{32}y_{iA} + r_{33}z_{iA} + p_3 \end{cases} \quad (4)$$

In above equations, there are 12 unknown variables. However, they are not all independent. Since  ${}^C R_A$  is an orthogonal matrix, there are six constraining equations which are referred to as the compatibility constraints:

$$\left\{ \begin{array}{l} r_{11}^2 + r_{21}^2 + r_{31}^2 = r_{12}^2 + r_{22}^2 + r_{32}^2 = r_{13}^2 + r_{23}^2 + r_{33}^2 = 1 \\ r_{11} \cdot r_{12} + r_{21} \cdot r_{22} + r_{31} \cdot r_{32} = 0 \\ r_{12} \cdot r_{13} + r_{22} \cdot r_{23} + r_{32} \cdot r_{33} = 0 \\ r_{13} \cdot r_{11} + r_{23} \cdot r_{21} + r_{33} \cdot r_{31} = 0 \end{array} \right. \quad (5)$$

Therefore, there are 6 independent variables in Eq.(4). The least squares optimization method will be adopted:

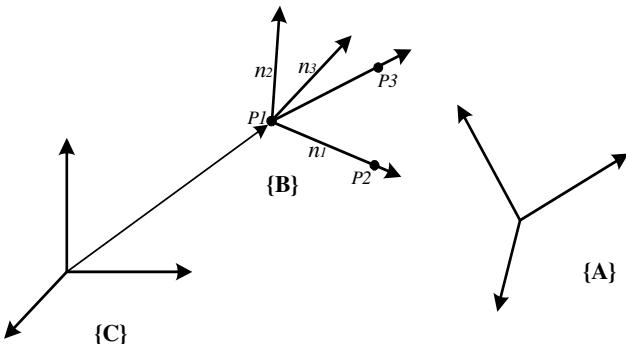
$$\min \sum_{i=1}^n \left| P_{Ci} - {}^C T_A P_{Ai} \right|^2 \quad (6)$$

*subject to: Compatibility constraints*

To solve this nonlinear optimization problem, an initial  ${}^cT_A$  need to be found first. This can be done by using three conjugate points to establish a local coordinate system and will be introduced in the following section.

## 2.2 Algorithm for Solving Initial Matrix

It is assumed that there are three non-collinear points  $P_1, P_2, P_3$  in the Cartesian frame (Fig.3). Denote  $\overrightarrow{P_2P_1}$  as  $n_1$ , vector  $n_2$  is calculated by  $\overrightarrow{P_2P_1} \otimes \overrightarrow{P_3P_1}$ , and then another vector  $n_3 = n_2 \otimes n_1$ . A local orthogonal coordinates is thus established, which is denoted by  $\{B\}$ .



**Fig. 3.** Transformation from  $\{A\}$  frame to  $\{C\}$  frame

Through the local coordinate, we can get the transformation from  $\{B\}$  to  $\{C\}$  can be expressed:

$${}^cT_B = \begin{pmatrix} {}^c n_{1B} & {}^c n_{2B} & {}^c n_{3B} & | & {}^c p_{1Bo} \\ 0 & 0 & 0 & | & 1 \end{pmatrix} \quad (7)$$

The same derivation for  ${}^AT_B$ , Eq. (8):

$${}^AT_B = \begin{pmatrix} {}^A n_{1B} & {}^A n_{2B} & {}^A n_{3B} & | & {}^A p_{1Bo} \\ 0 & 0 & 0 & | & 1 \end{pmatrix} \quad (8)$$

The matrix  ${}^AT_B$  reverse is:

$${}^BT_A = {}^AT_B^{-1} \quad (9)$$

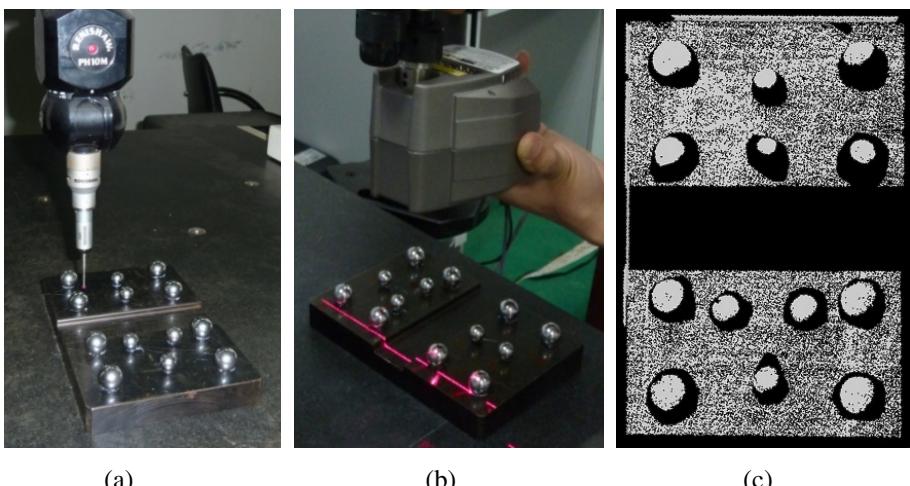
Then, the transformation matrix  ${}^C T_A$  can be calculated by:

$${}^C T_A = {}^C T_B \cdot {}^B T_A \quad (10)$$

### 3 Experiment and Results

#### 3.1 Reference Features Data Acquisition and Pre-processing

To verify the data fusion algorithm we proposed in this paper, a hybrid measurement experiment was performed. The measurand are balls attached to a plane, the diameter of the measured ball is 16mm. Ball feature is easy to be measured by CMM. Five points around spherical surface can be used for calculating the spherical center coordinate and diameter. In addition, the sphere is good for laser scanning measurement. The points on the top of the ball are used to fit the spherical center, which avoids the boundary noise influence.



**Fig. 4.** Experimental setup and point clouds processing (a)Inspection by CMM; (b)Scanning by AACMM; (c)Point clouds pre-processing

In this experiment, the balls are inspected by Renishaw TP2 tactile probe on the MISTRAL bridge-type CMM at first. And the balls are scanned by laser line scanner on the CimCore AACMM. The point clouds pre-processing and sphere fitting process were completed in ImageWare<sup>R</sup>(Fig.4). The measurement results are listed in Table 1 and Table 2, where D(C) represents diameter measured using CMM, while D(A) is the diameter measured using AACMM.

**Table 1.** Measurement results of the spherical center in CMM and AACMM coordinate

	CMM coordinate system			AACMM coordinate system		
	X coordinate	Y coordinate	Z coordinate	X coordinate	Y coordinate	Z coordinate
Ball#1	192.0876	304.7398	-543.5773	539.5306	378.2568	287.0670
Ball#2	232.3332	304.8608	-543.6082	528.3363	339.6163	287.1587
Ball#3	292.2886	304.5796	-543.7822	511.1811	282.2070	287.2348
Ball#4	332.5018	304.5609	-543.8164	499.8414	243.5322	287.4289
Ball#5	332.6681	384.7365	-543.7145	576.6983	220.9604	288.2011
Ball#6	292.5501	384.8053	-543.6743	588.0535	259.4438	288.0609
Ball#7	232.4241	384.7438	-543.4551	604.9509	317.1568	288.0817
Ball#8	192.1128	384.9507	-543.7092	616.4408	355.7035	287.5507

**Table 2.** Measurement results of the reference spheres's diamete (unit: mm)

	Ball#1	Ball#2	Ball#3	Ball#4	Ball#5	Ball#6	Ball#7	Ball#8
D (C)	16.018	16.014	16.015	16.013	16.012	16.001	16.007	15.993
D(A)	15.875	15.988	16.059	16.036	16.142	16.092	15.872	16.029

### 3.2 Error Analysis for Laser Scanner on AACMMs

From the data in Table 2, the standard deviation of the two measurements method is calculated as below:

$$\sigma_{CMM} = 0.008; \quad \sigma_{AACMM} = 0.097$$

It is obvious to find that the measurement accuracy of AACMM system is at least one order less than CMM system. There are some facts to influence the accuracy of the AACMM.

1. Laser scanning depth;
2. Projected angle(In-plan and out-plan angle);
3. Measured surface characters, such as specular, color, roughness, and materials;
4. Environment effects;
5. Operation error;
6. Data processing, fitting, transformation error.

All these facts have a comprehensive effect on the laser scanner's accuracy. Usually, calibration process can identify and compensate the laser scanner's systematic error. In this paper, the data from CMM measurement system are assumed ideal data and can be used as a reference artifact. To achieve the error analysis and compensation, the point data in AACMM coordinate system should transform to the CMM coordinate system.

### 3.3 Transformation between AACMM Coordinate System to CMM Coordinate System

Based on the measured spherical center coordinate, the coordinate transformation matrix can be calculated by Eq.(6) using three conjugate pairs. More conjugate pairs be chosen, better accuracy will be. The transformation matrix  ${}^cT_A$  is obtained as below:

$${}^cT_A = \begin{bmatrix} -0.2808 & 0.9597 & -0.0074 & 296.9347 \\ -0.9598 & -0.2808 & 0.0066 & 651.7932 \\ 0.0043 & 0.0090 & 1 & 827.0640 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11)$$

With the transformation matrix, the points in AACMM coordinate system can be transform to the CMM coordinate system by Eq.(1). The results are shown in Table 3:

**Table 3.** Results of transformation between CMM and AACMM

	CMM coordinate system			AACMM coordinate system		
	X coordinate	Y coordinate	Z coordinate	X coordinate	Y coordinate	Z coordinate
Ball#1	1k92.0876	304.7398	-543.5773	192.1029	304.7977	-543.5868
Ball#2	232.3332	304.8608	-543.6082	232.3321	304.9036	-543.6748
Ball#3	292.2886	304.5796	-543.7822	292.2485	304.5579	-543.8616
Ball#4	332.5018	304.5609	-543.8164	332.5520	304.5348	-543.8464
Ball#5	332.6681	384.7365	-543.7145	332.6370	384.6414	-543.8167
Ball#6	292.5501	384.8053	-543.6743	292.5131	384.7336	-543.7795
Ball#7	232.4241	384.7438	-543.4551	232.3780	384.7475	-543.4917
Ball#8	192.1128	384.9507	-543.7092	192.1539	384.9476	-543.8458

### 3.4 Error Compensation for Laser Scanning Data

From above discussion, the accuracy of laser scanner on AACMM is lower than that of CMM using probe. After transformation, the distance from the spherical centers coordinate of AACMM measured data to the centers of CMM measured can be obtained. It is assumed that the centers coordinate obtained by the CMM measurement system is ideal, then the single point error of the laser scanner on the AACMM can be calculated by Eq.(12):

$$d = \sqrt{(x_{iC} - {}^c x_{iA})^2 + (y_{iC} - {}^c y_{iA})^2 + (z_{iC} - {}^c z_{iA})^2} \quad (12)$$

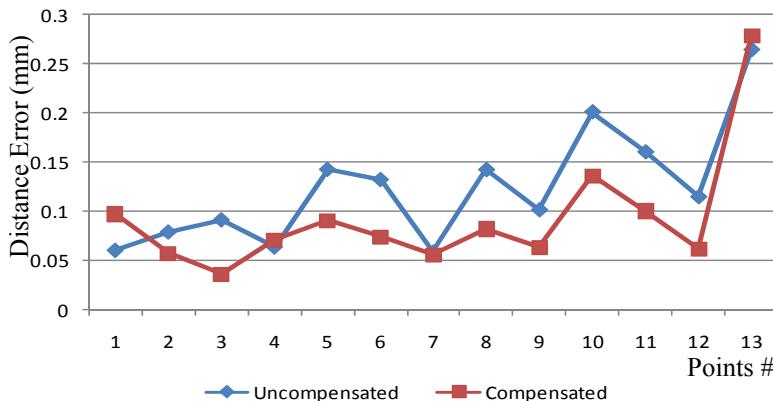
where  $(x_{iC}, y_{iC}, z_{iC})$  is the spherical center coordinate in CMM frame;  $({}^c x_{iA}, {}^c y_{iA}, {}^c z_{iA})$  is the spherical center coordinate transform to CMM frame from

AACMM frame. We then use the error obtained by equation (12) to compensate data obtained from laser scanner under the {C} frame. The compensation improve the accuracy of the data obtained from the relatively low precision measuring equipment (the laser scanner data), while the capability of scanning large quantity of data points from the laser scanning machines can be exploited. The comparison results between uncompensated and compensated are list in Table 4 and Fig. 5.

**Table 4.** Results comparison of error compensation

P		$x_c$	$y_c$	$z_c$	D
1	Uncompensated	192.0876	304.7398	-543.5773	0.0606
	Compensated	192.1090	304.8119	-543.5160	0.0970
2	Uncompensated	232.3332	304.8608	-543.6082	0.0792
	Compensated	232.3382	304.9178	-543.6040	0.0573
3	Uncompensated	292.2886	304.5796	-543.7822	0.0916
	Compensated	292.2546	304.5721	-543.7908	0.0359
4	Uncompensated	332.5018	304.5609	-543.8164	0.0640
	Compensated	332.5581	304.5490	-543.7756	0.0705
5	Uncompensated	332.6681	384.7365	-543.7145	0.1430
	Compensated	332.6431	384.6556	-543.7459	0.0904
6	Uncompensated	292.5501	384.8053	-543.6743	0.1326
	Compensated	292.5192	384.7478	-543.7087	0.0738
7	Uncompensated	232.4241	384.7438	-543.4551	0.0590
	Compensated	232.3841	384.7617	-543.4209	0.0555
8	Uncompensated	192.1128	384.9507	-543.7092	0.1427
	Compensated	192.1600	384.9618	-543.7750	0.0818
9	Uncompensated	201.6821	345.0907	-543.4327	0.1018
	Compensated	201.7282	345.1285	-543.4525	0.0628
10	Uncompensated	229.6773	345.1917	-543.3391	0.2015
	Compensated	229.7129	345.2206	-543.4671	0.1360
11	Uncompensated	296.6391	361.3481	-543.4150	0.1608
	Compensated	296.6977	361.3540	-543.4960	0.1002
12	Uncompensated	296.9463	328.4646	-543.4757	0.1151
	Compensated	296.9970	328.4574	-543.5088	0.0610
13	Uncompensated	326.6474	345.4085	-543.4085	0.2648
	Compensated	326.6874	345.6824	-543.3770	0.2785

From Table 4 and Fig. 5, it can be seen that the measurement error of laser scanner on AACMMs was reduced about 25%. As expected, the measurement systematic error has been compensated effectively, but the random error has not significant change. To improve measurement accuracy, more sampling features should be scanned.



**Fig. 5.** Efficiency of the compensation method

## 4 Conclusions

In this paper, a measured data fusion between high level accuracy coordinate and low level accuracy coordinate was presented. Based on the constrained optimization algorithm, the transformation matrix was solved using three or more conjugate pairs. In the same coordinate frame, the fitted spherical center coordinates' distance were compared. Through the error analysis and compensation, the points data obtained from laser scanner on AACMM were corrected to the high accuracy measurement data on CMM. Experimental results demonstrated that this method is effective to improve measurement correctness and accuracy for laser scanning inspection.

For free-form features, it is not possible to capture the global characteristics of the geometric feature within limited number of measuring operations using the tactile probe measurement. The proposed method exploits both advantages of the optical scanner and the tactile probe: the optical scanner can obtain large quantities of data within reasonable time, while the tactile probe provides precise positions of the measured points. Therefore, in the proposed data fusion method, data obtained from the tactile probe can serve as the reference, while the data obtained from the optical scanner captures global characteristics of the measured surface.

## Acknowledgments

This work is supported by the “973”-National Key Basic Research Program (Grant No. 2005CB724103), and the National Natural Science Foundation of China (Grant No. 50835004). The author would like to thank Tao Ye for fruitful discussions. The excellent technical assistance of Zhong Chu, Zhichao Zhao is gratefully acknowledged.

## References

1. Carbone, V., Carocci, M., Savio, E.: Combination of a Vision System and a Coordinate Measuring Machine for the Reverse Engineering of Freeform Surfaces. *International Journal of Advanced Manufacturing Technology* 17, 263–271 (2001)
2. Bradley, C., Chan, V.: A Complementary Sensor Approach to Reverse Engineering. *Journal of Manufacturing Science and Engineering* 123(1), 74–82 (2001)
3. Jamshidi, J., Owen, G.W., Mileham, A.R.: A New Data Fusion Method for Scanned Models. *Journal of Computing and Information Science in Engineering* 6(4), 340–348 (2006)
4. Sitnik, R., Śladek, J., Kupiecb, M.: New Concept of Fast Hybrid Contact and Non-contact Measurement for Automotive Industry. In: Proc. of SPIE. vol. 6198, pp. 619803-1–8 (2006)
5. Shen, T.S., Huang, J.B., Menq, C.H.: Multiple-sensor Integration for Rapid and High-precision Coordinate Metrology. *IEEE/ASME Transactions on Mechatronics* 5(2), 110–121 (2000)
6. Mears, L., Roth, J.T., Djurdjanovic, D.: Quality and Inspection of Machining Operations: CMM Integration to the Machine Tool. *Journal of Manufacturing Science and Engineering*, 131, 051006-1–13 (2009)
7. Che, C.G., Ni, J.: A Ball-target-based Extrinsic Calibration Technique for High-accuracy 3-D Metrology using Off-the-shelf Laser-stripe Sensors. *Precision Engineering* 24, 210–219 (2000)
8. Contri, A., Bourdet, P., Lartigue, C.: Quality of 3D Digitised Points with Non-contact Optical Sensors. *Annual of CIRP* 51(1), 443–446 (1999)
9. Feng, H.Y., Liu, Y., Xi, F.: Analysis of Digitizing Errors of a Laser Scanning System. *Precision Engineering* 25, 185–191 (2001)
10. Santolaria, J., Pastor, J.J., Brosed, F.J.: A One-step Intrinsic and Extrinsic Calibration Method for Laser Line Scanner Operation in Coordinate Measuring Machines. *Measurement Science and Technology*, 20, 045107-1–12 (2009)
11. Gestel, N.V., Cuypers, S., Bleys, P.: A Performance Evaluation Test for Laser Line Scanners on CMMs. *Optics and Lasers in Engineering* 47, 336–342 (2009)
12. Weckenmann, A., Jiang, X., Sommer, K.D.: Multisensor Data Fusion in Dimensional Metrology. *CIRP Annals-Manufacturing Technology* 58, 701–721 (2009)

# Evaluating Q-Learning Policies for Multi-objective Foraging Task in a Multi-agent Environment

Yogeswaran M. and Ponnambalam S.G.

School of Engineering, Monash University, Sunway campus, 46150 Petaling Jaya,  
Selangor, Malaysia

yogeswaran.mohan@eng.monash.edu.my, sgponnambalam@eng.monash.edu.my

**Abstract.** This paper evaluates the performances of the reported q-learning policies for multi-agent systems. A set of extensively used policies were identified in the open literature namely greedy,  $\epsilon$ -greedy, Boltzmann Distribution, Simulated Annealing and Probability Matching. Five agents are modeled to search and retrieve pucks back to a home location in the environment under specified constraints. A number of simulation-based experiments was conducted and based on the numerical results that was obtained, the performances of the learning policies are discussed.

## 1 Introduction

Multi-agent systems (MASs) are computational systems in which several agents, interact or work together to perform a set of tasks to satisfy a set of desired goals or objectives. MASs can be used to solve problems which are difficult or impossible for an individual agent to solve. Recently, MASs have attracted considerable attentions in various real-world applications such as hazardous waste cleanup, urban search and rescue, surveillance systems, planet exploration, construction and more. Typical problem domains for the study of MASs include foraging [1], box-pushing [2], formation forming [3], cooperative mapping [4] and soccer [5]. All of these systems consist of multiple agents acting autonomously based on their own individual decisions.

Foraging task has received the greatest attention in MASs [6] and it is a task which resembles most of the real-world applications. In this paper, a multi-objective foraging task was modeled consisting of 5 agents (Khepera II model) and 10 pucks which are placed within an environment. The agents will wander in the environment to search and retrieve pucks to the home location. Q-learning, which is a form of reinforcement learning will be used to tackle the foraging task modeled in this paper. Various q-learning policies namely greedy,  $\epsilon$ -greedy, Boltzmann Distribution (BD), Simulated Annealing (SA) and Probability Matching (PM) are adopted from the open literature to evaluate their performances. A proper learning policy is necessary to address the exploration and exploitation [7] issues that commonly arises in foraging tasks. A simple random search is

also included to justify the convergence of q-learning. The efficiency of the q-learning policies is defined by the number of pucks collected and the number of collisions made within a fixed amount of time. The modeling and the simulation experiments were completely carried out using Webots [8].

## 2 Related Work

Several attempts have been made to study the impact of the policies in q-learning. One of most recent work was by [7] where the paper examines multi-armed bandit tasks to solve the balancing problem of exploration and exploitation using policies such as  $\epsilon$ -greedy, BD, PM and adaptive pursuit method. In [9] an improved immune mechanism in q-learning and compared the results against  $\epsilon$ -greedy and BD. The immune mechanism proved to be superior to both  $\epsilon$ -greedy and BD which was tested on a simple maze problem.

Tuyls et al. [10] used the same game theory and compared the results with BD to show the convergence of q-learning in MASs. BD was also used in [11] where a new framework for MASs flocking using q-learning was modeled. Dahmani and Benyettou [12] applied a simple q-learning algorithm to solve the problem of the displacement of an agent within a 3 by 5 matrix. BD was used as the learning policy in this paper.

Some detailed analysis was done on  $\epsilon$ -greedy, BD and interval estimation policy in [13]. The authors have applied the policies in 2 different tasks specifically mountain car task and server job scheduling. Implementation of greedy and  $\epsilon$ -greedy can be seen in [14], where they described a formal and principled approach to imitation called implicit imitation between two agents. Achbany et al. [15] have presented a model integrating continual exploration and exploitation in a common framework and shown that the BD is optimal compared to  $\epsilon$ -greedy policy within this framework. Guo et al. [16] proposed SA algorithm to balance between exploration and exploitation in comparison with greedy,  $\epsilon$ -greedy and BD. Thierens [17] have compared adaptive pursuit allocation rule and PM algorithm in a non-stationary environment.

Most of the researchers implemented the policies in game theories for instance the prisoner's dilemma [18][19], rock-paper-scissors [20], pennies game [20], tic-tac-toe [21] and multi-armed bandit problems [17]. A more refined implementation of the policies can be viewed in [20][22] where 4 by 6 grids and 2 by 4 grids were used to model a soccer game between 2 agents. Based on the reviews conducted, it is clear that none of the researchers have studied the effects of the q-learning policies in a sophisticated task such as foraging. The experiments reported in the papers are sufficient to distinguish the differences in those policies but does not carry enough evidence that the policies will converge favorably towards a more sophisticated and challenging task such as foraging. Therefore, a multi-objective foraging environment is modeled and a critical study on the reported policies was conducted. Khepera II mobile robots which are equipped with 8 proximity sensors and a gripper module each is used in the model to search, retrieve and deposit the pucks at the home location in the environment.

### 3 Q-Learning

Q-learning is a form of model-free learning, where an agent does not need to have any model of the environment, instead the agent only needs to know the available states and possible actions in those states. The agent observes the environment and determines its current state  $s \in S$  and chooses an action  $a \in A$ . Once action  $a \in A$  have been executed, a reward  $r \in R$  is generated which is used to evaluate the quality of the action taken by updating the corresponding  $Q_{(s,a)}$  values. The  $Q_{(s,a)}$  values represents the sum of immediate reward  $r$  obtained by executing action  $a$  at state  $s$  and the total discounted expected future rewards obtained from new available states. Equation (1) used to update the  $Q_{(s,a)}$  values, where  $0 \leq \alpha \leq 1$  is the learning rate and  $0 \leq \gamma \leq 1$  is the discount rate.

$$Q_{(s,a)_{new}} \leftarrow Q_{(s,a)_{old}} + \alpha [r + \gamma \max_a Q_{(s',a')} - Q_{(s,a)_{old}}] \quad (1)$$

The learning rate weighs the influence of the received rewards in the learning process. A rate of 0 will make the agent not learn anything, while a rate of 1 would make the agent consider only the most recent reward. The discount factor weighs the influence of the future rewards. A rate of 0 will make the agent opportunistic by only considering current rewards, while a rate approaching 1 will make it venture for a long-term high reward. In this paper,  $\alpha$  and  $\gamma$  are adjusted to 0.8 and 0.2 respectively to obtain a marginally satisfactory result.

## 4 Policies

A policy defines the learning agent's way of behaving at a given time. The policy is a mapping from perceived states of the environment to actions to be taken when in those states. The policy is the core of a reinforcement learning agent in the sense that it alone is sufficient to determine behavior [23]. The following section presents a fundamental equation of the mentioned policies. Sensitivity studies have been carried out to determine the parameters in the reported policies. Due to page restrictions, the related data is not presented.

### 4.1 Random Search

$$a_i = \text{rand}(a_n) \quad (2)$$

An agent under random search policy will select its action  $a \in A$  without any influence of the rewards that are observed from the environment. This makes the agent to explore around in the environment with no proper directions or goals.

### 4.2 Greedy

$$a_i = \max_a Q_{(s,a)} \quad (3)$$

Following the greedy policy, the agent selects action  $a \in A$  based on the highest  $Q_{(s,a)}$  estimate of the available actions. The action considered by the greedy policy may depend on actions made so far but not on future actions.

### 4.3 $\epsilon$ -Greedy

$$a_i = \begin{cases} \text{rand}(a_n) & \text{rand}(0, 1) \leq \epsilon \\ \max_a Q_{(s,a)} & \text{otherwise} \end{cases} \quad (4)$$

The  $\epsilon$ -greedy policy is an extension of the greedy policy. The only difference is  $\epsilon$ -greedy allows a certain degree of random exploration with a probability of  $\epsilon$ . The  $\epsilon$  is a small positive value,  $0 < \epsilon < 1$  which is determined by trial and error. High values of  $\epsilon$  will force the agent to explore more frequently and as a result will prevent the agent from concentrating its choices to the optimal action, while giving the agent the ability to react rapidly to changes that takes place in the environment. Low value of  $\epsilon$  will drive the agent to choose optimal actions. For the studied case,  $\epsilon=0.2$  improved the quality of the solution marginally out of the set of tested  $\epsilon$  values.

### 4.4 Boltzmann Distribution

$$P = \frac{e^{(\max_a Q_{(s,a)}/T)}}{\sum e^{(Q_{(s,a)}/T)}} \quad (5)$$

$$a_i = \begin{cases} \max_a Q_{(s,a)} & P > x \\ \text{rand}(a_n) & \text{otherwise} \end{cases} \quad (6)$$

$$T_{new} = e^{(-dj)} T_{max} + 1 \quad (7)$$

An agent under Boltzmann distribution will choose an action based on the probability  $P$  shown in (5). High temperature causes the available actions to be equally preferable. As the temperature decays based on (7) when the iteration  $j$  increases, Boltzmann distribution becomes closer to greedy policy. For the foraging task in this paper, temperature  $T_{max}$  set to 500, decay rate  $d$  is set 0.009 and the threshold value  $x$  is set to 0.26.

### 4.5 Simulated Annealing

$$a_i = \begin{cases} \text{rand}(a_n) & \xi < e^{(\frac{\text{rand}_a Q_{(s,a)} - \max_a Q_{(s,a)}}{T})} \\ \max_a Q_{(s,a)} & \text{otherwise} \end{cases} \quad (8)$$

Following this policy, the agent selects an arbitrary action  $a \in A$  and executes the action based on the probability defined in (8) where  $\xi \in (0, 1)$ . Otherwise the agent selects and executes the optimal action.  $T$  is a positive parameter called the temperature which is set through trial and error. At the begining of the iteration, when the temperature is high, the SA algorithm policy allows more exploration. As the temperature drops based on (7), the SA algorithm policy reduces the exploration rate and drives the agent to select more optimal solutions. In this paper,  $T$  is set to 500.

## 4.6 Probability Matching

$$P_{max} = 1 - (K - 1) P_{min}, P_{min} \in (0, 1) \quad (9)$$

$$P_{a_i} = P_{min} + (1 - K \cdot P_{min}) \frac{Q_{(s,a_i)}}{\sum_{n=1}^K Q_{(s,a_n)}} \quad (10)$$

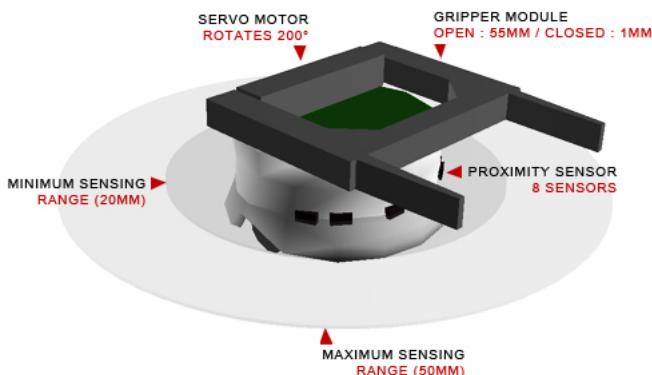
The PM rule computes each action's selection probability  $P_{a_i}$  as the proportion of the action's  $Q_{(s,a_i)}$  to the sum of all  $Q_{(s,a_n)}$  estimates. To enforce a sufficient amount of exploration, the actions exploration rate above a threshold of  $P_{min}$ . Otherwise, an action which is inefficient in the early iterations would never be considered. The exploration rate for each action is also kept below  $P_{max}$  being  $K$  the number of actions available in the observed state.  $P_{min}$  is set to 0.1 in this experiment.

## 5 Problem Environment

In this section, an integrated model representing the foraging task and q-learning is proposed and the influences of various policies mentioned in the previous sections are studied. The foraging task consists of 5 agents, an environment which is divided into few rooms by walls, a home which is located exactly in the middle of environment and 10 pucks which are placed randomly in the environment.

### 5.1 The Agent

Figure 1 shows the model of the agent. The agent is modeled strictly based on Khepera II mobile robot. It consists of 8 proximity sensors which are reflected in Fig. 1. The maximum sensing range of the proximity sensors are set to 50 mm and the minimum range to 20 mm. When an obstacle enters the minimum sensing

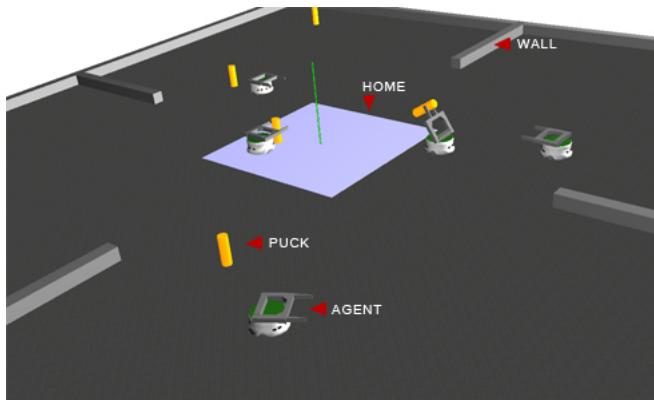


**Fig. 1.** Khepera II model

range of the agent, the agent is considered to have collided with an obstacle. There is a proximity sensor used to sense the presence of puck in the gripper module. Based on the input from the proximity sensor, the agent will react accordingly by gripping or releasing the puck from or to the desired location. The simulation model can be directly downloaded into a real Khepera II mobile robot from Webots for real time experiments.

## 5.2 The Environment

Figure 2 shows the dimension and layout of the environment. The width and length of the environment each are respectively 2000mm. The environment is separated into grid of 100 by 100 mm each resulting in 441 states for the agent to explore and exploit. Walls are located in the environment to divide the environment into four rooms and also act as a boundary for the environment. The wall also acts as obstacles for the agents. The home location is located exactly in the middle of the environment. The agent will deposit the collected pucks at the home position. Pucks are placed in the environment at predetermined locations. There are ten pucks distributed evenly in the environment.



**Fig. 2.** Environment model

## 6 Methodology

### 6.1 Objective Functions

$$F_1 = \min \frac{\sum_{i=1}^n E_i}{\sum_{i=1}^n A_i} \quad (11)$$

$$F_2 = \max \frac{\sum_{i=1}^m P_i x_i}{P} \quad (12)$$

$$F_3 = \min \frac{S_p}{S} \quad (13)$$

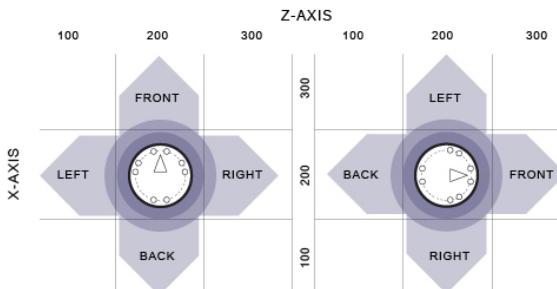
$$F_{all} = \max [w_1 (1 - F_1) + w_2 F_2 + w_3 (1 - F_3)] \quad (14)$$

The less number of collisions made over the total sum of actions taken throughout the simulation characterizes a healthier performance for the agents which can be defined via (11). Higher number of pucks collected throughout the entire simulation period characterizes a healthier performance for the agents which can be defined via (12). Less amount of time taken to retrieve all the pucks characterizes a healthier performance for the agents which is defined via (13). The agents aim are to minimize (11), maximize (12) and minimize (13). This can be achieved by using equation derived in (14). A weight factor of  $w_1$ ,  $w_2$  and  $w_3$  were used to introduce selective preferences on the objective functions. In this paper,  $w_1$  is set to 0.2,  $w_2$  is set to 0.5 and  $w_3$  is set to 0.3. In an ideal case, (14) will produce the value of 1. Total number of pucks in the environment,  $P$  is set to 10, while the sum of actions throughout the simulation period,  $A$  and the sum of collision made,  $E$  varies depending on the agent's reaction based on the simulation's conditions. Available simulation time  $S$  is set to 36000 milliseconds and  $S_p$  is the time taken to retrieve all the pucks in the environment.

## 6.2 States and Actions

Set of states  $s_i$  observed by the agents are coordinate values in the form of  $\{x_i, z_i\}$ . Referring to Fig. 3, at a given time, the agents have four available states  $s = \{(300, 200), (200, 300), (100, 200), (200, 100)\}$  and four available actions  $a = \{\text{front; back; left; right}\}$ .

Based on the policy adopted the agents will choose the desirable action  $a_i$  at the current available states  $s_i$ . The set of actions changes relatively to the agents. The  $x$ -axis and  $z$ -axis has a coordinate value ranging from  $\pm 1000$  to 0 which results in 441 states overall for the agents to explore and exploit. Q-values that are updated or reinforced are stored into a table called Q-table.



**Fig. 3.** States and actions

### 6.3 Rewards

Reward shows the desirability of the action taken by the agent towards the perceived states. If the action taken is attractive, a positive reward is given to the agent. If the action taken is not attractive then a negative reward is given to the agent. A +1.0 reward is given to the agents for picking up a puck from the environment. This triggers the agents to visit the same location again to look for more rewards. A reward of +1.0 is also given to the agents for dropping the puck at home location. The agents are given a punishment of -0.1 if the agents wander around the home without a puck in hand. This will lead the agents to escape the home location and perform search in the environment. A punishment of -1.0 is given if the agents experience any collision with the obstacles in the environment.

### 6.4 Experimental Procedures

In the model the agents are always performing one of four sub-tasks:- wander, grab, homing or deposit. The main sub-tasks are wander and homing. The Q-table for wandering and homing are separated. This is to avoid the agents from confusing the rewards received from both of the main-tasks. The Q-table<sub>wander</sub> will prompt the agents to leave the home location and wander into the environment to search for pucks. Therefore the Q-table<sub>wander</sub> will have negative rewards congregated at the home location. Once the agents encounters with pucks in the environment, the agents will pick up the pucks, updates the  $Q_{(s,a)}$  value in Q-table<sub>wander</sub> and that point forward the Q-table<sub>homing</sub> will used to lookup on the  $Q_{(s,a)}$  values. The Q-table<sub>homing</sub> will have positive rewards congregated at the home location and therefore the agents will pursue the home location to deposit the puck.

The simulation starts by setting the episode count and  $Q_{(s,a)}$  values in the both Q-tables are set to 0. In the initialization stage, the timer count is set to 0 followed by the resetting of agent's and the puck's locations to their respective coordinates in the environment. The agents and the pucks are located to their relative coordinates in the environment. Then the condition to verify whether the agent is holding a puck in its gripper is checked. Since the episode count is 0 at the beginning of the simulation, the condition will be false and therefore the agent goes into a wandering mode. All the relevant  $Q_{(s,a)}$  values are updated and referred from the Q-table<sub>wander</sub>. Then the agent will observe the available states in the environment. The agent will then select an action based on the policy given. Rewards for the action taken is calculated. The relevant  $Q_{(s,a)}$  value in the Q-table<sub>wander</sub> is updated using (14).

If 36000ms of the simulation time is reached, the episode counter will be incremented by one indicating the completion of one episode and the condition is checked for the completion of 30 episodes. Else the agents will continue to map the state-action pairs until simulation time and proceed into checking for the completion of 30 episodes. If the 30 episodes have been completed, the simulation will be terminated. Otherwise the same process will commence until the

termination condition is met. If the agent finds a puck, it will pick up the puck and a reward of +1 will be given to the agent. This reward will be updated in the Q-table<sub>wander</sub>. After this state, the agents will be using Q-table<sub>homing</sub> to update the relevant  $Q_{(s,a)}$  values until the puck is deposited at the home location. This process continues until all the pucks are collected or the termination condition is met.

## 7 Results and Discussion

The performances of the learning policies in the experiment conducted are reported in this section. The performances of the policies are evaluated based on the objective functions specified in previous sections. The simulation was carried out using Khepera II which was modeled using Webots [8]. Each of the policies is tested for 30 runs and the results are presented below. Fig. 4 shows the collision that have occurred throughout simulation period. The figure shows how the policies influence the collision occurrence in the environment as the q-learning converges for the foraging task being studied. It is observed that greedy policy have better performance compared to all the other policies studied. Fig. 5 shows the cummulative number of pucks collected by the agents throughout the entire simulation period.  $\epsilon$ -greedy policy out-performed the other policies in collecting pucks with a sum of 297 pucks. Greedy, BD, SA and PM policies performed well compared to random search. From fig. 6, it is observed that  $\epsilon$ -greedy policy out-performed the rest of the policies when measured using [4]. The greedy policy out-performed the other policies in overcoming collision in the environment. The greedy policy doesn't allow any exploration of the states once the learning process starts converging. So the agent will be forced to select the best action in every state and through that the states that are not desirable to the agent will not be explored again. This prevents the agent from experiencing the same

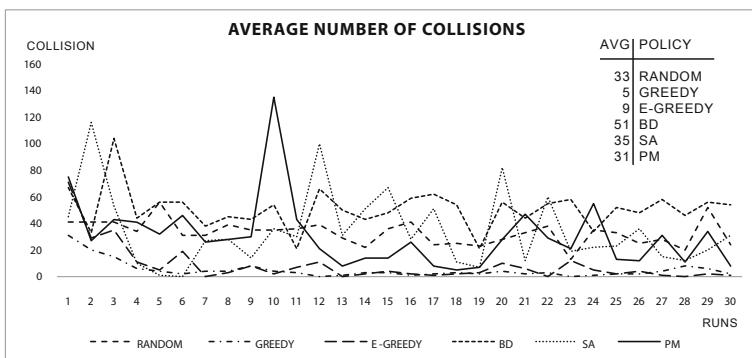


Fig. 4. Average number of collisions

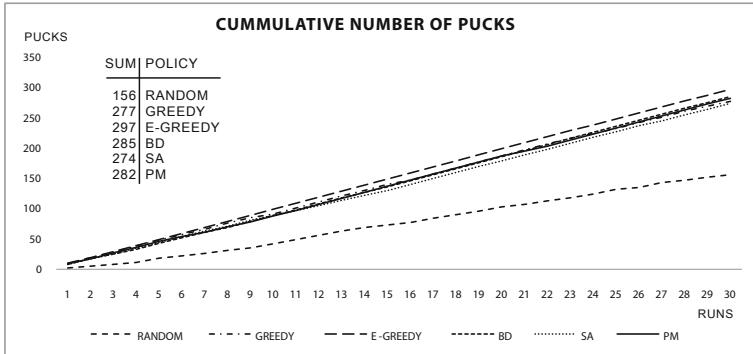


Fig. 5. Cummulative number of pucks

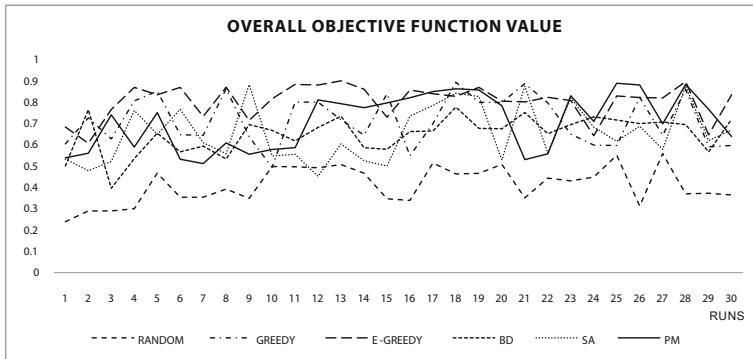


Fig. 6. Overall objective function value

collision again. Therefore greedy policy is well suited for collision avoidance. The  $\epsilon$ -greedy policy on the other hand allows the agent to explore the undesirable states with a probability of  $\epsilon$ . So at some point, even if the agent has found the optimal solution, the agent will still choose a random action  $\epsilon$  of the time. Therefore there is a probability that the collision may occur in the environment again.

BD allows exploration in the beginning of the episodes and slowly diverges to greedy policy's characteristics as the temperature drops. A low temperature means that the selection will be greedier while a high temperature means that it will be more random. The major advantage of this selection method is that the  $Q_{(s,a)}$  value is taken into account when selecting an action, meaning that more exploration will be done when the  $Q_{(s,a)}$  values are similar, and less when there is one action that has a significantly higher  $Q_{(s,a)}$  value than the others. The down side of the BD is that the policy generally requires more time to converge compared to other policies. So in this case BD is not suitable for the task due to its slow convergence. The random search, SA and PM performed equally poor compared to the policies modeled in this paper.

## 8 Conclusions

There are a variety of q-learning policies available in the open literature that works efficiently on a variety of small problem domains. But very few work have been carried out to prove that these policies scale well to larger problems. In this paper, the behavior of a q-learning agent using the reported policies namely greedy policy,  $\epsilon$ -greedy policy, BD, SA and PM have been studied in a foraging task and the results are reported. Through the experiment conducted, although all the policies carry their own advantages and disadvantages, it is clear that  $\epsilon$ -greedy policy is much more practical and effective compared to the other policies studied. It is necessary to identify the suitable policies in q-learning for different types of tasks as the suitable policies helps to accelerate the convergence of the agent's learning process. There are more directed policies available in the open literature which can be also considered for the study. The policies studied in this paper can be further improved based on the limitation and the advantages observed throughout the experiment. The simulation carried out in this paper can be easily uploaded and implemented into Khepera II mobile robot for further analysis.

## Acknowledgment

This research is funded by e-Science grant provided by Malaysia, Ministry of Science, Technology and Innovation (MOSTI), Project Number: 03-02-10-SF0036.

## References

1. Krieger, M., Billeter, J., Keller, L.: Ant-like task allocation and recruitment in cooperative robots. *Nature* 406(6799), 992–995 (2000)
2. Mataric, M., Nilsson, M., Simsarian, K.: Cooperative multi-robot box-pushing. In: *Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, pp. 556–561. Citeseer (1995)
3. Spletzer, J., Das, A., Fierro, R., Taylor, C., Kumar, V., Ostrowski, J.: Cooperative localization and control for multi-robot manipulation. In: *Proceedings IROS 2001*, Citeseer (2001)
4. Simmons, R., Apfelbaum, D., Burgard, W., Fox, D., Moors, M., Thrun, S., Younes, H.: Coordination for multi-robot exploration and mapping. In: *Proceedings of the National Conference on Artificial Intelligence*, AAAI Press, MIT Press, Menlo Park, Cambridge (1999)
5. Markus, T., Nebel, D.: CS Freiburg: coordinating robots for successful soccer playing. *IEEE Transactions on Robotics and Automation* 18(5) (2002)
6. Cao, Y., Fukunaga, A., Kahng, A.: Cooperative mobile robotics: Antecedents and directions. *Autonomous robots* 4(1), 7–27 (1997)
7. Koulouriotis, D.E., Xanthopoulos, A.: Reinforcement Learning and Evolutionary Algorithms for Non-stationary Multi-armed Bandit Problems. *Applied Mathematics and Computation* 196(2), 913–922 (2008)
8. Webots, Commercial Mobile Robot Simulation Software,  
<http://www.cyberbotics.com>

9. Ji, Z., Wu, Q., Sid-Ahmed, M.: An Improved Immune Q-learning Algorithm. In: IEEE International Conference on Systems, Man and Cybernetics, pp. 1636–1641 (2007)
10. Tuyls, K., Verbeeck, K., Lenaerts, T.: A Selection-Mutation Model for Q-learning in Multi-agent Systems. In: Proceedings of the second international joint conference on Autonomous agents and multiagent systems, pp. 693–700 (2003)
11. Morihiro, K., Isokawa, T., Nishimura, H., Matsui, N.: Emergence of Flocking Behavior Based on Reinforcement Learning. In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) KES 2006. LNCS (LNAI), vol. 4253, pp. 699–706. Springer, Heidelberg (2006)
12. Dahmani, Y., Benyettou, A.: Seek of an Optimal Way by Q-Learning. *Journal of Computer Science* 1(1), 28–30 (2005)
13. Whiteson, S., Taylor, M., Stone, P.: Empirical Studies in Action Selection with Reinforcement Learning. *Adaptive Behavior* 15(1), 33–50 (2007)
14. Price, B., Boutilier, C.: Accelerating Reinforcement Learning Through Implicit Imitation. *Journal of Artificial Intelligence Research* 19, 569–629 (2003)
15. Achbany, Y., Fouss, F., Yen, L., Pirotte, A., Saerens, M.: Tuning continual exploration in reinforcement learning: An optimality property of the Boltzmann strategy. *Neurocomputing* 71(13–15), 2507–2520 (2008)
16. Guo, M., Liu, Y., Malec, J.: A New Q-learning Algorithm Based on the Metropolis Criterion. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 34(5), 2140–2143 (2004)
17. Thierens, D.: An adaptive pursuit strategy for allocating operator probabilities. In: Proceedings of the 2005 Conference on Genetic and Evolutionary Computation, p. 1546. ACM, New York (2005)
18. Gomes, E., Kowalczyk, R.: Dynamic Analysis of Multiagent Q-learning with E-greedy Exploration. In: Proceedings of the 26th International Conference on Machine Learning, vol. 382, pp. 369–376 (2009)
19. Brafman, R., Tennenholtz, M.: Efficient Learning Equilibrium. *Artificial Intelligence* 159(1–2), 27–47 (2004)
20. Bowling, M., Veloso, M.: Rational and Convergent Learning in Stochastic Games. In: International Joint Conference on Artificial Intelligence, vol. 17, pp. 1021–1026 (2001)
21. Banerjee, B., Stone, P.: General Game Learning using Knowledge Transfer. In: The 20th International Joint Conference on Artificial Intelligence, pp. 672–677 (2007)
22. Greenwald, A., Hall, K., Serrano, R.: Correlated Q-learning. *Machine Learning* 20, 242–249 (2003)
23. Sutton, R., Barto, A.: Reinforcement Learning: An Introduction. MIT Press, Cambridge (1998)

# A Hybrid GA-AIS Heuristic for Optimization of Multipass Turning Operations

Lee Yi Zheng and S.G. Ponnambalam

School of Engineering, Monash University, Sunway campus, 46150 Petaling Jaya,  
Selangor, Malaysia  
vleeyz@yahoo.com, sgponnambalam@eng.monash.edu.my

**Abstract.** In this paper, three versions of a hybrid heuristic algorithm by combining genetic algorithm (GA) and artificial immune system (AIS) is implemented to optimize the multipass turning process which has rough machining and then a finish machining. The objective function considered is minimization of unit production cost that optimizes the machining parameters. The performance of the GA-AIS is evaluated by comparing it with the best performing heuristics reported in the literature. It is observed that the proposed GA-AIS perform better than the heuristics reported in the literature.

## 1 Introduction

Machining operations play an important role in the manufacturing of a test piece. There is more than one objective function to optimize the machining conditions. In this paper, minimizing the unit production cost is considered as the objective function. Turning operation is affected by three process parameters: the feed rate, cutting speed and depth of cut. In multipass, the machining operations are separated into two different distinct stages, the rough machining stage and the finish machining stage. Thus, in a multipass turning operation manufacturing problem, there are six parameters to be optimized in order to find the minimum unit production cost. The following lists the parameters to be optimized: Feed rate for rough and finish stage, Cutting speed for rough and finish stage, and Depth of cut for rough and finish stage.

However, for some problems, the same feed rate, cutting speed and depth of cut is used for both the rough and finish stages. At the earlier stages of research, the studies carried out had to be limited to single-pass operations as the computational power could not cope for complex mathematical calculations [1-3]. Optimization of machining parameters in multipass operations are researched by Iwata et al. [3], Ermer and Kromodihardo [4], Alberti [5], Shin and Joo [6], and Tan and Creese [7]. The local search techniques including simulated annealing algorithm [8], the genetic algorithm [9, 10], the tabu search (TS) by Glover [11] and the ant systems (AS) by Dorigo et al. [12] are applied to this problem. Lee Yi Zheng and Ponnambalam [13] proposed a particle swarm optimization (PSO) algorithm and its better performance is reported by comparing it with existing genetic algorithm (GA) [10] and simulated annealing (SA)

[14] algorithm. There is a scope of developing a hybrid search algorithm for this problem. In this paper, a hybrid of GA and AIS is proposed and its performance is reported.

## 2 Methodology

### 2.1 Notations

The notations used in the paper are explained in this section.

$UC$	unit production cost except material cost (\$/piece)
$C_M$	cutting cost by actual time in cut (\$/piece)
$C_I$	machine idle cost due to loading and unloading operations and tool idle motion time (\$/piece)
$C_R$	tool replacement cost (\$/piece)
$C_T$	tool cost (\$/piece)
$V_r$	cutting speed in rough machining (m/min)
$V_s$	cutting speed in finish machining (m/min)
$V_{rL}$	lower bound of cutting speed in rough machining (m/min)
$V_{rU}$	upper bound of cutting speed in rough machining (m/min)
$V_{sL}$	lower bound of cutting speed in finish machining (m/min)
$V_{sU}$	upper bound of cutting speed in finish machining (m/min)
$f_r$	feed rate in rough machining (mm/rev)
$f_s$	feed rate in finish machining (mm/rev)
$f_{rL}$	lower bound of feed rate in rough machining (mm/rev)
$f_{rU}$	upper bound of feed rate in rough machining (mm/rev)
$f_{sL}$	lower bound of feed rate in finish machining (mm/rev)
$f_{sU}$	upper bound of feed rate in finish machining (mm/rev)
$d_r$	depth of cut for each pass of rough machining (mm)
$d_s$	depth of cut for each pass of finish machining (mm)
$d_{rL}$	lower bound of depth of cut in rough machining (mm)
$d_{rU}$	upper bound of depth of cut in rough machining (mm)
$d_{sL}$	lower bound of depth of cut in finish machining (mm)

$d_{sU}$	upper bound of depth of cut in finish machining (mm)
$n$	number of rough cuts (an integer)
$d_t$	depth of material to be removed (mm)
$D$	diameter of work-piece (mm)
$L$	length of work-piece (mm)
$k_0$	direct labour cost + overhead (\$/min)
$k_t$	cutting cost (\$/edge)
$t_{mr}$	rough machining time (min)
$t_{ms}$	finish machining time (min)
$t_m$	actual machining time (min)
$t_c$	preparation time for loading/unloading (min)
$t_v$	idle tool motion time such as tool travel and tool approach/departure time (min)
$t_i$	total machine idle time (min)
$t_e$	tool exchange time (min)
$t_r$	tool replacement time (min)
$h_1$	constant relating to tool travel time (min)
$h_2$	constant relating to approach/departure time (min)
$T$	tool life (min)
$T_r$	expected tool life for rough machining (min)
$T_s$	expected tool life for finish machining (min)
$T_p$	tool life of weighted combination of $T_r$ and $T_s$ (min), a weight for $T_p$ [0 1]
$T_U$	upper bound for tool life (min)
$T_L$	lower bound for tool life (min)
$\alpha$	constant for tool-life equation
$\beta$	constant for tool-life equation
$\gamma$	constant for tool-life equation
$C$	constant for tool-life equation

$p$	$p = 1/\alpha$
$r$	$r = \gamma/\alpha$
$C_0$	$C_0 = C^{1/\alpha}$
$SR$	maximum allowable surface roughness (mm)
$R$	nose radius of cutting tool (mm)
$F_r$	cutting force during rough machining (kgf)
$F_s$	cutting force during finish machining (kgf)
$F_U$	maximum allowable cutting force (kgf)
$k_1$	constant of cutting force equation
$\mu$	constant of cutting force equation
$v$	constant of cutting force equation
$P_r$	cutting power during rough machining (kW)
$P_s$	cutting power during finish machining (kW)
$P_U$	maximum allowable cutting power (kW) efficiency
$\lambda$	constant related to expression of stable cutting region
$\nu$	constant related to expression of stable cutting region
$SC$	limit of stable region cutting constraint
$Q_r$	chip-tool interface rough machining temperature (oC)
$Q_s$	chip-tool interface finish machining temperature (oC)
$Q_U$	maximum allowable chip-tool interface temperature (oC)
$k_2$	constant related to equation of chip-tool interface temperature
$\tau$	constant related to equation of chip-tool interface temperature
$\phi$	constant related to equation of chip-tool interface temperature
$\delta$	constant related to equation of chip-tool interface temperature
$X$	$\{x_1, x_2, \dots, x_m\}$ machining parameter set
$n_i$	possible value of $n$ under a given $X$
$d_{ri}$	possible value of $d_r$ under a given $X$
$N_L$	lower bound of $n$
$N_U$	upper bound of $n$

$k_3$	constant for roughing and finish parameter relations, $k_3 \geq 1$
$k_4$	constant for roughing and finish parameter relations, $k_4 \geq 1$
$k_5$	constant for roughing and finish parameter relations, $k_5 \geq 1$
$\phi_1$	constant used for velocity update. Determines the weightage of particle best in the update.
$\phi_2$	constant used for velocity update. Determines the weightage of global best in the update.
$P_{xover(out)}$	constant used for outer GA loop.
$P_{mut(out)}$	constant used for outer GA loop.
$P_{xover(in)}$	constant used for inner GA loop.
$P_{mut(in)}$	constant used for inner GA loop.

## 2.2 List of Constants

The list of constants used in this paper is given in figure 1. These are used by Chen and Tsai [14].

## 2.3 Objective Functions

In this paper, the objective function chosen is the unit production cost that is to be minimized. The unit production cost consists of [6] Cutting cost by actual time in cutting operation, Machine idle cost due to loading and unloading operations and idle tool motion, Cost for tool replacement, and Tool cost. Both Chen and Tsai [14] and Onwubolu and Kumalo [10] has adopted the unit production cost calculation proposed by Shin and Joo [6] in their respective study. The first of the four cost components of the unit production cost is the cutting cost. The cutting cost is divided into multipass rough machining and a finish machining [6]. The cutting cost is therefore expressed as:

$$C_M = k_0 \left[ \frac{\pi D L}{1000 V_r f_r} \left( \frac{d_t - d_s}{d_r} \right) + \frac{\pi D L}{1000 V_s f_s} \right] \quad (1)$$

The second component is the machine idle cost. The machine idle cost is highly dependent on the machine idle time. This idle time consists of a constant term which is due to the loading and unloading operations, and a variable term which is due to the idle tool motion [6] and is expressed as:

$$C_I = k_0 \left[ t_c + (h_1 L + h_2) \left( \frac{d_t - d_s}{d_r} + 1 \right) \right] \quad (2)$$

---

$D = 50 \text{ mm}$	$L = 300 \text{ mm}$	$d_t = 6 \text{ mm}$
$V_{rv} = 500 \text{ m/min}$	$V_{rl} = 50 \text{ m/min}$	$V_{sv} = 500 \text{ m/min}$
$f_{rv} = 0.9 \text{ mm/rev}$	$f_{rl} = 0.1 \text{ mm/rev}$	$f_{sv} = 0.9 \text{ mm/rev}$
$d_{rv} = 3.0 \text{ mm}$	$d_{rl} = 1.0 \text{ mm}$	$d_{sv} = 3.0 \text{ mm}$
$T_{rl} = 25 \text{ min}$	$T_v = 45 \text{ min}$	$F_v = 5.0 \text{ kgf}$
$SR_v = 10 \mu\text{m}$	$Q_v = 1000^\circ\text{C}$	$SC = 140$
$k_1 = 108$	$k_2 = 132$	$k_3 = 1.0$
$k_5 = 1.0$	$p = 5$	$q = 1.75$
$\mu = 0.75$	$\nu = 0.95$	$\eta = 0.85$
$\nu = -1$	$\tau = 0.4$	$\phi = 0.2$
$k_0 = 0.5 \text{ \$/min}$	$k_t = 2.5 \text{ \$/min}$	$k_1 = 7 \times 10^{-4}$
$t_e = 0.75 \text{ min/piece}$	$t_e = 1.5 \text{ min/edge}$	

---

**Fig. 1.** List of constants used

The third component is the tool replacement cost. In order to obtain the cost, a tool-life equation is necessary, and in this problem, the Taylor tool-life equation proposed by Armarego and Brown [15] will be used. The equation is expressed as:

$$T = \frac{C_0}{V^p f^q d^r} \quad (3)$$

Since the rough machining and finish machining stages have different machining conditions, the wear rate of tools differ between the stages. Thus, the tool life can be expressed in the following way[6]:

$$T_p = \theta T_r + (1 - \theta) T_s \quad (4)$$

where  $T_r = \frac{C_0}{V_r^p f_r^q d_r^r}$  (5)

and  $T_s = \frac{C_0}{V_s^p f_s^q d_s^r}$  (6)

Therefore, the tool replacement cost is expressed as [6]:

$$C_R = k_0 \frac{t_e}{T_p} \left[ \frac{\pi DL}{1000V_r f_r} \left( \frac{d_t - d_s}{d_r} \right) + \frac{\pi DL}{1000V_s f_s} \right] \quad (7)$$

The fourth component, the tool cost, is expressed as the following [6]:

$$C_T = \frac{k_t}{T_p} \left[ \frac{\pi DL}{1000V_r f_r} \left( \frac{d_t - d_s}{d_r} \right) + \frac{\pi DL}{1000V_s f_s} \right] \quad (8)$$

Finally, by adding all four components (equations (1), (2), (7) and (8)), the unit production cost, UC, can be expressed as:

$$\begin{aligned} UC = & k_0 \left[ \frac{\pi DL}{1000V_r f_r} \left( \frac{d_t - d_s}{d_r} \right) + \frac{\pi DL}{1000V_s f_s} \right] + k_0 \left[ t_c + (h_1 L + h_2) \left( \frac{d_t - d_s}{d_r} + 1 \right) \right] + \\ & k_0 \frac{t_e}{T_p} \left[ \frac{\pi DL}{1000V_r f_r} \left( \frac{d_t - d_s}{d_r} \right) + \frac{\pi DL}{1000V_s f_s} \right] + k_t \frac{\pi DL}{1000V_r f_r} \left( \frac{d_t - d_s}{d_r} \right) + \frac{\pi DL}{1000V_s f_s} \end{aligned} \quad (9)$$

## 2.4 Objective Functions

The parameters to be optimized are cutting speed, feed rate and depth of cut for both rough machining and finish machining and they are constrained within boundaries so that the optimal parameter values obtained are achievable in real life machining. The number of passes will be fixed and the rest of the constraints will be altered to correspond to the fixed number of passes. The cutting force constraints and the power constraint equations will be adapted from Shin and Joo [6] and are expressed as:

$$F = k_1 f^\mu d^\nu \quad (10)$$

$$P = \frac{FV}{6120\eta} \quad (11)$$

The chip-tool interface constraint will follow the formulation of Hati and Rao [16] and is expressed as:

$$Q = k_2 V^\tau f^\phi d^\delta \quad (12)$$

The stable cutting region constraint equation and the surface finish constraint equation follows the formulation of Narang and Fischer [17] and are expressed as:

$$SC = V^\lambda f d^\nu \quad (13)$$

$$SR = \frac{f^2}{8R} \quad (14)$$

Thus, the equations of constraints for the optimization model are expressed in a list below. For the rough machining constraints:

$$V_{rL} \leq V_r \leq V_{rU} \quad (15)$$

$$f_{rL} \leq f_r \leq f_{rU} \quad (16)$$

$$d_{rL} \leq d_r \leq d_{rU} \quad (17)$$

$$T_L \leq T_r \leq T_U \quad (18)$$

$$F_r \leq F_U \quad (19)$$

$$P_r \leq P_U \quad (20)$$

$$Q_r \leq Q_U \quad (21)$$

$$SC_r \leq SC \quad (22)$$

For the finish machining constraints:

$$V_{sL} \leq V_s \leq V_{sU} \quad (23)$$

$$f_{sL} \leq f_s \leq f_{sU} \quad (24)$$

$$d_{sL} \leq d_s \leq d_{sU} \quad (25)$$

$$T_L \leq T_s \leq T_U \quad (26)$$

$$F_s \leq F_U \quad (27)$$

$$P_s \leq P_U \quad (28)$$

$$Q_s \leq Q_U \quad (29)$$

$$SC_s \leq SC \quad (30)$$

$$SR_s \leq SR_U \quad (31)$$

For the rough machining and finish machining parameter relationship constraints:

$$V_s \geq k_3 V_r \quad (32)$$

$$f_r \geq k_4 f_s \quad (33)$$

$$d_r \geq k_5 d_s \quad (34)$$

### 3 Hybrid GA-AIS Algorithm

#### 3.1 Input Representation

The input parameters for GA-AIS are represented in binary strings. Each parameter is represented as a 10-bit binary string, and one set of data is a string of 60 bit characters. An example of the input string generated is shown in figure 2.

#### 3.2 Input Acquisition

The input acquisition method applied is similar to that applied by Onwubolu and Kumalo [10]. Each input parameter is represented by a 10 bit binary number.

**Fig. 2.** Binary representation of the input string

However, binary number is of no use as decimal value of the same number will be preferred. Thus, a conversion method needs to be implemented in order to convert the 10 bit binary number into a decimal number. Furthermore, the conversion method must ensure that the decimal obtained is within the range of interest as other values will be of no use. Thus, the 10 bit binary number should be within range of consideration. This means that 0000000000 should represent the lower boundary and 1111111111 should represent the upper boundary. In the method used in this paper, the 10 bit binary number is first converted to its corresponding decimal value. For instance, a 0000000100 will convert to a decimal value of 4. Then, the obtained decimal value is being converted to another decimal value within range according to the following formula:

$$\text{PreferredDecimal} = \frac{\text{ConvertedDecimal} \times (\text{upbound} - \text{lowbound})}{2^{10} - 1} + \text{lowbound} \quad (35)$$

An example of a conversion of a value for cutting speed will be provided below.

Binary Bit representation = 1011001010

Converted Decimal = 714,      Upper boundary = 500,    Lower boundary = 50

$$\Pr[eferredDecimal] = \frac{714 \times (500 - 50)}{2^{10} - 1} + 50 = 364.07$$

Thus, the binary bit 1011001010 represents a cutting speed of 364.07 m/min. This method is used to convert the binary string into decimal values representing the cutting speed, feed rate and depth of cut for both rough machining and finish machining.

### 3.3 The Proposed GA-AIS

The GA-AIS proposed by Coello and Cortes [18] is adopted with modifications to suit to this problem. The details of the proposed algorithm are presented in this section. There are two Genetic Algorithm loops, the inner GA and the outer GA loop in the GA-AIS. In this paper, three versions of GA-AIS implementation are implemented. First version is with no selection operators at the outer GA loop. The second version is the direct implementation of Coello and Cortes's [18] method in which a binary tournament selection operator is included. The third version is with a roulette wheel selection operator instead of the binary tournament selection. The results obtained from all three implementations are compared and the best GA-AIS implementation is used to compare with the other best performing heuristics reported in the literature.

### Outer GA loop

Step 1: A population of 20 chromosomes of 60 bit binary strings is randomly generated.

### Inner GA loop

Step 2: Each chromosome is checked with constraint equations. If the chromosome fulfills all the constraint equation, it is categorized as an antigen. Else, it is categorized as an antibody.

Step 3: An antigen is randomly selected from the antigen pool and a group of 5 antibodies are randomly selected from the antibody pool.

Step 4: Each antibody selected will be compared with the antigen selected. The matching function of each antibody is calculated by the following formula.

$$Z = \sum_{i=1}^{60} (\text{antibody}(i) = \text{antigen}(i))$$

The matching function is the total number of times in which the i-th bit position of both the antibody and antigen strain are the same.

Step 5: The antibody with the highest matching function will have its matching function value updated whilst the rest will not be updated.

Step 6: The matching function of the pool of antibody (not the selected ones only) is checked. If it converges, the inner GA loop ends here.

Step 7: The termination criterion (number of generations) is checked. If it is fulfilled, the inner GA loop ends here.

Step 8: Inner loop crossover and mutation are performed.

Step 9: Steps 2 to 8 are repeated until the inner GA loop ends.

### Outer GA loop

Step 10: The population is updated.

Step 11: Each chromosome is checked using the constraint equations. If it fulfills the constraints, the chromosome is continued on to steps 12 and 13. Else, the chromosome continues at step 14.

Step 12: The objective function is evaluated. If it is the best for this current generation, the value is stored.

Step 13: The best obtained value from this current generation is compared with those of the previous generation. If it is better, the best solution obtained from this current generation is stored.

Step 14: The termination criterion (number of generations) is checked. If it is not fulfilled, steps 2 to 13 are repeated. If it is fulfilled, the outer GA loop ends here.

Step 15: If the algorithm has a selection operator, selection is carried out. Then, outer loop crossover and mutation are performed.

Step 16: The best obtained solution is printed if the termination criterion is fulfilled.

This algorithm has a very high tendency towards localization. Thus, crossover and mutation are performed twice, once in the inner GA loop, and once at the outer GA loop. The AIS is implemented solely on the selection basis.

### 3.4 Parameters Used

Several parameter values are vital in the implementation of the algorithms developed and the values used in this paper are listed below.

$$P_{xover(out)} = 0.8 \quad P_{mut(out)} = 0.1 \quad P_{xover(in)} = 0.8 \quad P_{mut(in)} = 0.1$$

Number of generations = 50 Population size = 20

Selection = no selection, Binary Tournament and Roulette wheel

Crossover = Single point operator, Mutation = unary operator

## 4 Results and Discussion

Table 1 shows the average (of 100 runs) Unit Production Cost obtained for PSO and all three versions of GA-AIS. Each sample consists of 100 runs and the shown UC value is an average value obtained from 100 runs. From table 2, it can be seen that the GA-AIS with binary tournament selection performs best among the three versions.

**Table 1.** Unit Cost (UC) obtained for PSO and GA-AIS (all three implementations) for 100 runs

Sample	UC for GA-AIS (no selection)	UC for GA-AIS (binary tournament selection)	UC for GA-AIS (roulette wheel selection)
1	1.89856	1.57657	2.44129
2	1.86718	1.59093	2.46096
3	1.89551	1.56815	2.39288
4	1.88277	1.57907	2.40319
5	1.89567	1.58072	2.41840
6	1.89021	1.58260	2.46171
7	1.90383	1.55289	2.37578
8	1.89387	1.57108	2.39393
9	1.89656	1.57195	2.43183
10	1.89876	1.56826	2.42365
Average	1.89229	1.57422	2.42036

**Table 2.** Comparison of unit cost obtained using GA, SA, ACO, FEGA, PSO and GA-AIS(binary tournament selection)

Sample	GA	SA	ACO	FEWA	PSO	GA-AIS
1	1.761	2.2971	1.6262	2.3065	1.65849	1.57657
2	1.761	2.2939	1.6262	2.3058	1.64863	1.59093
3	1.761	2.2960	1.6262	2.3028	1.63443	1.56815
4	1.761	2.2969	1.6262	2.3022	1.64549	1.57907
5	1.761	2.2956	1.6262	2.3057	1.63711	1.58072
6	1.761	2.2965	1.6262	2.3096	1.64573	1.58260
7	1.761	2.2973	1.6262	2.3058	1.65000	1.55289
8	1.761	2.2922	1.6262	2.3013	1.63934	1.57108
9	1.761	2.2981	1.6262	2.3085	1.65141	1.57195
10	1.761	2.2959	1.6262	2.3091	1.61704	1.56826
Average	1.761	2.2960	1.6262	2.3057	1.64277	1.57422

Thus, GA-AIS with the binary tournament selection is used for comparison with existing heuristics reported in the literature. The heuristics considered for evaluation are GA by Onwubolu and Kumalo [10], SA by Chen and Tsai [14], ACO by Vijayakumar et al., [19] and FEGA by Chen and Chen [20]. From table 2, it is clear that GA-AIS is performing better than the other heuristics that are considered for evaluation.

## 5 Conclusion

In this paper, three versions of GA-AIS heuristic are proposed and their performances are evaluated with existing heuristics reported in the literature. GA-AIS has proven to be the best heuristic for the stated problem as the returned average UC value as shown in Table 2 is the lowest obtained from all the heuristics that are compared.

## References

1. Ermer, D.S., Wu, S.M.: The effective experimental error on the determination of the optimum metal-cutting conditions. *Journal of Engineering for Industry, Transaction ASME, Series B* 89, 315–322 (1967)
2. Ermer, D.S., Morris, S.M.: A treatment of error of estimation in determining optimum machining conditions. *International Journal of Machine Tool Design and Research* 9, 357–362 (1969)
3. Iwata, K., Murotsu, Y., Iwatubo, T., Oba, F.: Optimization of cutting conditions for multi-pass operations considering probabilistic nature in machining conditions. *Journal of Engineering for Industry, Transaction ASME, Series B* 76, 211–217 (1977)
4. Ermer, D.S., Kromodihardo, S.: Optimization of multi-pass turning with constraints. *Journal of Engineering for Industry, Transaction ASME* 103, 462–468 (1981)
5. Alberti, N.: Optimization of multi-pass turning. *Transactions of the 14th NAMRC* (1986)
6. Shin, Y.C., Joo, Y.S.: Optimization of machining conditions with practical constraints. *International Journal of Production Research* 30, 2907–2919 (1992)
7. Tan, F.P., Creese, R.C.: A generalized multi-pass machining model for machining parameter selection in turning. *International Journal of Production Research* 33, 1467–1487 (1995)
8. Kirkpatrick, S., Gelatt Jr., C.C., Vecchi, M.P.: Optimization by simulated annealing. *Science* 220, 671–680 (1983)
9. Goldberg, D.E.: *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading (1989)
10. Onwubolu, G.C., Kumalo, T.: Optimization of multipass turning operations with genetic algorithms. *International Journal of Production Research* 39, 3727–3745 (2001)
11. Glover, F.: Tabu search - Part I. *ORSA Journal on Computing* 2, 4–32 (1989)
12. Dorigo, M., Maniezzo, V., Colomi, A.: Ant system: optimization by a colony of cooperating agents. *IEEE Transaction on Systems, Man, and Cybernetics, Part B: Cybernetics* 26, 29–41 (1996)
13. Zheng, L.Y., Ponnambalam, S.G.: Optimization of multipass turning operations using particle swarm optimization. In: *Proceeding of the 7th International Symposium on Mechatronics and its Applications (ISMA 2010)*, Sharjah, UAE (April 20-22, 2010)
14. Chen, M.-C., Tsai, D.-M.: A simulated annealing approach for optimization of multi-pass turning operations. *International Journal of Production Research* 34, 2803–2825 (1996)

15. Armarego, E.J.A., Brown, R.H.: *The Machining of Metal*. Prentice-Hall, Englewood Cliffs (1969)
16. Hati, S.K., Rao, S.S.: Determination of optimum machining conditions – deterministic and probabilistic approaches. *Journal of Engineering for Industry, Transaction ASME* 98, 354–359 (1976)
17. Narang, R.V., Fischer, G.W.: Development of a frame work to automate processplanning functions and determine machining parameters. *International Journal of Production Research* 31, 1921–1942 (1993)
18. Coello, C.A.C., Cortés, N.C.: Hybridizing a genetic algorithm with an artificial immune system for Global optimization. *Engineering Optimization* 36(5), 607–634 (2004)
19. Vijayakumar, K., Prabhaharan, G., Asokan, P., Saravanan, R.: Optimization of multi-pass turning operations using ant colony system. *International Journal of Machine Tools & Manufacture* 43, 1633–1639 (2003)
20. Chen, M.C., Chen, K.Y.: Optimization of multipass turning operations with genetic algorithms: a note. *International Journal of Production Research* 41(14), 3385–3388 (2003)

# **Robot Assisted Stroke Rehabilitation: Joint Torque/Force Conversion from EMG Using SA Process**

S. Parasuraman, Arif Wicaksono Oyong, and Veronica Lestari Jauw

School of Engineering,  
Monash University Malaysia,  
Jalan Lagoon Selatan, Bandar Sunway, Selangor 46150  
Malaysia

**Abstract.** This paper focuses on the implementation of robot-assisted stroke rehabilitation using electromyography (EMG) as the interface between the robot and subjects. The key issue in implementing EMG for this application is the conversion process of EMG signal into torque/force, which is used as a input to the control system. This paper presents a methodology of EMG signal conversion into estimated joint torque by using simulated annealing (SA) technique. Basic principle of SA, formulation, and implementation to the problem are discussed in this paper. Experimental studies with real life EMG data have been carried out for five subjects. These studies are used to evaluate the feasibility of the methodology proposed for robot-assisted stroke rehabilitation problem. Experimental investigations and results are discussed at the end of the paper.

**Keywords:** Stroke rehabilitation, robotics, electromyography (EMG), simulated annealing (SA).

## **1 Introduction**

Stroke is the most common cause of disability in the United States and the other part of the world. A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee shows that in 2006, there are 6.4 million of stroke occurrence and 795,000 new and recurrent strokes. The direct and indirect cost of stroke in 2010 is estimated to grow up to \$73.7 billion [1, 2].

Despite of the increasing number of stroke prevalence, it can be seen that there is significant improvement in the rate of mortality. The number goes down, from 150,100 in 2004 to 137,100 in 2006 [1, 2]. The significant reduction of mortality due to stroke can be attributed to improvement in technology and recover techniques. According to National Institute of Neurological Disorders and Stroke, only about 10% of stroke survivors recover almost completely. About 65% of the survivors recover with minor, moderate, or severe impairments that require special treatments [3].

Several researches reported that robot-assisted therapy significantly recovers from stroke through motor learning [5, 6, 7, 8]. P.S. Lum reported that compared to

conventional therapy, robot assisted rehabilitation has advantages in terms of clinical and biomechanical measures [4].

Recently, Electromyography (EMG) has been proposed as an interface between the user and machine. EMG signal shows the energy level of muscles, which are essential for the joint movement. Mulas, M. has developed an EMG-exoskeleton for hand rehabilitation program [10]. The system employs a threshold value, above which, the motion is activated. However, the system acts only as an on-off switch, without the capability of predicting the motion torque.

Various researches have investigated the use of neural network to predict joint torque corresponding to EMG signal [11, 12]. Different approach of neural network application was investigated by DaSalla [13]. In his work, neural network was used to estimate posture of human wrist and forearm based on the measured EMG signals. Although results from neural network applications show positive results, there are still several issues regarding computational load associated with neural network. In large training data, the computational time might affect the feasibility of neural network for rehabilitation applications. Different approach was carried out by Rosen by constructing biomechanical model of human arm [14]. The biomechanical model was developed by using Hill-Based muscle model to construct mathematical representative of muscle mechanics. The use of EMG together as input to the biomechanical model was able to predict joint torque of the corresponding motion. Similar approaches of the use of biomechanical model to predict joint torque have been implemented in various studies [15, 16].

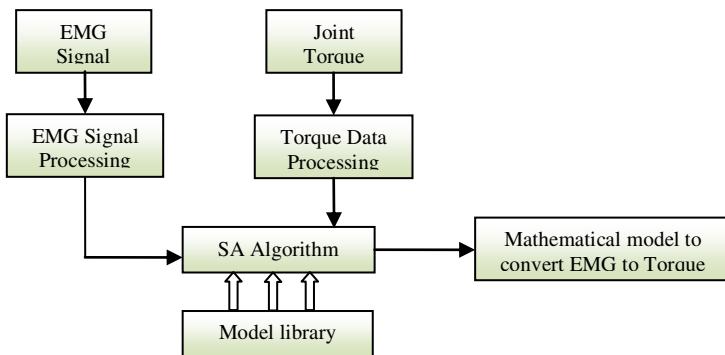
This project presents a methodology to establish a mathematical model that is able to estimate joint torque from the corresponding EMG signal. The system uses collections of EMG as the training data and estimated the corresponding torque using a Simulated annealing.

## 2 Methodology

### 2.1 System Overview

The system focuses on the use of EMG for human upper limb rehabilitation, which covers shoulder (3DOFs) and elbow joint (2 DOFs). Presently, only 3 sets of motion are considered for EMG /torque conversion, which are shoulder flexion-extension, shoulder abduction-adduction, and elbow flexion-extension. Different muscles are responsible for their respective joint movements. Table 1 shows the Upper limb motion, muscles group and descriptions of each muscle.

Figure 1 shows the overview diagram of the system. EMG signal and corresponding joint torque are measured through data acquisition system. Signal processing algorithms are implemented in order to extract the useful information from noise and signal artifacts. SA algorithm is implemented to find a mathematical model that fits processed EMG data into joint torque. However, the model construction itself has to be provided by user. In this case, several possible models are provided into a model library. SA algorithm picks the best fit model from library and finds the optimum parameters corresponding to the selected model.

**Fig. 1.** System Overview**Table 1.** Joint Movement and Muscle group

movement	muscles group	descriptions
 		Shoulder Flexion: Deltoid Anterior Muscle: It covers shoulder joint up to lateral humerus. This muscle originates from clavicle and scapula, and insert into humerus.
		Shoulder Extension: Deltoid posterior Muscles: Greatly responsible for shoulder extension.
		Shoulder Abduction: Deltoid Middle portion: Plays major role for shoulder abduction movements
		Shoulder Adduction: Pectoralis major is a major muscle located in human chest.. Pectoralis major originates from clavicle, sternum, and ribs cartilages, and covers superior, anterior chest, up to humerus.
		Infraspinatus is a triangular muscle at located at human back. It originates from scapula and inserts to humerus. Infraspinatus covers infraspinatus fossa at human back, and mainly responsible for shoulder external rotation.

## 2.2 Simulated Annealing (SA)

SA is a global optimization method to find optimum solution in large optimization problem. Annealing is a process used to improve material properties through relieving the internal stress within the material and refining the crystal structure. The process starts at high temperature, at which disordered internal crystal structure (high energy level) appears, and gradually decreases the temperature until thermodynamic equilibrium is reached. As temperature decreases, the crystal structure is arranged into configuration with lower energy level. At thermodynamic equilibrium the system reaches its lowest energy level.

By following the concept of material annealing, the algorithm starts at certain high temperature  $T_0$  with corresponding energy level of  $E(s_1)$ . The algorithm than evaluates a random neighbor candidate  $s_2$  to obtain the energy level  $E(s_2)$ . The algorithm can choose to accept the new candidate  $s_2$  or remain with the current candidate  $s_1$ . The acceptance of the new candidate is based on probability function that depends on the energy level difference ( $E(s_2) - E(s_1)$ ) and time-dependence variable ‘temperature’ ( $T_i$ ). The probability function is designed according to the following rules:

$$\begin{aligned} P_a &= \text{high} & \left\{ \begin{array}{l} E(s_2) < E(s_1) \\ T_i \text{ is high} \end{array} \right. . \\ P_a &= \text{low} & \left\{ \begin{array}{l} E(s_2) > E(s_1) \\ T_i \text{ is low} \end{array} \right. \end{aligned} \quad (1)$$

The above design enables the algorithm to accept large search space at the beginning of iteration (high temperature) and increasingly converge as the temperature decreases.

**Implementation.** In order to implement SA in our problem, there are few issues that have to be defined:

### 1. Energy function

Energy function is the cost function that is to be minimized. In this case, the energy function  $E$  is the sum of squared difference between actual torque and estimated torque:

$$\begin{aligned} E &= \sum_{k=1}^N (\tau_{\text{ref}} - \tau_k)^2 \\ &= \sum_{k=1}^N (\tau_{\text{ref}} - F(E, (s_i)))^2 \end{aligned} \quad (2)$$

It is to be noted that the form of  $s_i$  is greatly affected by the mathematical model selected from model library. For example:

If the selected model is in the following form:

$$MM1: \tau_{\text{est}} = x(1).E_i + x(2).E_i^{1/2}$$

then,  $s_i$  is in the form of  $s_i = \begin{bmatrix} x(1) & 0 \\ 0 & x(2) \end{bmatrix}$

### 2. Candidate generation procedure

In each iteration, SA algorithm chooses a nearby candidate for evaluation. The selection of this candidate is done randomly within a search diameter. The design

of search diameter will affect the convergence of the algorithm. Large search diameter enables the algorithm to evaluate more solution at the expense of computation time. Small diameter enables faster convergence, but algorithm might be trapped at local minimum. To overcome this problem, candidate generation procedure is designed based on the annealing temperature. When temperature is high, the search distance is large, as the temperature cools down the search distance is reduced, enabling faster convergence.

$$s_i = s_{i+1} + \sqrt{T} * \text{Rand}() . \quad (3)$$

### 3. Acceptance probability

The acceptance of the new state is depending on the energy level of the new state  $S_i$  compared with the energy level of the current state  $S_{i-1}$ . If  $E(S_i)$  is smaller than  $E(S_{i-1})$  than the new state is accepted. However, if  $E(S_i)$  is larger than  $E(S_{i-1})$  the algorithm might still accept it depending on acceptance probability. The acceptance probability is designed based on the following:

#### 1. T (annealing temperature)

The acceptance probability is designed to be large when T is high and small when T is low.

#### 2. dE is the difference between the $E(S_i)$ and $E(S_{i-1})$ . Higher dE will result in lower acceptance probability and vice versa.

To facilitate the above requirement the acceptance probability is designed based on Boltzmann probability density.

$$P = \frac{1}{1 + \exp\left(\frac{dE}{T}\right)} . \quad (4)$$

### 4. Temperature function

Temperature function determines the rate at which the annealing process cools down. This function will greatly affect the acceptance probability as shown previously. In this project, the temperature function is designed as follow:

$$T = T_0 * 0.95^k . \quad (5)$$

The following pseudo-codes show detailed procedure for SA algorithm presented. The overall flowchart diagram of the implemented algorithm is shown in Figure 2.

```

1 Begin
2   T = initial T;
3   Generate initial state
4     s = [s(1); s(2); s(3); ...; s(i)]
5   for m = 1: EMG_Data_Size
6     Δτ = τ_ref - τ(EMG, s)
7     Evaluate initial energy level E
8   End
9   Repeat
10    s_new = candidate generation procedure

```

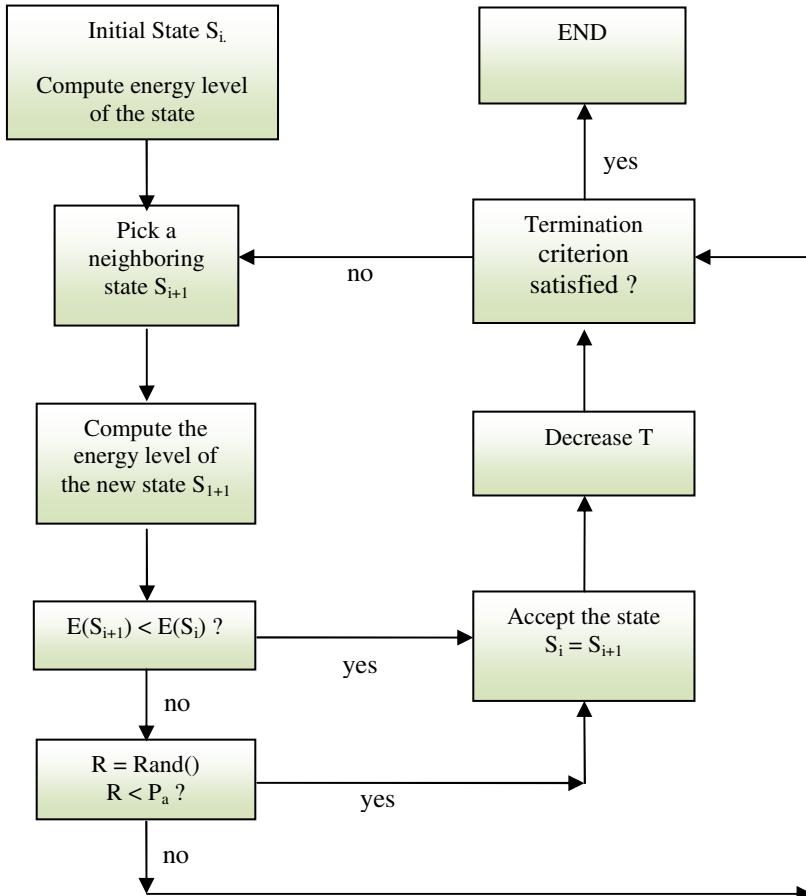
```

11  for m = 1 : EMG_Data_Size
12      Δτ = τ_ref - τ(EMG, s)
13      Evaluate new energy level E_new
14  End
15  Generate random number R
16  Calculate acceptance probability Pa
17  If (E_new < E)
18      s = s_new;           E = E_new;
19  elseif (R < Pa)
20      s = s_new           E = E_new;
21  else
22      s = s;             E = E;
23  endif
24  Decrease temperature T
25  Until termination criteria is TRUE
26  END

```

**Table 2.** List of Main Symbols

Symbols	descriptions
$\tau$	torque
$\tau_{ref}$	measured torque
$P_a$	acceptance probability
$T_i$	SA temperature at $i^{\text{th}}$ iteration
$T_0$	initial temperature
E	energy level or energy function
$E(s1)$	energy level at state s1
$E(s2)$	energy level at state s2
$s_i$	model parameters associated with selected model
u	processed EMG signal
$F(u, (s_i))$	model function, depending on EMG signal and model parameters
Rand()	random number generator
dE	difference in energy level
k	annealing constant associated with temperature function
MM1	mathematical model no. 1
MM2	mathematical model no. 2
R	Pearson correlation
$x_i$	coefficient actual torque
$y_i$	estimated torque
$\sigma$	standard deviation



**Fig. 2.** Flow Chart Diagram of Implemented SA

### 3 Experiment

Experiments were carried out on five different subjects (Table-3), with age range from 20 – 80 years old. Due to safety reason, only healthy able body was selected for the experiment. Various movements were carried out (refer Table ) and the corresponding EMG data were recorded. Surface EMG technique was employed with MP150WSW Acquisition System from Biopac Systems Inc. For each muscle group, silver-silver chloride (Ag-AgCl) electrodes were used with conductive gel to provide better bio-potential readings. The corresponding joint torque could not be directly measured due to limitations on available equipment. The joint torque associated with experimented motion was acquired through Matlab SimMechanics.

Raw data obtained from the data acquisition system cannot be used directly. Raw EMG data has to be processed through signal processing algorithm. The signal processing consists of 5 stages:

- Low-pass filter
- Average rectified
- Moving average
- Muscle activation

The output of signal processing algorithm (processed EMG signal) is passed to SA algorithm for estimation of joint torque.

**Table 3.** Test Subjects

	Gender	Height (cm)	Weight (kg)	Age
S1	Male	170	60	26
S2	Male	175	76	24
S3	Male	160	52	22
S4	Male	167	58	20
S 5	Female	150	43	80

**Table 4.** Motion Performed during Experiment

	Without weight			With weight (5kg)		
	Elbow Flexion-Extension	Shoulder Abduction-Adduction	Shoulder Flexion-Extension	Elbow Flexion-Extension	Shoulder Abduction-Adduction	Shoulder Flexion-Extension
S1	x	x	x	x	x	x
S2	x	x	x	x	x	x
S3	x	x	x	x	x	x
S4	x	x	x	x	x	x
S5	x	-	-	x*	-	-

\*movements were carried out with 500g weight.

The model library consists of various mathematical models that were tested during the experiment, however for the purpose of this paper, the following are presented:

$$\begin{aligned}
 \text{MM1: } \tau_{\text{est}} &= x(1).E_i + x(2).E_i^{\frac{1}{2}} \\
 \text{MM2: } \tau_{\text{est}} &= x(1) + x(2).E_i^2 + x(3).e^{E_i} \\
 \text{MM3: } \tau_{\text{est}} &= x(1) + x(2).\cos(E_i) + x(3).\sin(E_i) \\
 \text{MM4: } \tau_{\text{est}} &= x(1).E_i^{x(2)} \\
 \text{MM5: } \tau_{\text{est}} &= x(1) + x(2).e^{E_i} + x(3).e^{-E_i} \\
 \text{MM6: } \tau_{\text{est}} &= x(1).E_i^{x(2)} + x(3).E_i^{x(4)} \\
 \text{MM7: } \tau_{\text{est}} &= x(1).E_i^4 + x(2).E_i^3 + x(3).E_i^2 + x(4).E_i^1 + x(5).E_i^0 . \quad (6)
 \end{aligned}$$

## 4 Results and Discussions

### 4.1 Calibration

Table 5 shows the result of the experimental training. The fitness value indicates the fitness of a model in predicting joint torque from EMG. The lower the fitness indicates good model predictability. However, it can be seen that for the same model, the fitness value is higher when motion is performed with weight. Therefore evaluation based on fitness value must be performed in relative to the experiment category (results from experiment with weight cannot be compared with results from experiment without weight). Further analysis shows that, fitness value alone is not enough for determination of best fit model. Based on fitness value, the results indicate that MM3 is a better fit compared to MM7. However, graphical analysis shows that MM3 does not show correlation between processed EMG and joint torque. Graphical comparison of MM3 and MM7 are shown in Figure 3.

**Table 5.** Results of Training of Different Models

	Without weight		With weight	
	Fitness	R	Fitness	R
<b>MM1</b>	<u>16.4762</u>	<b>0.8415</b>	<u>10153.25</u>	<b>0.8079</b>
<b>MM2</b>	1.31E+03	0.0923	328020	0.00
<b>MM3</b>	63.4096	0.0220	30782.5	-0.0957
<b>MM4</b>	3.47E+08	-0.1036	328020	0.00
<b>MM5</b>	1.31E+03	0.0923	328065.5	0.00
<b>MM6</b>	6.99E+02	0.8215	328020	-0.0566
<b>MM7</b>	<u>225.5363</u>	<b>0.8023</b>	<u>38744.25</u>	<b>0.7741</b>

To overcome the problem, Pearson correlation coefficient was introduced. Pearson correlation coefficient (R) is a statistical method to determine the relationship between two random variables. Correlation coefficient can be computed as follows:

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)\sigma_x \sigma_y}. \quad (7)$$

The value of R ranges from -1 to 1. R = 1 indicates that the two variables are proportional to each other, while -1 indicates inverse proportionality. R = 0 shows that there is no correlation between two random variables. Correlation coefficients of associated models are shown in Table 5.

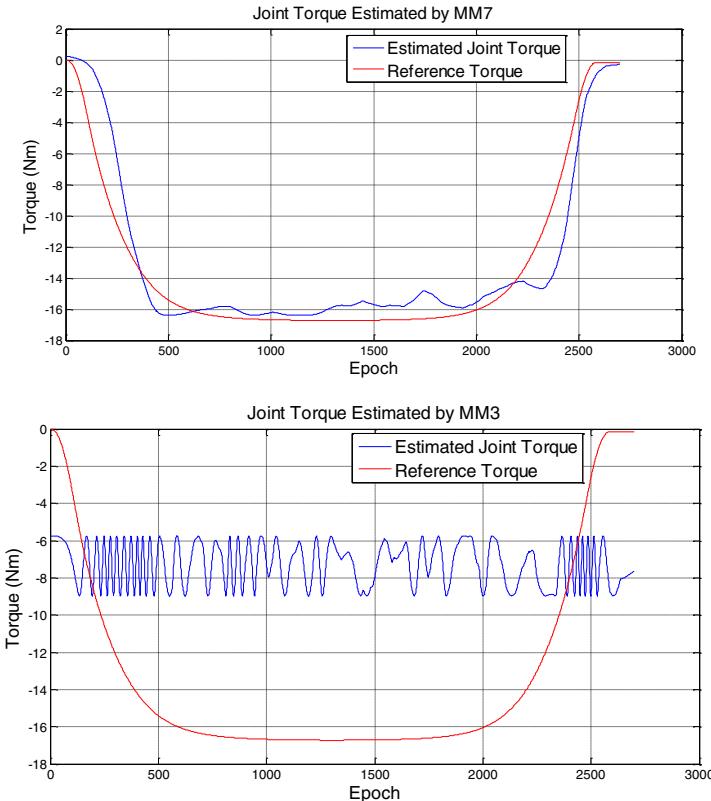
For the above results, correlation analysis shows that only MM1 and MM7 can be used for joint torque estimation. Comparing the fitness value between the two, we can indicate that MM1 is the best fit model among the 7 models presented.

$$MM1: \tau_{est} = x(1).E_i + x(2).E_i^{\frac{1}{2}} . \quad (8)$$

From the experiment, the followings parameters are to give good result.

$$x(1) = 0.0495 \quad x(2) = -0.461$$

It is to be noted that model and associated parameters obtained are specific for an individual. Furthermore, further training might result in slight difference parametrical values.



**Fig. 3.** Comparisons of Joint Torque Estimated by MM7 and MM3

## 4.2 Model Validation

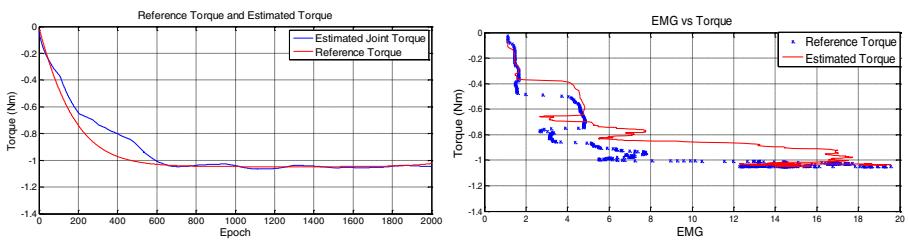
Results presented in Table 5 were obtained with training data. In order to validate the model, the model was tested with EMG data that has never been used in the calibration.

Figure 4 and Figure 5 show that the model obtained through SA is able to predict joint torque for untrained EMG data. One important finding is that a model calibrated with a certain EMG data cannot be used to predict joint torque corresponding to EMG signal beyond that data. That means that the model calibrated with motion with 5kg

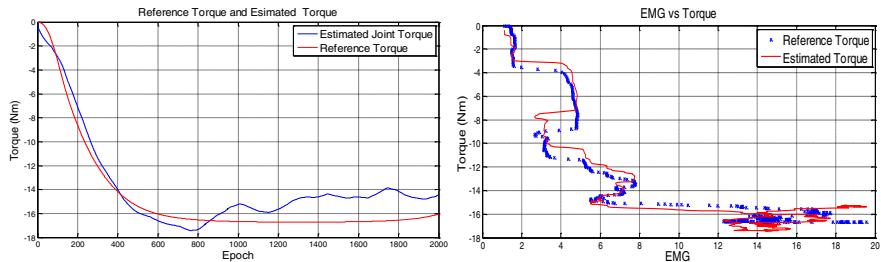
weight can be used to predict joint torque corresponding to motion without weight. However, the same results are not true in opposite condition.

Possible errors may occur during the calibration of the mathematical model. Torque data used during this experiment was ideal torque based on simulation. On the other hand, the EMG data was obtained from actual experiment. Because of this issue, the calibration data was not synchronized properly. Introducing torque sensor will solve this problem, enabling synchronized recording of EMG and corresponding torque. However, additional noise will be incurred by the sensor. Further considerations have to be focused on EMG and torque recording synchronization and signal processing of data obtained.

Calibration by GA and SA can be further improved with more available data. Ideally, the calibration data should range from the free motion up to maximal voluntary contraction. The availability of these data means that the calibration covers the full range of muscle contractions. Therefore, the resulted mathematical model will be able to provide better estimate of joint torque.



**Fig. 4.** Results with Untrained Data 1



**Fig. 5.** Results with Untrained Data 2

## 5 Conclusion and Future Works

A methodology for EMG to joint torque conversion by using SA approach has been presented in this paper. Experimental investigation shows that the system is able to find mathematical model that predicts joint torque from associated EMG signal.

Better experimentation can be achieved by using simultaneous measurement of EMG and the corresponding joint torque. Synchronization between EMG and torque data will enable more accurate training data. The investigation was done with healthy

subjects. Further experimentation has to be carried out with stroke affected patient, in order to assess the feasibility of the proposed methodology for stroke rehabilitation purpose.

## References

1. American Heart Association Statistics Committee and Stroke Statistics Subcommittee: Heart Disease and Stroke Statistics – 2007 Update (2007)
2. American Heart Association Statistics Committee and Stroke Statistics Subcommittee: Heart Disease and Stroke Statistics – 2010 Update (2010)
3. National Institute of Neurological Disorders and Stroke,  
<http://www.ninds.nih.gov/>
4. Jette, D., Warren, L., Wirtalla, C.: The relation between therapy intensity and outcomes of rehabilitation in skilled nursing facilities. *Archives of Physical Medicine and Rehabilitation* 86(3), 373–379 (2006)
5. Lum, P.S., Burgar, C.G., Shor, P.C., Majmundar, M., Van der Loos, M.: Robot-Assisted Movement Training Compared with Conventional Therapy Techniques for the Rehabilitation of Upper-Limb Motor Function After Stroke. *Archives of Physical Medicine and Rehabilitation* 83 (2002)
6. Neale, T., Writer, S.: ASA: Robot-Assisted Stroke Rehab has Some Benefit, In: MedPage Today, SA, USA (March 1, 2010)
7. Colombo, R., Pisano, R., Micera, S., Mazzone, A., Delconte, C., Carrozza, M.C., Dario, P., Minuco, G.: Robotic Techniques for Upper Limb Evaluation and Rehabilitation of Stroke Patients. *IEEE Trans neural System Rehabilitation Engineering* 13, 311–324 (2005)
8. Aisen, M.I., Krebs, H.I., Hogan, N., McDowell, F., Volpe, B.T.: The Effect of Robot-Assisted Therapy and Rehabilitative Training on motor Recovery Following Stroke. *Archives of Neurology* 54(4), 443–446 (1997)
9. Fasoli, S., Krebs, H., Stein, J., Frontera, W., Hogan, N.: Effects of Robotic Therapy on Motor Impairment and Recovery in Chronic Stroke. *Archives of Physical Medicine and Rehabilitation* 84(4), 477–482
10. Mulas, M., Folgheraiter, M., Gini, G.: An EMG-controlled Exoskeleton for Hand Rehabilitation. In: International Conference on Rehabilitation Robotics, pp. 371–374 (2005)
11. Luh, J.J., Chang, G.C., Cheng, C.K., Lai, J.S., Kuo, T.S.: Isokinetic Elbow Joint Torques Estimation from Surface EMG and Joint Kinematic Data: Using an Artificial Neural Network Model. *Journal of Electromyography and Kinesiology* 9(3), 173–183 (1999)
12. Song, R., Tong, K.Y.: Using recurrent Artificial Neural Network Model to Estimate Voluntary Elbow Torque in Dynamic Situations. *Medical & Biological Engineering & Computing* 43(4), 473–480 (2005)
13. DaSalla, C., Kim, J., Koike, Y.: Robot Control Using Electromyography (EMG) Signals of the Wrist. *Applied Bionics and Biomechanics* 2(2), 97–102 (2005)
14. Rosen, J., Brand, M., Fuchs, M.B., Arcan, M.: A Myosignal-Based Powered Exoskeleton System. *IEEE Transaction on Systems, Man, and Cybernetics – Part A: Systems and Humans* 31(3) (2001)
15. Shin, D., Kim, J., Koike, Y.: A Myokinetic Arm Model for Estimating Joint Torque and Stiffness from EMG Signals during Maintained Force. *Journal of Neurophysiology* 101, 387–401 (2008)
16. Fleischer, C., Wege, A., Kondak, K., Hommel, G.: Application of EMG Signals for Exoskeleton Robots. *Biomed. Tech.* 51, 314–319 (2006)

# Genetic Goal Oriented Path Planning Algorithm for Acute Obstacle Avoidance in Mobile Robot Navigation

Soh Chin Yun<sup>1</sup>, S. Parasuraman<sup>1</sup>, and V. Ganapathy<sup>2</sup>

<sup>1</sup> School of Engineering, Monash University, Sunway Campus,  
Jalan Lagoon Selatan, 46150 Bandar Sunway, Malaysia

<sup>2</sup> Department of Electrical Engineering, Faculty of Engineering,  
University of Malaya, 50603 Kuala Lumpur, Malaysia

mscysoh@gmail.com, s.parasuraman@eng.monash.edu.my

**Abstract.** Current researches in mobile robots aim to build an autonomous and intelligent system which can plan its motion in a static, dynamic and uncertain environment. In this research, Genetic Algorithm (GA) is used to assist the autonomous mobile robot to move, identify the obstacles by learning the environment and reach the desired goal. This study is focused on the development of algorithm for the acute obstacle avoidance behavior. The proposed navigation technique is capable of re-planning the new optimum collision free path in the event of mobile robot encountering any dynamic obstacles. The method is verified using MATLAB simulation and validated the algorithm using Team AmigoBotTM robot. The results obtained from simulation and experiments are discussed at the end of the paper.

**Keywords:** Genetic Algorithm (GA), Genetic Controller, Goal Oriented Path Planning Algorithm (GOPPA), Team AmigoBotTM robot and MATLAB.

## 1 Introduction

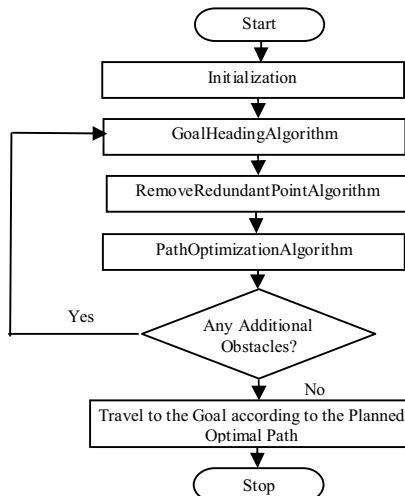
Mobile robot navigation is the technique to guide the mobile robot moving from the starting position towards the desired goal, along a desired path in a static, dynamic, known and unknown environment. The environment of mobile robot navigation is distinguished by variable terrain, a set of distinct objects such as obstacles, milestones, and landmarks [1]. Robot navigation can be accomplished through behavior arbitration [2] and hierarchical behavior control [3]. The influence AI of techniques further enlightened the development of behavior arbitration for robot navigation. In the last decade, genetic algorithms have been widely used as alternative method to generate the optimum path by taking advantage of its strong optimization ability [4]. Generally, the initial population of the classic Genetic Algorithm is generated randomly. This is the easiest method but will lead to large quantities of infeasible paths if it is to be used in path planning [9, 10, 11 and 12]. Qing Li et al [5, 10, 7] have proposed an algorithm in order to generate feasible initial population. Although the existing algorithms have rapid search and high search quality, there are six problems associated with the existing methods. Firstly, the initial population contains many

infeasible paths, which have negative influence on the performance of the genetic algorithm. Secondly, there are not sufficient heuristic knowledge based genetic operators. Thirdly, after each generation, offspring may contain infeasible path. Forth, the guidance of the mobile robot was from the starting position to the desired goal and do not take into consideration the circumstances when mobile robot is trapped inside acute ‘U’ or ‘V’ shaped obstacles. In the cases when mobile robot encounters acute ‘U’ or ‘V’ shaped obstacles, mobile robot will be trapped and unable to come out from the trapped environment. Fifth, during navigation, robot might encounter dynamic obstacles. The existing algorithms do not take into consideration about these circumstances. Lastly, there is no reported real time implementation to prove the applicability of the algorithm in the actual mobile robot.

In order to generate a feasible collision free path for an acute ‘U’ or ‘V’ shaped obstacle and dynamic obstacles, a Goal Oriented Path Planning Algorithm (GOPPA) is proposed. The proposed GOPPA is goal oriented and thus reduces unnecessary search time. The effectiveness of this algorithm is demonstrated by simulation studies and experiments using Team AmigoBot™ [8] robot. The performances of GOPPA in both simulation and real time implementation are investigated and evaluated on various aspects.

## 2 Proposed Methodology

The GOPPA is designed to create a hybrid genetic algorithm capable of identifying the obstacles in the environment and guide the robot to navigate autonomously. The flowchart of the proposed Goal Oriented Path Planning Algorithm (GOPPA) is shown in Figure-1. Details of how GOPPA works are described in the following steps:



**Fig. 1.** Goal Oriented Path Planning Algorithm (GOPPA)

**Step 1: Initialization**

GOPPA based controller starts with the initialization of the parameters.

**Step 2: GoalHeadingAlgorithm**

*GoalHeadingAlgorithm* is activated to assist the mobile robot travel from the starting position to the desired goal in a collision free path. The pseudocode of *GoalHeadingAlgorithm* is as below:

```

while not reaching goal
    identify the nearest obstacle and its distance from the
    mobile robot

    if no obstacle is detected
        define rotational angle of the mobile robot, θ,
    else

        define the fitness function
        define and compute the values of all the
        inequality/equality constraints
        define the parameters for genetic operation
        run genetic algorithm

        if same path is traversed again
            build virtual wall
            guide the mobile robot to escape from the trap
        endif

        if the proposed_next_position is not inside obstacle area
            proposed_next_position is taken as the next position
        endif

        if the mobile robot is not crossing any obstacle
            mobile robot moves forward
        endif
    endif...else
endwhile

```

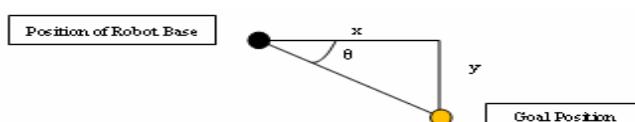
Figure 2 depicts the rotation angle of the mobile robot towards the predefined goal. In this controller, rotational angle of the robot towards the goal is controlled by the angle ‘theta’ which is defined as below:

$$\theta = \tan^{-1} \frac{y}{x} \quad (1)$$

where  $y$  =  $y$ \_coordinate of goal position and

$x$  =  $x$ \_coordinate of goal position

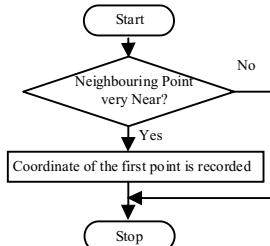
both with reference to the robot’s base position.



**Fig. 2.** The Rotational Angle of the Mobile Robot,  $\theta$

### Step 3: RemoveRedundantPointAlgorithm

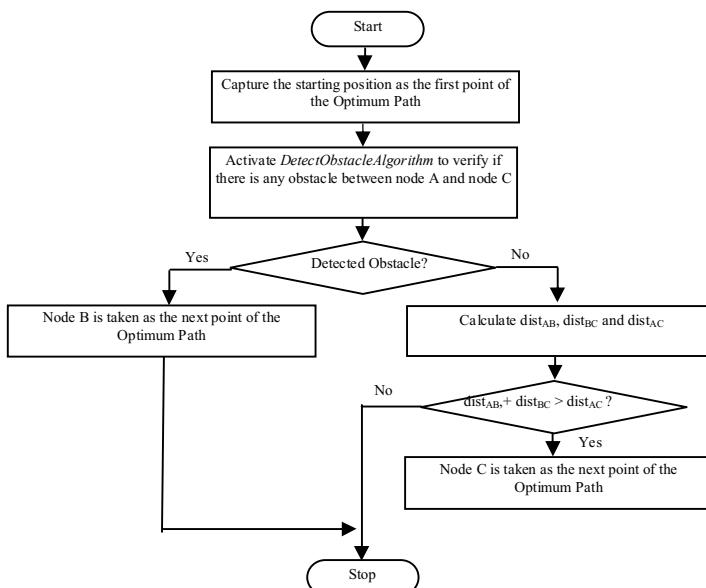
After new collision free path is found, unnecessary neighbouring coordinate(s) is/are screened out to remove redundant points lying along the path through *RemoveRedundantPointAlgorithm*. The flowchart of *RemoveRedundantPointAlgorithm* is shown in Figure 3.



**Fig. 3.** The flowchart of RemoveRedundantPointAlgorithm

### Step 4: PathOptimizationAlgorithm

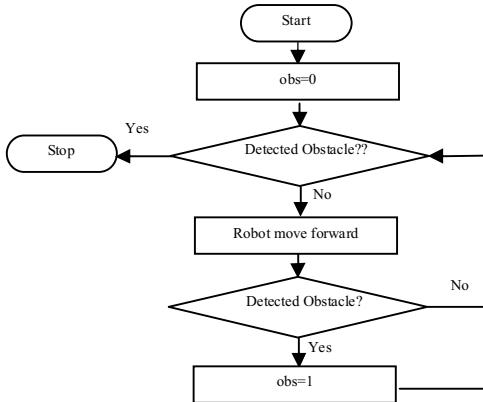
When the collision free path without redundant point is found, *PathOptimizationAlgorithm* is utilized to find the optimum path from the obstacle free path obtained after *GoalHeadingAlgorithm* and *RemoveRedundantPointAlgorithm*. The flowchart of *PathOptimizationAlgorithm* is shown in Figure 4.



**Fig. 4.** The flow chart of Path Optimization Algorithm

### Step 5: DetectObstacleAlgorithm

*DetectObstacleAlgorithm* is called to identify if there is/are any dynamic obstacle(s) while guiding the mobile robot to travel from the starting position to the desired goal according to the optimized collision free path. The flowchart of *DetectObstacleAlgorithm* is shown in Figure 5.



**Fig. 5.** The flowchart of *DetectObstacleAlgorithm*

### Step 6: No Dynamic Random Obstacle

If no dynamic random obstacle is detected during the process of guiding the mobile to the desired goal, the mobile robot will travel from the starting position to the desired goal according to the optimized collision free path determined by *PathOptimization-Algorithm*.

### Step 7: Encounter Dynamic Random Obstacle

In the cases where dynamic random obstacle(s) is/are positioned in the environment, the current position is taken as the new starting position. The original goal remains as the new desired goal of the mobile robot and Step 2 – Step 7 are repeated. *GoalHeadingAlgorithm* is activated to search for new obstacle free path from the new starting position to the desired goal.

## 3 Results and Discussions

### 3.1 MATLAB Simulation

Based on Table 1 and 2, simulation is carried out and the result is shown in Figure 6. In the simulation environment, the robot is shown as solid circle, obstacles are rectangular blocks and target is a square box. After identifying the collision free path, unnecessary paths are removed to create or obtain an optimum collision free path from the starting position to the desired goal position. The Test-1a and Test-1b are carried out to find the optimum path without and with dynamic obstacles respectively.

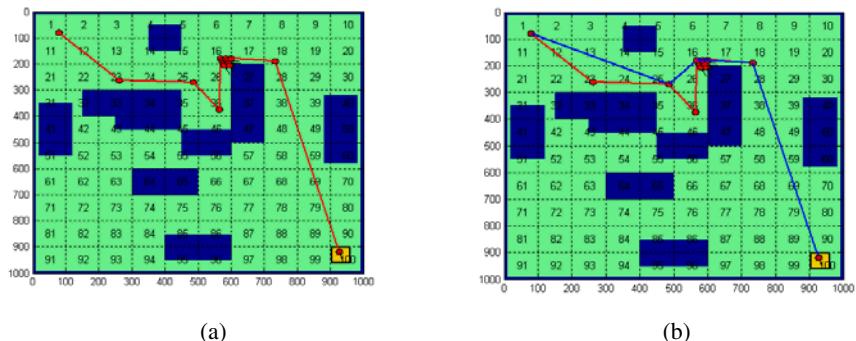
The final path of Test-1a is shown in Figure 7. The final path of Test-1b is shown in Figure 10.

**Table 1.** Input Parameters for Test-1a

<b>Environment size</b>	1000 x 1000, shown as 10 x 10 grid
<b>Starting position</b>	(80, 80)
<b>Goal position</b>	(930, 930)
<b>Obstacles</b>	Obstacles defined by user
<b>Dynamic Obstacles after Path Optimization</b>	No.

**Table 2.** Input Parameters for Test-1b

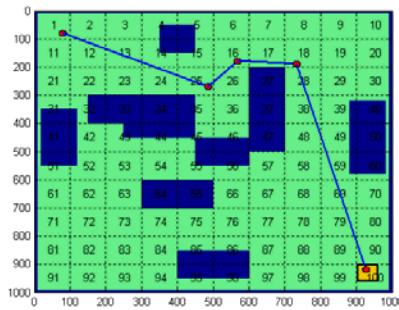
<b>Environment size</b>	1000 x 1000, shown as 10 x 10 grid
<b>Starting position</b>	(80, 80)
<b>Goal position</b>	(930, 930)
<b>Obstacles</b>	Obstacles defined by user
<b>Dynamic Obstacles after Path Optimization</b>	Dynamic obstacles encountered.



**Fig. 6.** Results of Test-1 (a) Collision Free Path from the Starting Position to the Desired Goal; (b) Optimized the Collision Free Path

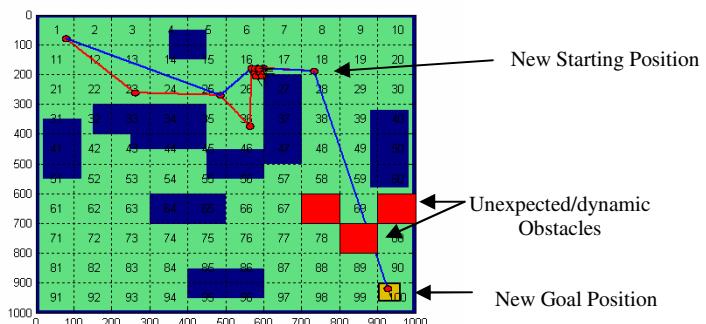
**Static Obstacles:** Referring to the results of Test-1 as in Figure 6, GOPPA is capable of coming out with the optimum collision free path to guide the mobile robot moving from the starting position to the desired goal. Unnecessary pathways were removed in order to come out with an optimum collision free path. A minimum clearance distance was provided between mobile robot and obstacles to avoid the mobile robot from colliding with the obstacles. Virtual walls were created so that the same grids and positions would not be traversed, travelled or analyzed again. Moreover, mobile robot was able to move away from ‘U-shaped’ type obstacles and it did not traverse in between two obstacles.

The final optimum collision free path from the starting position to the goal position without dynamic obstacles is shown in the trace line in Figure 7.

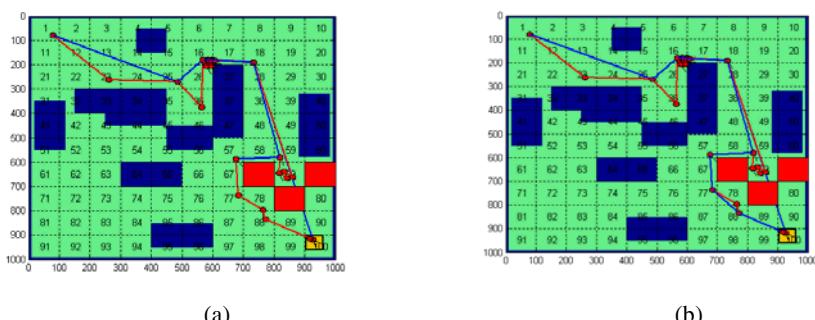


**Fig. 7.** Mobile Robot Travels from the Starting Position to the Desired Goal in the Optimum Collision Free Path WITHOUT Dynamic Obstacles (Test-1a)

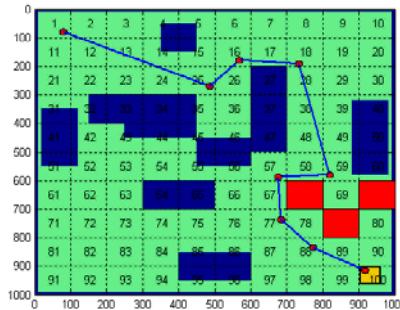
**Dynamic Obstacles:** Figure 8 depicts the simulated environment of Test-1 with dynamic obstacles. Figure 9 (a) and (b) are the possible paths from the starting point to the target after applying step 5. Optimization process is carried out using step-4 to achieve the near optimum path planning in goal seeking. Figure 10 shows the result obtained by Test-1b.



**Fig. 8.** Dynamic Obstacle(s) is/are in the Simulated Environment of Test-1



**Fig. 9.** Results of Test-1b (a) Collision Free Path from the Starting Position to the Desired Goal when Dynamic Obstacle(s) are encountered; (b) Optimized the Collision Free Path



**Fig. 10.** Mobile Robot Travels from the Starting Position to the Desired Goal in the Optimum Collision Free Path WITH Dynamic Obstacles (Test-1b)

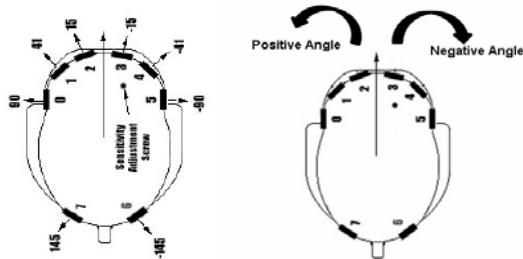
Referring to Figure 8, mobile robot does not detect any dynamic obstacle in the optimum collision free path before grid 18. Mobile robot travels according to the collision free path determined earlier and no path searching is required until grid 18. Since *DetectObstacleAlgorithm* is in activation, the system could find dynamic obstacles in between grid 18 and grid 100. In this instance, a new starting position is re-located at the grid 18 and the new target position is re-positioned at grid 100. *GoalHeadingAlgorithm* is activated to re-plan the new path to guide the mobile robot travelling from the current position to the desired goal without colliding with any obstacles in the environment.

Referring to the results of Test-1b as in Figure 9, GOPPA is capable of finding the optimum collision free path. The resultant path from the starting position (grid 1) to the goal position (grid 100) with dynamic obstacles is shown in the trace line in Figure 10.

From the simulation results, it is found that proposed GOPPA was successfully implemented to (i) generate the feasible path, (ii) resolve the circumstances where mobile robot is trapped in the acute ‘U’ or ‘V’ shaped obstacle, (iii) find an optimum and shortest collision free path with a given starting position and goal position, and (iv) avoid dynamic obstacle(s). In short, MATLAB simulation has verified and has validated the effectiveness of GOPPA.

### 3.2 Real Time Implementation

Team AmigoBot<sup>TM</sup> robot is used for real time implementation of GOPPA. The standard AmigoBot comes with eight range-finding sonars: one on each side, four forward facing and two in the rear for 360-degree sensing coverage. The direction rotation of Team AmigoBot<sup>TM</sup> robot is shown in Figure 11. The actual working environment and two dimensional top view are shown in Figure 12 and Figure 13.

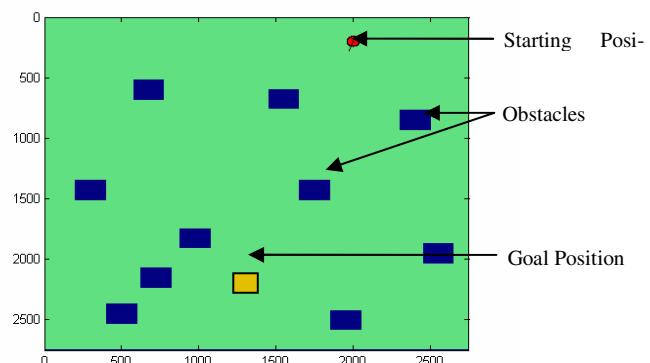


**Fig. 11.** Top View of Team AmigoBot<sup>TM</sup> robot with Mounted Angles of Sensors [8]

The real time implementation of the algorithm real world experiments was carried out for static and dynamic environment. In both cases, more than one possible collision free paths are identified using the steps 1 and 2. Finally the optimal path compared to the other possible paths is explored by the steps 3 and 4 of the algorithm. This is shown in Figure 14(a) and Figure 15(a) for static and dynamic obstacles respectively.



**Fig. 12.** Actual Environment (2750mm x 2750mm) of Randomly Located Obstacles, Team AmigoBot<sup>TM</sup> Robot is Located at (a) Starting Position (b) Goal Position



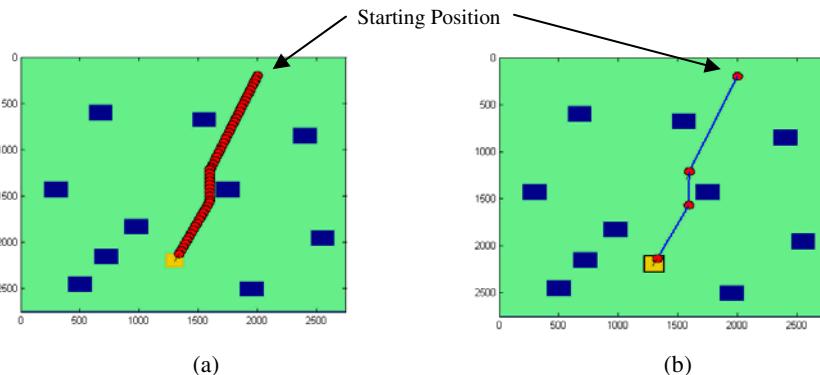
**Fig. 13.** Real Time Implementation. Test-2

**Table 3.** Input Parameters for Test-2a

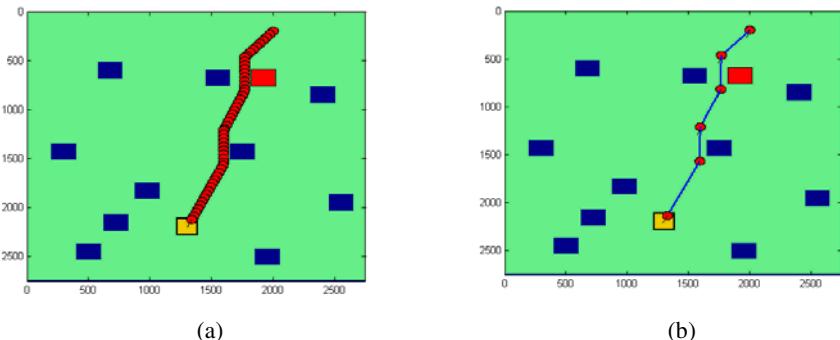
Environment size	2750 x 2750
Starting position	(2000, 200)
Goal position	(1300, 2090)
Obstacles	10.
Dynamic Obstacles after Path Optimization	NO.

**Table 4.** Input Parameters for Test-2b

Environment size	2500 x 2500
Starting position	(2000, 200)
Goal position	(1300, 2090)
Obstacles	10.
Dynamic Obstacles after Path Optimization	YES.

**Fig. 14.** Result of Test-2a (a) in Real Time Implementation Representation Environment; (b) in MATLAB Simulation Environment

Referring to the results of Test-2b, the same path pattern was obtained for both real time implementation and MATLAB simulation. For Test-2b, a dynamic obstacle was detected when guiding the mobile robot to travel according to the optimum collision free path determined by the controller. Thus, when the sensors detected the dynamic obstacles, which are blocking the originally planned path, *GoalHeadingAlgorithm* is turned on to re-plan the new path and guide the robot to localize the current position as starting point. The current position is set to become new starting position and the original goal position is set to become new goal position. Thus, the trace lines as seen in Figure 15.



**Fig. 15.** Result of Test-2b (a) in Real Time Implementation Representation Environment; (b) in MATLAB Simulation Environment

## 4 Conclusions

In this research, a Goal Oriented Path Planning Algorithm (GOPPA) was established and investigated using simulation studies and real world experiments. Using GOPPA, the robot seeks the best possible path and navigated without any collision for static and dynamic environment. The proposed Path remembering algorithm was able to drive the robot autonomously and seek best possible path to reach the target, even in acute situation such as corners, U/V shaped obstacles and dead ends. Optimization approach proposed in this research assists the mobile robot to remove the redundant path ways to reduce the total travelled distance. In the event of dynamic obstacle, GOPPA was capable of re-plan the new start point, and able to map for the best collision free path and drive the mobile robot. The results obtained from real time implementation were compared with the MATLAB simulation study to verify the practicality and robustness of the proposed algorithm. The effectiveness of the proposed GOPPA have been validated using MATLAB simulation and real world experiments using Team AmigoBot™ robot in the aspects of (i) best possible path, (ii) acute obstacle avoidances, (iii) static and dynamic obstacles, and (iv) find a shortest path from starting position to the goal position.

## References

1. Kam, M., Zhu, X., Kalata, P.: Sensor Fusion for Mobile Robot Navigation. Proceedings of the IEEE 85(1), 108–119 (1997)
2. Saffiotti, A.: Fuzzy Logic in Autonomous Robotics: Behavior Coordination. In: Proceedings of the 6th IEEE Int. Conf. on Fuzzy Systems, Pages, pp. 573–578 (1997)
3. Tunstel, E., Lippincott, Y., Jamshidi, M.: Behaviour Hierarchy for Autonomous Mobile Robots - Fuzzy-behaviour Modulation and Evolution. International Journal of Intelligent Automation and Soft Computing 3(1), 37–50 (1997)
4. Tunstel, E., Jamshidi, M.: On Genetic Programming of Fuzzy Rule-Based Systems for Intelligent Control. International Journal of Intelligent Automation and Soft Computing 2(3), 273–284 (1996)

5. Li, Q., Zhang, W., Yin, Y., Wang, Z., Liu, G.: An Improved Genetic Algorithm of Optimum Path Planning for Mobile Robots. In: Proceeding of the Sixth International Conference on Intelligent Systems Design and Applications (ISDA 2006), Jinan, China, vol. 2, pp. 637–642 (2006)
6. Li, Q., Tong, X., Xie, S., Zhang, Y.: Optimum Path Planning for Mobile Robots Based on a Hybrid Genetic Algorithm. In: Proceeding of the Sixth International Conference on Hybrid Intelligent Systems (HIS 2006), Rio de Janeiro, Brazil (2006)
7. Li, Q., Tong, X., Xie, S., Liu, G.: An Improved Adaptive Algorithm for Controlling the Probabilities of Crossover and Mutation Based on a Fuzzy Control Strategy. In: Proceeding of the Sixth International Conference on Hybrid Intelligent Systems (HIS 2006), Rio de Janeiro, Brazil (2006)
8. Mobile Robots Inc., Team AmigoBotTM Operations Manual version 4, Mobile Robots Inc. (2007)
9. Fujimura, K., Samet, H.: A Hierarchical Strategy for Pat Planning Among Moving Obstacles. IEEE transactions on robotics and Automation 5(1) (February 1999)
10. Sanjiv Singh, J., Wagh, M.D.: Robot Path Planning using Intersecting Convex Shapes: Analysis and Simulation. IEEE Journal of Robotics and Automation FA-3(2) (1987)
11. Saripalli, S., Sukhatme, G.S., Montgomery, J.F.: An Experimental Study Of The Autonomous Helicopter Landing Problem. In: The Eight International Symposium on Experimental Robotics, July 2002, pp. 8–11 (2002)
12. Gerkey, B.P., Mataric, M.J.: Principled Communication for Dynamic Multi-Robot Task Allocation. In: Experimental Robotics VII. LNCIS, vol. 271, pp. 353–362. Springer, Heidelberg (2001)

# Services Integration Framework for Vehicle Telematics

S. Veera Ragavan<sup>1</sup>, S.G. Ponnambalam<sup>1</sup>, Velappa Ganapathy<sup>2</sup>, and Joshua Teh<sup>1</sup>

<sup>1</sup> Monash University, Sunway campus, Bandar Sunway,  
46150 Selangor, Malaysia

{Veera.ragavan, Name}@eng.monash.edu.my  
<sup>2</sup> Department of Electrical Engineering, Faculty of Engineering,  
University of Malaya, 50603 Kuala Lumpur, Malaysia

**Abstract.** Development of a General Application Framework based on SOA for a Robust Service Platform that can be used in dynamic environments employing Open Systems, Open Standards and Open Source is being proposed here to provide a basis for research and development of complete systems and solutions specifically targeted for Telematic Applications. Example Service Implementations using Telematics and Web Service Techniques illustrate the efficient Service Invocation and interaction in the Service Platform architecture as Service Components and Applications. The Functional, Design and Structural Aspects of the Service Architecture are also described.

**Keywords:** Service Oriented Architecture (SOA), Telematics, Services Computing (SC), Service Platform (SP), OSGi Framework, Open Source, Open Systems, Vehicle Telematics (VT), Java Virtual Machine (JVM).

## 1 Introduction

The use of modern software engineering architectures like Service Oriented Architecture (SOA), Object-Oriented (OO) , Distributed Computing (DC) and Client/Server architectures are essential for Rapid Application Deployment and effective maintenance of program code and Services. “Services” are now fairly mature and can be made available “On-Demand” and distributed over networks as standalone Embedded Services or as standards-based web services (XML, SOAP, WSDL, and UDDI) which can be ‘plugged’ into the Frameworks. It has been rightly predicted that widely commercialized Service areas in telecommunications and Intelligent Vehicles will be based on SOA [1].

The major problems for Rapid Application Development today lie not only in the hardware side but also on the software side, and effort and time can be saved if hardware components are available Completely-Off-The-Shelf (COTS) and software processes lend itself to “Service Oriented Architecture” (SOA). SOA which is an IT approach or concept for making systems robust against change has to be gainfully utilized in the framework for Telematic applications, in which applications either provide services or rely on services made available by other applications on the networks. It is quite ironical that Telematics Systems today are not quite Service Oriented at least in a Software Architectural sense. Hence it is only logical to expect and

extend that Telematic Service Frameworks should also leverage on these architectures. Moreover Services become cost effective only when “outsourced and distributed” and hence a low cost would also mean Distributed Services “On-Demand” through a Service Oriented Architecture core.

Development of a General Application Framework based on OSGi [2] that can be used in all environments employing open systems, open standards and open source is being proposed here to provide a basis for research and development of complete systems and solutions specifically targeted for Telematic Services applications capable of running on any Operating System(OS). The term ‘open standard’ used here refers to the royalty-free terms for the unregistered usage of the software libraries, modules and components [3]. In addition, OSGi has been used to integrate existing automotive protocols and networks in Vehicle Telematics [4]. Some of the key benefits of OSGi such as platform and application independence [5] and integration communication Protocols, are salient features of this development.

As OSGi is a service-oriented platform, services can be implemented and aggregated by the deployment of bundles in OSGi framework’s [6] and executed using Java Virtual Machines (JVM’s). Any OS running a JVM can deploy the OSGi Service platform [6]. These OSGi services run in a single OSGi container executing on a single JVM, thus posing minor limitations for vehicle telematics. However, the distribution of OSGi services is possible via a middleware platform known as R-OSGi [7]. R-OSGi permits seamless embedding in OSGi, such that local and remote services are indistinguishable to the OSGi framework [7]. This allows the distribution of services to be seamlessly integrated.

Finally, the success of this work also relies on the services and system integration and testing of the hardware components to form a fully functional system. However, it should be emphasized that the design of the framework and system is not restricted only to these hardware components. Hardware independence (through COTS hardware components) is a core feature of the implemented System.

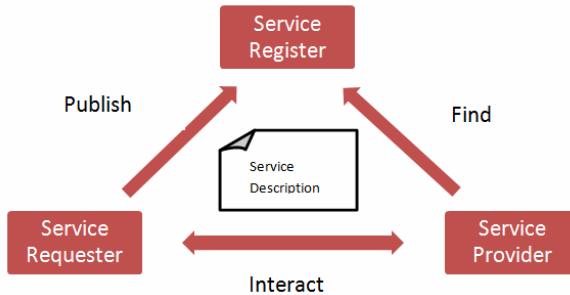
## 2 Evolution of Software Frameworks

In spite of an explosion of technology and methods, Service frameworks are still in early stages of development. Although Bottom –up Framework approaches that exploit new methods or technologies have yielded good results, due to the diversity and convergence of many new and available technologies, a paradigm shift in approach to a “Top down approach” for convergence of relevant technologies is the need of the hour [8]. Frameworks typically provide a way to manage a system of interacting objects, develop objects that will integrate seamlessly into the framework, suggest patterns of collaboration between components and allow flexible channels of collaboration that suit the application at hand. Components emerging from a finished framework share consistent design attributes, and may even share common implementations. The benefit is a more maintainable and consistent software system [9].

## 3 Service Oriented Architectures (SOA) and Challenges

SOA is not a new concept. Sun inc. defined SOA in the late 1990's to describe Jini, an environment for dynamic discovery and use of services over a network. [10]

SOA is an architectural style whose goal is to achieve loose coupling among interacting software agents for Service Component Reuse. Applications in SOA are built based on services. A service is an implementation of well-defined business functionality, and such services can then be consumed by clients in different applications or business processes. In general terms a “Service” is a unit of work done by a service provider to achieve desired end results for a service consumer. Both provider and consumer are roles played by software agents on behalf of their owners.



**Fig. 1.** Service Models [11]

The idea of SOA departs significantly from that of object oriented programming (where data and processing are bound together). A Service usually results in the change of state for the consumer or for the provider or for both. Consuming a service is usually cheaper and more effective and smarter way than doing the work ourselves. The same rule applies to building software systems. Also called "separation of concerns", it is regarded as a principle of software engineering.

SOA permits reuse of existing assets. New services can be created from existing IT systems and businesses can leverage existing investments through reuse of existing applications. This promises interoperability between heterogeneous applications and technologies. SOA provides a level of flexibility that wasn't possible before in the sense that:

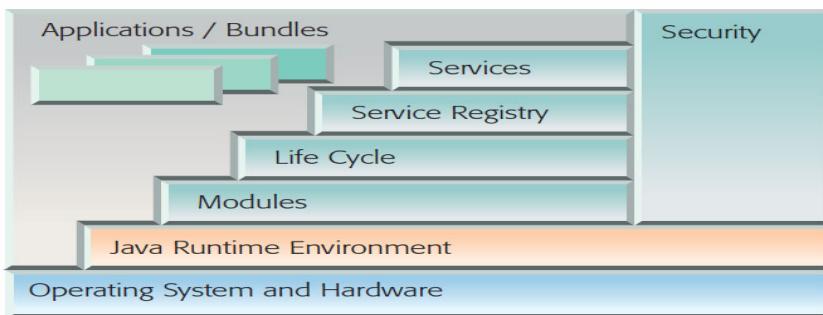
- Services are implementation-independent software components with well-defined interfaces separating service interfaces (What's) from its implementation (How's) for consumption by clients who are not concerned with how these services will execute their requests.
- Services are self-contained (perform predetermined tasks) and loosely coupled (for independence) and can be Services can be dynamically discovered.
- Composite services can be built by aggregating other smaller services.

SOA is emerging as the premier integration and architecture framework in today's complex and heterogeneous computing environment. While any of the Telematic functions such as Navigation can be made into a service, the challenge is to define a service interface that is at the right level of abstraction such that Services provide coarse-grained functionality. The automotive industry worldwide has done a lot of research on telematics platforms, services and infrastructures[1]. While most of the research projects successfully demonstrated the technical feasibility as well as the

value of telematics systems and services, no mass deployments on the market was possible. While the reasons ranging from lack of defined responsibility for operating network of telematics vehicles to legal and business issues are discussed what concerns researchers most is that they face a potential gap – i.e. The missing Software Framework, Standards and Platforms for Telematics Services[ 13].

### 3.1 OSGi Framework

OSGi technology [3] provides a service-oriented, component-based environment for developers and offers standardized ways to manage the software lifecycle. These capabilities greatly increase the value of a wide range of computers and devices that use the Java<sup>TM</sup> platform.



**Fig. 2.** OSGi Framework [2, 3]

The OSGi framework follows a Service-Oriented Approach. Recent trends marry the notion of service orientation to Web services, but Web services are just but one realization of service-oriented principles based on Web-related standards and protocols. Service orientation is more general and is founded on the concepts of late non-explicit bindings between clients and servers using mechanisms of functionality description and discovery. In service oriented terminology, service providers (i.e., servers) offer services (i.e., functionality) to service requesters (i.e., clients). Service functionality is characterized in a service description that includes some combination of explicitly declared syntax, behavior, and semantics. The service description represents a contract between a service provider and a service requester of the provided service.

Services are discovered using their service descriptions (i.e., non-explicit bindings) at execution time (i.e., late bindings) by service requesters querying a service registry, where service providers publish service descriptions. These key concepts form the quintessential service-oriented interaction pattern.

### 3.2 OSGi Key Features

- Provides the ability to develop software that is much less dependent on the used hardware and operating system.

- Allows the software to be structured as dynamic components which can be installed, started, stopped, updated, and uninstalled without affecting other components being executed within the framework.
- Declarative Services specification that addresses memory footprint issues that can prevent small embedded devices from using service oriented architecture to support multiple applications. Additionally, it significantly simplifies the service-oriented programming model by declaratively handling the dynamics of services.
- Provides a reliable secure environment restricting components to access only the resources and perform only the operations they are allowed to.

OSGi's roots in the embedded device segment has an advantage - it's lean and mean whereas other SOA frameworks are usually "bloated, do-everything frameworks". OSGi has been tried for Vehicular telematics [1]. A survey by the authors establishes that very little work has been done in this area and hence it is a fertile ground for research. In this work, we use this framework for Service orchestration for Telematic services.

### 3.3 OSGi Service Platform – Release 3(OSGi SP-R3)

The OSGi Alliance publishes specifications [3] that define the Service Platform (SP) depicted in Figure 3. The SP was initially targeted at residential Internet gateways with home automation applications. It consists of a small layer above a Java™ Virtual Machine (JVM™) that provides a shared platform for network provisioned components and services. It is shared by different providers of components, potentially across organizational boundaries. It provides an extensive security model and at the same time promotes cooperation and reuse between components. The most attractive features of the platform are its long-running design based on service oriented architecture and its ability to support dynamic updates with minimal perturbation of the running environment.

The OSGi SP-R3 embodies a component and service model. An OSGi component, or bundle, is a set of Java packages containing both classes and resources, essentially what would be traditionally packaged in a JAR (Java archive) file. The difference, however, is that the SP manages such bundles and their dependencies that are expressed as meta-data attached to each bundle.

## 4 Design of Services and Experimental Investigations

In order to obtain the characteristics of an open standard platform, the system had to be designed according to specifications of an OSGi model [3]. A Service-Oriented Architecture (SOA) is a flexible set of design principles, used during the phases of systems development and integration, essentially as loosely coupled collection of services that coordinate activities.

**Stage 1: Determining the Services and the Clients:** Services (Service providers) in the OSGi framework refer to bundles that register and provide services to other bundles. The hardware components in this System each require a dedicated bundle to

implement, register and provide services to enable the employment of these hardware components by client bundles.



**Fig. 3.** General Service System Architecture

Clients (Service users) in the OSGi framework discover and employ services from service providers. Therefore, the applications are written within the bundles of service users. In Service design and modelling, decomposition is the separation of a system into simpler or basic Services. It is often a difficult task to create an optimal decomposition of a system, as any experienced designer would vouch. An optimal decomposition requires gathering all the necessary requirements and then using those requirements to decide upon the appropriate decomposition of systems and services.

**Stage 2: Defining the Services:** For the Design of SOA is to be effective, a clear understanding of what constitutes a Service is required. A service is a function that is well-defined, self-contained, and does not depend on the context or State of other services and can involve both hardware and software resources.

The Services have been divided into two main groups: User Channel Services and Vehicle System Services. User Channel Services refer to the type of loosely coupled services which provide means for data to be transferred over telecommunication mediums. The data can either be transmitted or received from an external source, using Communication Services. Vehicle System Services refer to the type of Embedded Services which provide interaction with components of a vehicle. These services can either be providing information from the components or manipulating the components of the vehicle to provide an Aggregated Complex Service.

**Stage 3: Designing the Interfaces:** OSGi services interact with their clients through an interface. The design of the interface is specific to the intended service that the bundle offers. The interface is equivalent to the definition of the service.

An interface had to be designed for every type of service currently available and generalised to include future services that can be conceived. Bundles that offer the same type of services can reuse the same interface to interact with their clients. These Service interfaces have to be used by the clients to employ the respective services.

**Stage 4: Implementing the Services:** The implementation of each service is based on the interface designed for that particular type of service. Since the client only requires the service's interface to employ the services, the implementation of the service does not need to be exported from the service bundle.



**Fig. 4.** System Services

**Stage 5: Designing the Client Application:** Clients are OSGi bundles that employ services offered by other bundles. Applications within the client's bundle can use these services as part of the application. This contributes to the ever increasing complexity and functionality of the application as well as the dynamic aggregation of the System "On demand".

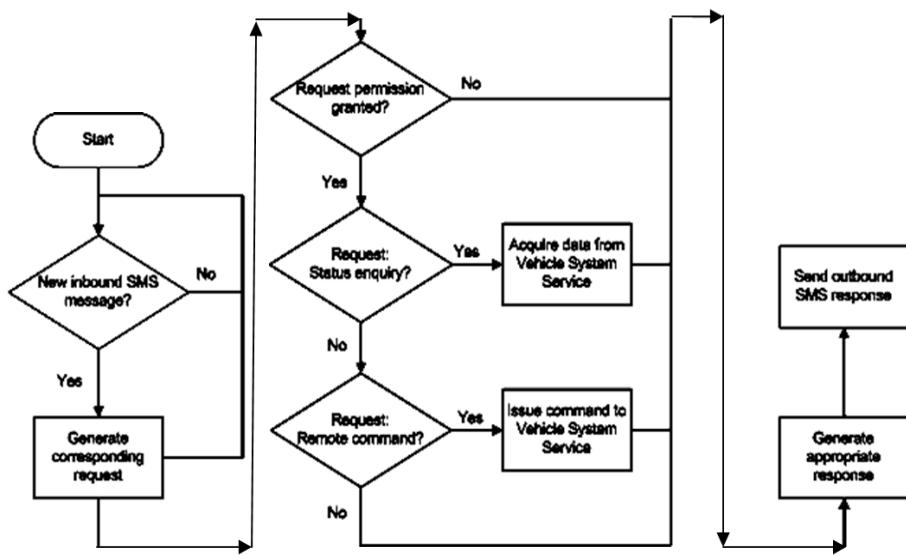
Client bundles can make use of service listeners to constantly be aware of service events taking place. When a service event occurs, the type of event can be determined, followed by the appropriate actions with compliance to the client bundle.

**Design Outcome:** As service-oriented architecture is essentially a collection of services that communicate with each other either through simple data passing or as two or more services coordinating some activity through some means of connection Services, for the actual implementation, a services core and three types of Services were Designed and Deployed: A Short Message (SMS) Communication Service, a Location Service (GPS based) and an Event-Based IO service as explained below.

## 5 Implementation of System Services

**Services Core:** In this Service Implementation, the Core contains the main Service Application. The SMS service, GPS service and the IO service are User Channel Services required to execute the core's application properly. The application processes an inbound SMS message and invokes an appropriate Service (Process) in response. The Core's application first checks the permission level of a request when it arrives. When permission is granted for the particular request, the validity of the request is checked. If the request is valid request, only then the corresponding vehicle system service will be utilized by the core to perform the request. An SMS response is then generated by the core to be sent back to the SMS service for acknowledgement and another once the process/service is complete as an action taken report.

**Service 1 - Short Message (SMS) Service:** The SMS Service developed utilizes a GSM modem with a SIM card to send and receive SMS messages. From an end-user's perspective, this service enables an SMS message to be sent to a client application, and allows the client to send a reply via SMS message. Upon the arrival of an SMS message, the source address (sender's phone number) is checked for the permission level. The permission level is determined by a list of phone numbers that were predefined in this service's implementation. The source address is then stored in a request queue, and an appropriate request id is generated for each new request. The request id, permission level and the body of the text message is visible to clients utilizing the SMS Service for a given time. Each request is kept alive for a set time period, approximately 2 seconds.



**Fig. 5.** System application flow diagrams

After the client has processed the SMS message, if the Service Client produces a response it is forwarded back to this service. The response has to include the request id of the request it is responding to with an “Action taken report”. The SMS service will then look up the request queue for the source address and forwards the report back to the original service requestor.

**Service 2 - Localisation (GPS) Service :** The Localisation (GPS) Service utilizes a serial port that receives data in the form of NMEA sentences [14] from a GPS. This service provides GPS data to the client that can be accessed when a GPS fix is present.

There are two NMEA sentences of interest in this implementation, GGA and VTG. The latitude, longitude, altitude and the time of fix are extracted from the GGA sentence, while the speed is extracted from the VTG sentence [14]. These NMEA sentences are constantly updated and read by the service implementation and sent to the appropriate consumers

**Service 3 - General Purpose Input-Output (GPIO) Services:** The General Purpose IO service developed interacts with Vehicle management System (VMS) through CAN- Bus and Vehicle Modules through the hard wired connections in a vehicle. Modules on today's vehicles are the Engine Control Unit (ECU), the Transmission Control Unit (TCU), the Anti-lock Braking System (ABS) and body control modules (BCM). Vehicle bus is a specialized internal communications network that interconnects components inside a vehicle using Protocols such as Controller Area Network (CAN), Local Interconnect Network (LIN) and others. Conventional computer networking technologies (such as Ethernet and TCP/IP) are rarely used [15]. This

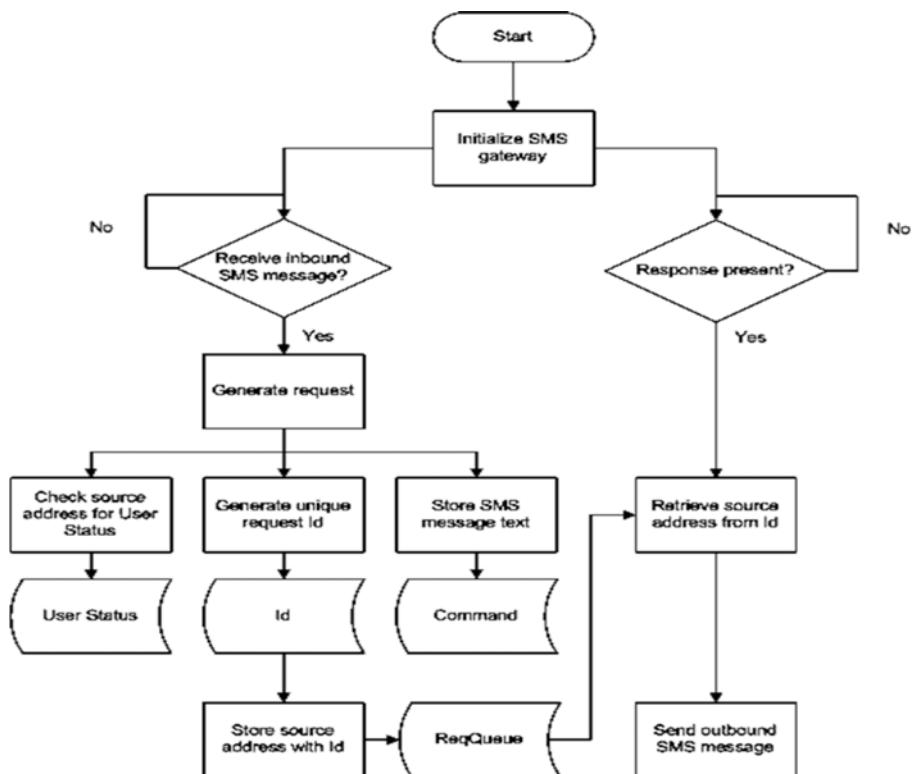


Fig. 6. SMS Service flow diagram

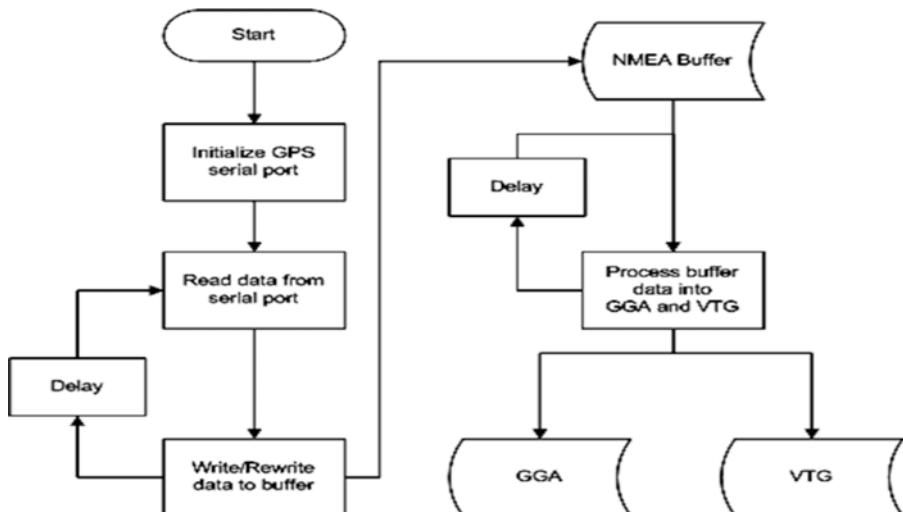


Fig. 7. GPS Service flow diagram

service allows a client to obtain live data of sensors and circuitry in the vehicle, as long as they are hard wired to an appropriate bus. It also provides the client the capabilities of sending outputs through this bus, providing a degree of remote controllability of the vehicle. In this work simulated inputs and outputs were used to replicate the interactions from a field bus (CAN/MOST/etc.) of the vehicle.

## 6 System Implementation and Outcome

The Implemented Experimental setup comprises of three Basic Services that can be consumed independently or which can be aggregated by one client Service. The services are capable of being distributed over TCP/IP networks and accessible for the client. For a demonstration of the system's capabilities three machines, each containing a Java Virtual Machine (JVM) and the developed Services were deployed.

The system's setup is illustrated in Figure 8. Each OSGi container represents a separate machine. More than one OSGi container can be run on one machine, however to illustrate distributed services three machines, containing the IO Service, the SMS Service with the Core and the GPS Service with the GPS Sensor module was deployed. Hardware for the services such as GSM modem for SMS service and a GPS Sensor module for Location info were connected via RS232 and distributed on a wireless local area network through a wireless router.

### 6.1 User Service Applications

A SOA based telematic system capable of receiving and processing service requests through an SMS channel was successfully developed. The system is capable of generating an appropriate response for each of these requests, as well as carry out commands to the vehicle. Also, requests are not case sensitive for the convenience of the user.

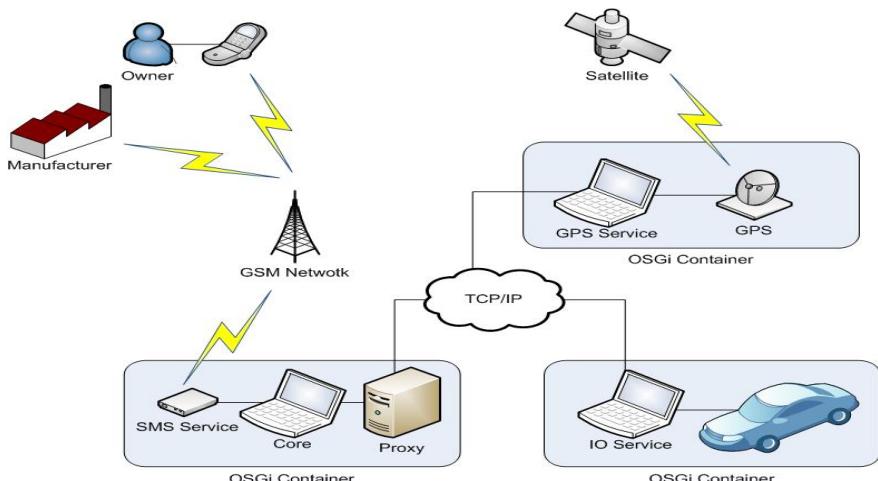


Fig. 8. System Setup diagram

## 6.2 Basic Services for User Applications

One of the basic features of SOA is that primitive services entities that are complete processes by themselves have to be developed to perform basic services so that these entities can be quickly aggregated to form complex and dynamic services.

**Table 1.** Basic Services Request list

Service Request	Description
<b>Status Alarm</b>	Status enquiry of the vehicle's alarm system.
<b>Status Doors</b>	Status enquiry of the vehicle's doors.
<b>Status Windows</b>	Status enquiry of the vehicle's windows.
<b>Status GPS</b>	Status enquiry of the vehicle's position.
<b>Status Speed</b>	Status enquiry of the vehicle's speed.
<b>Doors Lock</b>	Command for locking the vehicle's doors.
<b>Doors Unlock</b>	Command for unlocking the vehicle's doors.
<b>Windows Up</b>	Command for winding up the vehicle's windows.
<b>Windows Down</b>	Command for winding down the vehicle's windows.

## 6.3 Deployment of Services

The Implemented framework was able to deploy and utilize three distinct services that were required for the system's application. The system's GPS and IO service were successfully distributed over TCP/IP.

Since these services are distributed over TCP/IP, this system is capable of monitoring and managing a fleet of vehicles, where each vehicle would contain vehicle specific GPS and IO services. As the system was developed as an open standard platform, new services can be developed and added into the system, while existing services can continue to be updated.

First of the steps was to conceptually realize majority of the service functionality in SOA via abstract end points. Application and services that need to connect to each other and share data are treated as end points. End point could represent a local binding to an application adapter, a callout to an external web service or a single monolithic application such as a legacy Map server. Building a Messaging service through a communication server is the heart of the SOA integration network and has been attempted with some success.

Framework success needs to be measured by rapidity and robustness with which existing systems could be “woven” together with new capabilities such as geospatial data fusion, AGV tracking, event correlation and multi channel event notification.

More over the platform's success needs to be measured by parameters like being vendor agnostic, agile etc. it is needless to mention that testing SOA applications is a non trivial task and can be very challenging.

## 7 Building Complex and Dynamic Services through Aggregation of Basic Services

For example Emergency services (911 services) applications can be quickly aggregated from basic Services to provide more complex event based services. A remote emergency service (Service bundle) can be aggregated and a service bundle deployed on the vehicle dynamically. The Service is then instantiated to provide an emergency Service Application capable of notifying local authorities via the SMS channel when a vehicle's airbag is triggered or Impact Sensor detects an impact. As these applications employ existing services with ease, this telematic system caters to the numerous needs and requirements of the system's users.

## 8 Conclusion

The above prototype implementation shows that a SOA based General Frameworks can be developed. The success of this work has been established through a reliable integration and testing of the hardware components to form a fully functional System. However, it should be emphasized that the design of the framework and system is not restricted only to these hardware components in order to maintain independence of COTS hardware components forming the backbone of the implemented System.

Many other application areas are thinkable. They might require certain changes in the system prototype, especially in the user interface design and implementation, but the general concepts of building a general framework for integration of hybrid telematic entities under the supervisory control of a remote operator provide potential in other areas such as Home Automation, Tele-Maintenance and Tele -Diagnosis, Tele-Learning, Remote Control of Planetary Robots etc remain the same as discussed. The result of pilot work and experimental investigations prove that these services can be plugged into the framework and service functionality realised via abstract end points.

## References

1. Gordon, T., Lay, R., Misener, J., Shladover, S.: Research Needed to Support Vehicle-Infrastructure Cooperation. TRB Committee on Vehicle-Highway Automation (January 2005)
2. OSGi Service Platform, Release 3 Specifications, OSGi Alliance,  
[http://www.osgi.org/resources/spec\\_download.asp](http://www.osgi.org/resources/spec_download.asp)
3. Open Source Initiative, "The Open Source Definition" (2009),  
<http://opensource.org/docs/osd>
4. Li, Y., Wang, F., He, F., Li, Z.: OSGi-based Service Gateway Architecture for Intelligent Automobiles. IEEE, Los Alamitos (2005)

5. Marples, D., Kriens, P.: The Open Services Gateway Initiative: An Introductory Overview. IEEE Communications Magazine (December 2001)
6. Gartner: SOA Will Prevail in Mission-Critical Apps, Technology Letters (April 30, 2007), <http://www.gridtoday.com/grid/1536625.html>
7. Rellermeyer, J.S., Alonso, G., Roscoe, T.: Building, deploying, and monitoring distributed applications with Eclipse and R-OSGI. In: Proceedings of the 2007 OOPSLA Workshop on Eclipse Technology Exchange Montreal, Quebec, Canada, Eclipse 2007, October 21 - 21, pp. 50–54. ACM, New York (2007)
8. Ragavan, V., Ganapathy, V.: A General Telematics Framework for Autonomous Service robots. In: IEEE on Conference Automation Science and Engineering, Scottsdale USA (September 2007)
9. Lewandowski, S.M.: Frameworks for component-based client/server computing. Source ACM Computing Surveys (CSUR) archive 30(1), 3–27
10. <http://java.sun.com/>
11. Hall, R.S., Cervantes, H.: Challenges in building service-oriented applications for osgi. Communications Magazine, IEEE 42, 144–149 (2004)
12. Prinsloo Jaco, M., Schulz Christian, L., Kourie Derrick, G., Theunissen, W.H.M., Strauss, T., Van Den Heever, R., Grobbelaar, S.: A service oriented architecture for wireless sensor and actor network applications. In: SAICSIT 2006: Proceedings of the 2006 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries (2006)
13. Hackbarth, K.: Key to success in telematics – A standardized software platform, TELEMATICS UPDATE- ISSUE 37 SEPTEMBER (2006)
14. SiRF Technology, NMEA Reference Manual (December 2001), <http://www.sirf.com>
15. Cox, R.W.: Local Area Network Technology Applied to Automotive Electronics Communications. IEEE Transactions on Industrial Electronics IE-32(4), 327–333 (1985)

# Manipulating a Micro Stream by ‘Hydro Tweezers’ for Integration of Nanodevices

Mei Liu<sup>1</sup>, Yan Peng<sup>1</sup>, Qiuquan Guo<sup>2</sup>, Jun Luo<sup>1</sup>, and Jun Yang<sup>2</sup>

<sup>1</sup> School of Mechatronics Engineering and Automation, Shanghai University, China

<sup>2</sup> Department of Mechanical & Materials Engineering, University of Western Ontario, London, Canada

mliu@shu.edu.cn

**Abstract.** Integrating synthetic nanowires into functional nanodevices by bottom-up approaches is strategically important for building next generation of nanosystems. However, efficient manipulating of either big amounts of nanowires or individual ones is still a big challenge. Here we present a simple but versatile method of positioning and aligning nanowires by ‘hydro tweezers’: a micro stream hydrodynamically controlled by two sheath flows. By adjusting the flow rates of the sheath flows and the focused micro stream containing nanowires, width, density and position of the assembled nanowire array can be readily tuned using one single design. This method can be used to align and position both a large scale of nanowires and a single line of nanowires to designated positions, offering a general pathway for assembly of nanosystems.

**Keywords:** Hydrodynamic focusing, hydro tweezers, sheath flow, sample flow, nanowire.

## 1 Introduction

Synthetic nanowires (NWs) or nanotubes(NTs) are considered as highly promising building blocks for hierarchical nanodevices because of their unprecedented electronic, mechanical and optical properties and also their associated functions can be easily engineered by controllable doping. Functional NWs have already been integrated into various nanodevices for a wide range of applications such as biosensors, nanoelectronics, optics, optoelectronics, logic circuits, nanoscale solar cells, and so on. However challenges still remain in manipulation of NWs and NTs, for example, alignment of NWs.

In literatures, there are two major categories of techniques for assembling 1D nanostructures: bottom-up and top-down. The conventional top-down approaches mainly based on photolithography have many advantages such as mass production, compatibility with existing semiconductor techniques, as well as precise control of feature sizes and structures. But they are usually limited in diversity of materials and nanofabrication processes. In contrast to top-down approaches, bottom-up methods offer more flexibility in selecting nanowire materials and fabrication processes.

There are already many types of methods for assembling 1D nanomaterials on substrates, such as manipulating the nanomaterials directly with very fine probes or AFM tips [1], photolithography combined with Langmuir-Blodgett technique [2-3], intermolecular interactions method [4], shear force induced assembly [5], in-situ synthesis from surface-bound catalysts [6], or microflows method [7-8].

Though great progress has been made by many groups in developing bottom-up approaches of NW assembly, challenges still remain in fabricating NW based nanodevices in terms of yield, efficiency, reliability and controllability. Particularly, one of the biggest challenges is to position a single NW precisely to a pre-defined location with designated orientation, which is truly one of the most critical steps in fabricating nanoelectronics. And it would be very beneficial if a single NW can be placed between two electrodes skilfully, resulting in easy measurement of electrical property of the NW at a single NW level, since usually the electrical property of NWs must be tested before application.

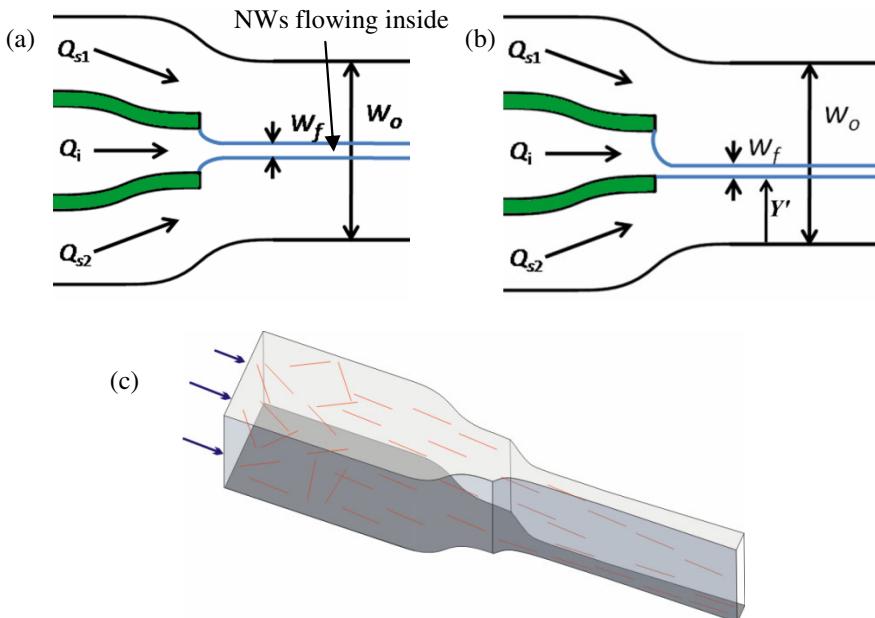
Here we propose a bottom-up approach to manipulate either big amounts of NWs or a single line of NWs by microfluidic ‘hydro tweezers’, which is formed by hydrodynamic focusing with two sheath flows and one sample flow. By adjusting the flow rates of the sheath flows and the sample flow that contains NWs, the width and position of the sample flow can be readily tuned. Further focusing the sample flow, NWs will flow in single file. Thus this approach enables positioning and alignment of both big amounts of NWs or single NWs in a line on a substrate.

## 2 Concepts and Principles

### 2.1 Principles of Two Dimensional Hydrodynamic Focusing

Hydrodynamic focusing is a common phenomenon that when one fluid flows in another fluid, if their velocities and densities are different, they do not mix, instead, they form a two-layer flow, with the sheath flow enveloping the inner flow, like tweezers. It has become one of the most utilized techniques in microfluidics. It has been employed in a wide variety of chemical/biological analyses, including the most popular on-chip flow cytometry for cell/particle counting and sorting [9-11], single molecule detection and measurement [12-13], and production of micro-droplets [14], bubbles [15-17] and microparticles [18].

In this study, hydrodynamic focusing is innovatively applied to assemble NWs in a controllable manner. Fig. 1(a) shows the nozzle section of the hydrodynamic focusing device. In the device, the sample flow from the inlet channel is held laterally by the hydro tweezers—two neighbouring sheath flows from the side channels, as shown in Fig. 1(a) and (b). The flow inside the microfluidic channel is considered to be laminar, and the diffusion and mixing between the focused stream and sheath flows are assumed negligible. Thus, the sample flow and the sheath flows will form a stable three-layer flow configuration, with the sheath flows enveloping the sample flow in a straight line. The width of the focused stream could be controlled through adjusting the flow ratio of the sheath flow to the sample flow. The position of the focused stream could be controlled through adjusting the flow ratio of the two sheath flows.



**Fig. 1.** The schematic of hydrodynamic focusing and NW assembly. (a) Symmetrical hydrodynamic focusing.  $Q_i$  is the volumetric sample flow rate; the sheath flow  $Q_{s1}$  is the flow rate in side channel 1, and  $Q_{s2}$  is the flow rate in side channel 2;  $w_f$  and  $w_o$  are the width of the focused stream and the outlet channel respectively. (b) Non-symmetric hydrodynamic focusing.  $Y'$  is the share one sheath flow occupies. (c) Side view of the focused stream showing NWs flowing in the sample stream and depositing on the device substrate.

In this study, the aspect ratio is much smaller than 1, so a parabolic velocity profile will be formed across the channel height and is independent of the position across the channel width (i.e. the velocity profile is ‘plug-like’ across the channel width). Under such conditions, based on the mass conservation law, in hydrodynamic focusing, it is calculated that width of the focused stream is proportional to the width of the outlet channel  $w_o$  and the share of the sample flow in the whole flow. The space each sheath flow occupies in the outlet channel is proportional to  $w_o$  and its share in the whole flow. In a word, the widths of the sheath flows and sample flow in the outlet channel are all proportional to  $w_o$  and their respective shares in the whole flow rate.

## 2.2 Principles of NW Assembly

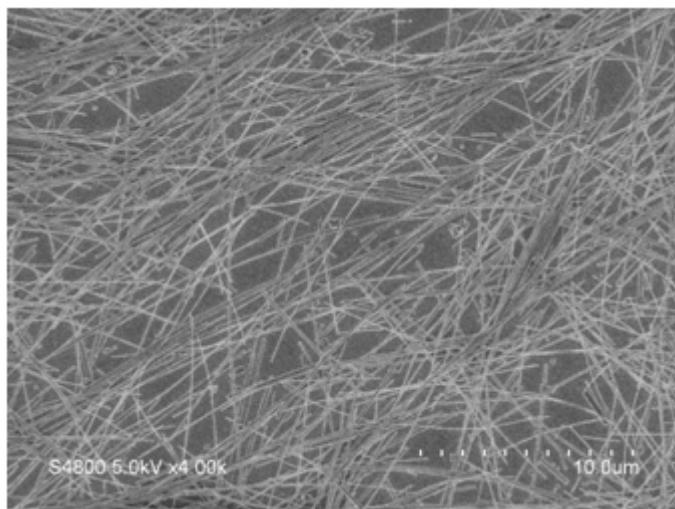
NWs assembly was carried out by injecting a NW suspension as the sample flow ( $Q_i$  in Fig. 1(a)), while the other fluid ( $Q_{s1}$  and  $Q_{s2}$  in Fig. 1(a)) without NW as the sheath flows. Since microscopic flows almost always occur at low Reynolds number ( $Re$ ), and are laminar and stable, therefore, NWs injected into the sample flow follow well defined streamlines along the flow direction. Shear forces in the flow aligns the NWs in the flow direction (shown in Fig. 1(c)) before they are immobilized on the substrate. The substrate is chemically treated so that the NWs in the sample flow tend to

reside on it. Since only the sample flow contains NWs, only the focused-stream covered area will be coated with NWs. By tuning the flow duration and flow rates of the sample flow and sheath flows, the width, density and positions of the NW array can be readily tuned.

### 3 Experiments

#### 3.1 Silver Nanowires Synthesis and Purification

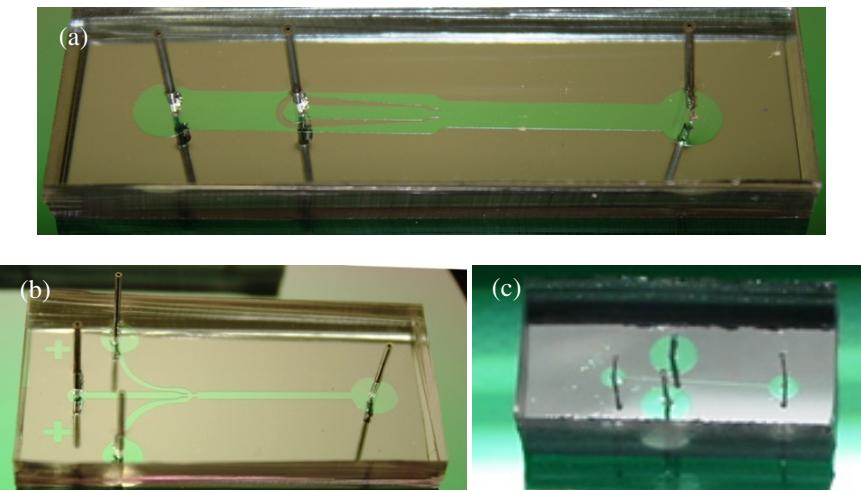
Silver nanowires (AgNWs) was selected as the test nanomaterial since they can be made in a relatively high yield and uniformity. AgNWs were synthesized with a solution-based method by reducing  $\text{AgNO}_3$  with EG (Ethylene glycol, Sigma Aldrich) in the presence of PVP (Polyvinylpyrrolidone, Sigma Aldrich). This process produced 100% pure AgNWs with a concentration of less than 0.8 wt%. Fig. 2 shows the synthesized AgNWs with diameters around 60 nm and lengths of about 10  $\mu\text{m}$ .



**Fig. 2.** Synthesized AgNWs under SEM

#### 3.2 Device Design and Fabrication

Three types of microfluidic device with the same nozzle shape were fabricated for hydrodynamic focusing: symmetrical hydrodynamic focusing, non-symmetric hydrodynamic focusing, and a size-reduced device for non-symmetric hydrodynamic focusing. Fig. 3 shows the fabricated devices on silicon substrates. The microfluidic devices were fabricated using the now popular soft lithography technique. Holes were punched at the center of the reservoir area and the reliefs were sealed to a silicon wafer to form a closed channel. The devices were connected to syringe pumps and a waste beaker by plastic tubing.



**Fig. 3.** The fabricated devices on silicon substrates. The three figures were drawn to different scales. (a) The symmetrical hydrodynamic focusing device, for this device,  $w_o=2.4$  mm; (b) The non-symmetric hydrodynamic focusing device,  $w_o=1.0$  mm; (c) The size-reduced hydrodynamic focusing device,  $w_o=0.1$  mm.

### 3.3 The NW Assembly Process

Before put into experiment, the microfluidic devices were treated so that AgNWs tend to residue on their surface. For all those three devices, the ethanol solution with suspended AgNWs was injected from the center channel and DI water was injected into the sheath channels by external pumps. As expected, when an appropriate flow protocol was applied, the sample solution and the DI water sheath flow forms a stable two-phase flow, with the sample flow enveloped in the sheath flow.

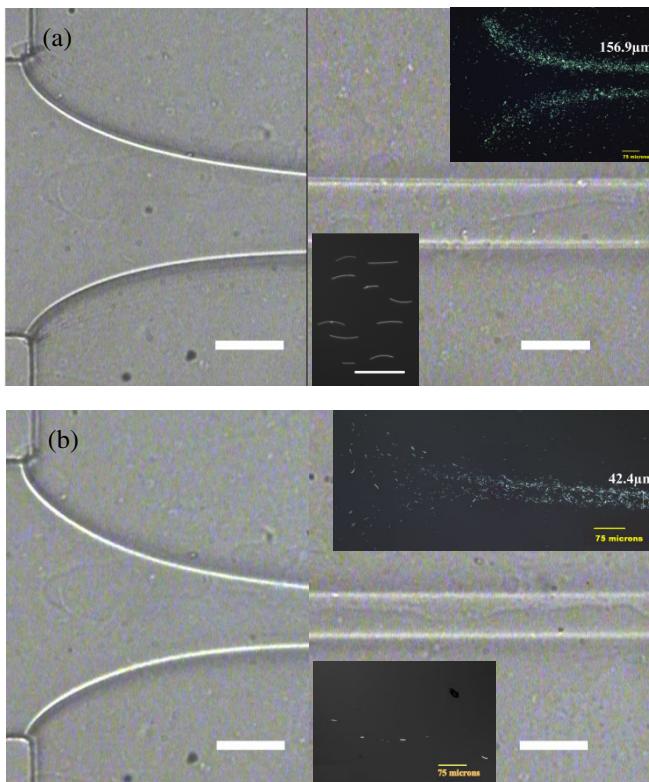
#### 3.3.1 Symmetrical Hydrodynamic Focusing Process

For the symmetrical hydrodynamic focusing device in Fig. 3(a), two specific protocols were applied to study the flow behaviour and the AgNW assembly process, as shown in Fig. 4.

In this process, the ethanol solution containing AgNWs was injected into the inner channel as the sample flow, and DI water was injected into the outer channel as the sheath flow. The sample flow was hydrodynamically focused into a narrow stream constrained by the tweezers (shown in Fig. 4). It was obvious in the experiment that as the sheath flow increases over the sample flow, the focused sample stream will be narrowed down. While flowing in the channel, AgNWs would reside and align on the substrate due to the interaction between AgNWs and the treated  $\text{SiO}_2$  surface. In the experiment, the focused stream could be tuned from up to 2.4 mm to down to 40  $\mu\text{m}$ .

After 20min flow durations, the flow is stopped and the assembled NW pattern was checked. In Fig. 4(a), the upper inset figure shows the whole AgNWs array, and the bottom one shows part of the assembly results at high magnification. In Fig. 4(b), the

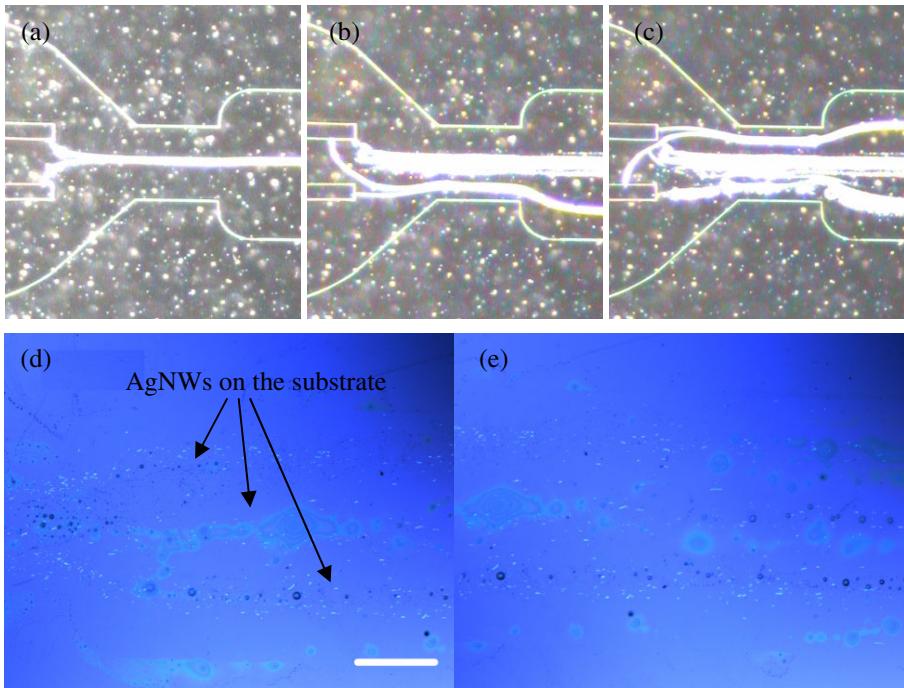
upper inset figure shows the assembly results of sample solution with high AgNWs concentration, the bottom one shows the results of very dilute AgNWs solution.



**Fig. 4.** Hydrodynamic focusing in the microfluidic device and AgNWs assembled on the substrate surface. On the left is the focused stream close to the nozzle; on the right is the steady stream away from the nozzle. (a) Flow Protocol 1, sheath flow  $Q_{s1}+Q_{s2}=1.62$  mL/h, sample flow  $Q_i=0.108$  mL/h; (b) Flow Protocol 2,  $Q_{s1}+Q_{s2}=2.70$  mL/h,  $Q_i=0.108$  mL/h. The scale bar in the flow figures were 150  $\mu\text{m}$ . The scale bars in the inset figures were 75  $\mu\text{m}$ , except the bottom one in (a), whose scale bar was 40  $\mu\text{m}$ . The outlet channel width  $w_o=2.4$  mm.

### 3.3.2 Non-symmetric Hydrodynamic Focusing Process

In this section the unequal sheath flows were applied in the flowing system, using the device shown in Fig. 3(b). By manipulating the two sheath flows, the focused sample stream could be shifted across the whole microchannel hydrodynamically, as shown in Fig. 5. As one sheath flow increases, it occupies more in the microchannel outlet. By varying the flow rate ratio of two sheath streams on either side of the sample stream, a controlled displacement of the latter happened towards the side of the chamber where the flow was reduced. Thus, AgNWs carried in the sample stream were deposited on functionalized chip surfaces as discrete wide lanes.

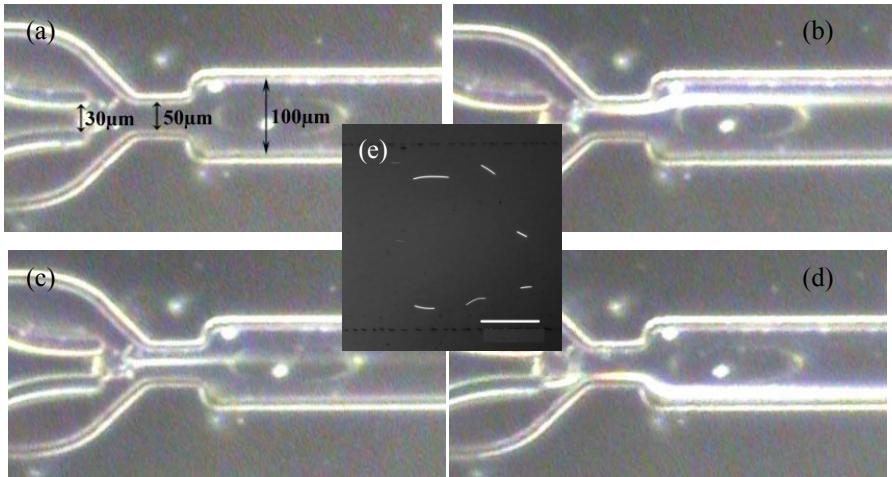


**Fig. 5.** Non-symmetrical hydrodynamic focusing for multiple AgNWs stripes assembly. (a) Step 1,  $Q_{s1}=Q_{s2}=1.0 \text{ mL/h}$ ,  $Q_i=0.003 \text{ mL/h}$ ; (b) Step 2,  $Q_{s1}=1.6 \text{ mL/h}$ ,  $Q_{s2}=0.4 \text{ mL/h}$ ,  $Q_i=0.003 \text{ mL/h}$ ; (c) Step 3,  $Q_{s1}=0.4 \text{ mL/h}$ ,  $Q_{s2}=1.6 \text{ mL/h}$ ,  $Q_i=0.003 \text{ mL/h}$ . (d) The assembled AgNWs array at the nozzle section. (e) The assembled three AgNWs array beyond the nozzle. The scale bar in (d) and (e) were  $300 \mu\text{m}$ . The outlet channel width  $w_o=1.0 \text{ mm}$ .

### 3.3.3 Size-Reduced Hydrodynamic Focusing Process

In addition to controllable alignment of big amounts of NWs, the hydrodynamic focusing approach can also align a single line of AgNWs by a size-reduced device, as shown in Fig. 3(c), which was experimentally confirmed in this section, as shown in Fig. 6. The new device has the size of about 10% of the previous non-symmetrical one (Fig. 3(b)), which acts like highly precise ‘hydro tweezers’, potentially being able to manipulate a single NW skilfully, comparable to AFM tips.

Based on calculations, the focused stream width is proportional to the outlet channel width. With the outlet channel size reduced, at the same condition, the focused stream will be reduced to 10% of the previous device. So the AgNWs contained in the focused stream will be largely reduced. And since the aspect ratio is far smaller than 1, the focused width will not be affected by its position in the channel, resulting in possible alignment of single NW. Preliminary results on single lines alignment of NWs have been obtained by this approach, as shown in Fig. 6(e).



**Fig. 6.** Non-symmetrical hydrodynamic focusing for single AgNWs stripes assembly. (a) Nozzle and dimension of the device; (b) Step 1,  $Q_{s1}=0.05\text{mL/h}$ ,  $Q_{s2}=0.20\text{mL/h}$ ,  $Q_i=0.0012\text{mL/h}$ ; (c) Step 2,  $Q_{s1}=Q_{s2}=0.15\text{mL/h}$ ,  $Q_i=0.0012\text{mL/h}$ ; (d) Step 3,  $Q_{s1}=0.20\text{mL/h}$ ,  $Q_{s2}=0.05\text{mL/h}$ ,  $Q_i=0.0012\text{mL/h}$ . (e) Three assembled single lines of AgNWs. The scale bar in (e) was 30  $\mu\text{m}$ .

## 4 Results and Discussions

Assembly results in Fig. 4-Fig.6 show that virtually all the NWs were aligned along the flow direction, though there were also some small deviations with respect to the flow direction, which were caused by instabilities during the flowing process.

An examination of the assembled NWs on larger length scales shows that the aligned NW array readily extends over hundreds of micrometers. Indeed, alignment of the NWs has been found to extend up to millimeter length scales and seems to be limited by the size of the fluidic channels only. Thus the length of the NW array can be easily chosen by setting length of the microchannel.

In the experiment, there are several factors controlling the alignment and average separation of the NWs: (1) the flow rate. With increasing flow rates, the deviation of the NWs' alignment with respect to the flow direction substantially becomes narrower. Higher flow rates produce larger shear forces and hence lead to better NW alignment. (2) The NW deposition coverage can be controlled by the flow duration. Experiments carried out at given operating conditions showed that the NW density increases systematically with increase of the flow duration. (3) The deposition rate and hence the average NW separation distance versus time also depend strongly on the surface chemical functionality. (4) NW concentration in the solution. With higher density of NWs, the assembled AgNW arrays become denser. Diluting NW concentration to an optimal value is important to obtain the line alignment of single NWs, as shown in Fig. 4(b). When the AgNWs solution is much diluted, one single line of NWs was obtained by the symmetrical hydrodynamic focusing device.

The fluidic hydrodynamic focusing approach, not like AFM tips manipulation or the photolithography method, the AgNWs held in the ‘hydro tweezers’ remains in the growth solution, and do not get into contact with any hard surfaces or strengths, thus the AgNWs are well protected and will keep their original shape and properties.

## 5 Conclusions

Assembly and alignment of both large-scale NW arrays and single lines of AgNWs have been demonstrated using the ‘hydro tweezers’, creating both big amounts of AgNWs arrays and parallel stripes, and also discrete lines of single AgNWs. The hydro tweezers offered many advantages among the bottom-up approaches, such as the unmatchable controllability, small sample consumption, high speed and yield, its suitability to a wide range of NWs, and good protection of the assembled NWs. The well aligned AgNWs can be employed in various highly sensitive sensors, making the best of the special characteristics of nanomaterials.

## Acknowledgement

The publication of this paper was supported by the National High Technology Research and Development Program of China (863, No. 2009AA04Z211) and the Innovation Program of Shanghai Municipal Education Commission (No. 09ZZ091).

## References

1. Zyvex Instruments, <http://www.zyvex.com/>
2. Whang, D., Jin, S., Wu, Y., Lieber, C.M.: Large-Scale Hierarchical Organization of Nanowire Arrays for Integrated Nanosystems. *Nano. Lett.* 3, 1255–1259 (2003)
3. Park, J., Shin, G., Ha, J.S.: Controlling Orientation of  $V_2O_5$  Nanowires Within Micropatterns via Microcontact Printing Combined with the gluing Langmuir-Blodgett Technique. *Nanotechnology* 19(395303), 6 (2008)
4. Heo, K., Cho, E., Yang, J.E., et al.: Large-Scale Assembly of Silicon Nanowire Network-Based Devices Using Conventional Microfabrication Facilities. *Nano. Lett.* 8, 4523–4527 (2008)
5. Fan, Z., Ho, J.C., Jacobson, Z.A., et al.: Wafer-Scale Assembly of Highly Ordered Semiconductor Nanowire Arrays by Contact Printing. *Nano. Lett.* 8, 20–25 (2008)
6. Sławiński, G.W., Zamborini, F.P.: Synthesis and Alignment of Silver Nanorods and Nanowires and the Formation of Pt, Pd, and Core/Shell Structures by Galvanic Exchange Directly on Surfaces. *Langmuir* 23, 10357–10365 (2007)
7. Huang, Y., Duan, X., Wei, Q., et al.: Directed assembly of one-dimensional nanostructures into functional networks. *Science* 291, 630–633 (2001)
8. Cui, Y., Wei, Q., Park, H., Lieber, C.M.: Nanowire Nanosensors for Highly Sensitive and Selective Detection of Biological and Chemical Species. *Science* 293, 1289–1292 (2001)
9. Huh, D., Gu, W., Kamotani, Y., et al.: Microfluidics for flow cytometric analysis of cells and particles. *Physiol. Meas.* 26(3), 73–98 (2005)
10. Wang, M.M., Tu, E., Raymond, D.E., et al.: Microfluidic sorting of mammalian cells by optical force switching. *Nat. Biotechnol.* 23(1), 83–87 (2005)

11. Wolff, A., Perch-Nielsen, I.R., Larsen, U.D., et al.: Integrating advanced functionality in a microfabricated high-throughput fluorescent-activated cell sorter. *Lab Chip* 3(1), 22–27 (2003)
12. de Mello, A.J., Edel, J.B.: Hydrodynamic focusing in microstructures: Improved detection efficiencies in subfemtoliter probe volumes. *J. Appl. Phys.* 101(8), 84903 (2007)
13. Wang, T.H., Peng, Y., Zhang, C., et al.: Single-molecule tracing on a fluidic microchip for quantitative detection of low-abundance nucleic acids. *JACS* 127(15), 5354–5359 (2005)
14. Anna, S.L., Bontoux, N., Stone, H.A.: Formation of dispersions using “flow focusing” in microchannels. *Appl. Phys. Lett.* 82(3), 364–366 (2003)
15. Gañán-Calvo, A.M.: Perfectly monodisperse microbubbling by capillary flow focusing: An alternate physical description and universal scaling. *Phys. Rev. E* 69(2), 27301 (2004)
16. Gañán-Calvo, A.M., Gordillo, J.M.: Perfectly monodisperse microbubbling by capillary flow focusing. *Phys. Rev. Lett.* 87(27), 274501 (2001)
17. Garstecki, P., Gitlin, I., Diluzio, W., et al.: Formation of monodisperse bubbles in a microfluidic flow-focusing device. *Appl. Phys. Lett.* 85(13), 2649–2651 (2004)
18. Martín-Banderas, L., Flores-Mosquera, M., Riesco-Chueca, P., et al.: Flow focusing: A versatile technology to produce size-controlled and specific-morphology microparticles. *Small* 1(7), 688–692 (2005)

# **Multi-modality — EMG and Visual Based Hands-Free Control of an Intelligent Wheelchair**

Lai Wei and Huosheng Hu

School of Computer Science & Electronic Engineering, University of Essex  
Wivenhoe Park, Colchester CO4 3SQ, United Kingdom  
lwei@essex.ac.uk, hhu@essex.ac.uk

**Abstract.** This paper presents a novel human machine interface for people with severe disabilities to control an electric powered wheelchair using face movements. Five face movements including jaw clenching and eye closing movements are identified by extracting movement features from both forehead Electromyography (EMG) signal and facial image information. A real-world indoor environment is setup for evaluating the performance of the new control method. Five subjects participated in the experiment to follow designated routes on a map using the new control method, as well as a traditional joystick control respectively. Comparison of two control methods are made in terms of easiness of control, time duration, wheelchair trajectory and error command rate etc. Wheelchair trajectory and time consumption are recorded for each task and results show the new control method are comparable to a joystick control and can be used as a hands-free control interface for disabled and elderly users.

**Keywords:** Intelligent wheelchairs, eye detection, multi-modality, hands-free control, EMG, HMI, HCI.

## **1 Introduction**

Since most of the human machine interface (HMI) or human computer interface (HCI) used today are operated by hand. For people with sever physical disabilities such as spinal cord injury, quadriplegia, hemiplegia or amputation or elderly people, it is necessary to develop hands-free control interfaces. Apart from hand movements, a number of other human movements can be deployed for HMI or HCI purposes, such as eye movement, limb movement, neck movement, jaw movement and tongue movement. For decades, these movements are intensively studied for the control of rehabilitation devices such as mobile wheelchairs, prosthetics and assistant computers etc., and for most of people who have severe disability; face movement may be their last possible movements to be used.

In the past decade, a number of simple yet effective hands-free HMIs are brought into applications such as sip & puff switch, head infrared pointer and head touch switch buttons and these interfaces are widely used for their simplicity to control and easiness to setup. At the same time, there have been some studies on developing HMIs based on human physiological signals such as electromyography (EMG), electrooculography (EOG) and electro-encephalography (EEG). As can be seen in literature [1-5],

HMIs developed from these signals are used for hands-free control of electric-powered wheelchairs. With the recent advancement of machine and computing ability, many HMIs have been developed from computer vision information [11], speech information and multi-modality phenomenon. Ju and Shin et al. introduce a vision-based intelligent wheelchair system developed for people with quadriplegic [6]. Li and Tan [7] propose a bimodal wheelchair control approach by integrating vision and speech controls. Matsumoto and Ino et al. [8] apply the recognition of head motion and eye gaze onto a locomotive wheelchair system. Ferreira and Silvaaet al. [9] proposed an HMI structure to control a robotic wheelchair by scalp EMG and EEG signals. Both eye blinking movements and eye close movements are used as command movement to control a mobile wheelchair through an onboard PDA.

This paper presents a new hybrid HMI developed to control an electric powered wheelchair via face movements resolved from forehead EMG signals and facial image information. To test the performance of this new control strategy, an indoor experimental environment is setup and five subjects are designated with some given tasks to follow a route on a designed map with either the joystick control or the face movement control in turn. The performance of the new method is evaluated by comparing control results with the joystick control in terms of time consumption, wheelchair trajectories and subject personal experience.

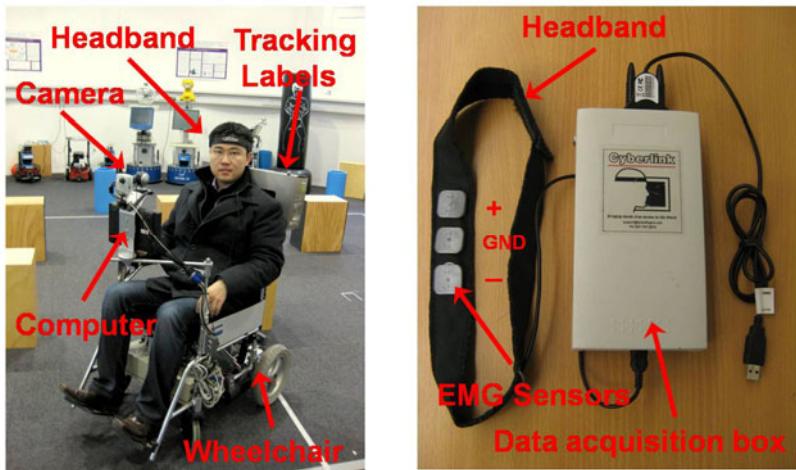
The rest of the paper is organized as follows. Section 2 describes the system architecture which consists of three parts: a wearable sensory device Cyberlink<sup>TM</sup> [10]; an Intelligent Wheelchair system and the HMI. Control movement description, feature selection procedure and control strategy are presented in Section 3. Section 4 shows the designed wheelchair experiment environment and real-world experiment results between two control methods. Finally, a brief summary and potential future extension of the system are given in Section 5 and an appendix shows recorded wheelchair trajectories of five subjects is attached thereafter.

## 2 System Architecture

The proposed experimental system contains three parts. The first part is the data acquisition device Cyberlink<sup>TM</sup> [10], which is composed of 1) a data processing box and 2) a wearable headband. The second part of the system is an intelligent wheelchair platform which is used for evaluating real-world performance of the control system. The third part is the human machine interface which is responsible for extracting features and classifying selected movement patterns and mapping the classified patterns into wheelchair control commands; Details of these parts will be presented in the following paragraphs.

### 2.1 Sensory Data Acquisition

The sensory system for detecting subject facial movements, as shown in Fig. 1, contains an EMG signal acquisition device Cyberlink<sup>TM</sup> and a Logitech S5500 web camera which is mounted at the front of the wheelchair pointing to wheelchair user's face as shown in Fig. 1 (Left).



**Fig. 1.** Intelligent wheelchair setup for a subject and indoor experiment environment (Left); The Cyberlink<sup>TM</sup> headband and data acquisition device (Right)

Fig. 1 (Right) shows the composition of EMG device which has a data processing box and a wearable headband with three flat attachable electrodes. The electrodes are attached to human forehead by fastening the headband to user's forehead. From three electrodes, one channel EMG signal can be obtained with one positive electrode, one negative electrode and a reference electrode. The raw EMG amplified from the electrodes are then processed in data acquisition box and turned into digital averaged EMG data sequence with a sampling frequency of 100Hz;

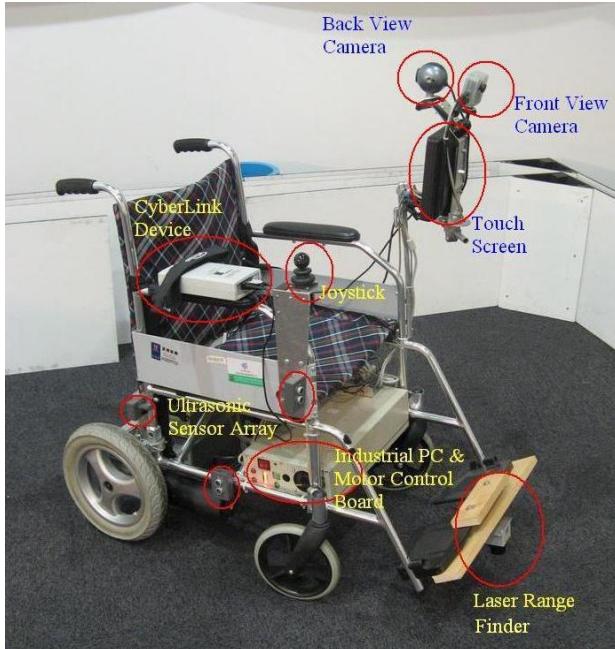
## 2.2 The Intelligent Wheelchair

As shown in Fig. 2, the intelligent wheelchair system employed in this experiment is an electric-powered mobile wheelchair equipped with an industrial computer and a built-in motor control board. The motor control board is embedded with a digital signal processing (DSP) unit which can implement open-loop liner motor control or closed-loop PID motor control by controlling two differentially-driven wheelchair wheels; the wheelchair is mounted with a laser range finder and sonar sensor array, from which the onboard computer can calculate real-time range information to avoidance obstacles such as boxes, people, walls etc. In this experiment, obstacle avoidance function has been disabled.

## 2.3 Human-Machine Interface (HMI)

The human machine interface is responsible for extracting features and classifying selected face movements and mapping the classified movement patterns into wheel-chair control commands; detailed procedure of this part is graphically depicted in Fig. 4, as shown, from the EMG sensor, five designed jaw and eye movements are classified into different patterns and with the eye close information detected from the camera, five face movements are identified and these movement patterns are mapped into six wheelchair control commands i.e. "GO Forward", "Go Backwards", "Turn Left",

“Turn Right”, “Reduce Speed” and “Stop”, which imitates direction and speed control of a joystick. The detailed pattern recognition process and control strategy mapping procedure including data segmentation, feature extraction, classification and logic control will be discussed further in Section 3.



**Fig. 2.** Intelligent wheelchair platform employed in the experiment

### 3 Control Movement and Pattern Recognition

#### 3.1 Control Movement Selection

Five designated face movements employed in this experiment are named as single jaw clenching (SJC), forehead double jaw clenching (DJC), left eye close (LEC), right eye close (REC) and continuous jaw clenching (CJC) movement. Detailed motion of the five selected control movements are described as follows:

As shown in Fig.4, SJC is a heavy jaw clenching movement which can be produced by the subject holding up their teeth and pushing the jaw against upper teeth and making a strong contraction of masseter and buccinators muscles with a fast jaw clenching movement. Similar to teeth holding up motion in SJC movement; CJC can be produced by a subject making gentle and repetitive clenching, compared with SJC clenching, CJC is gentle and persistent jaw clenching movement.

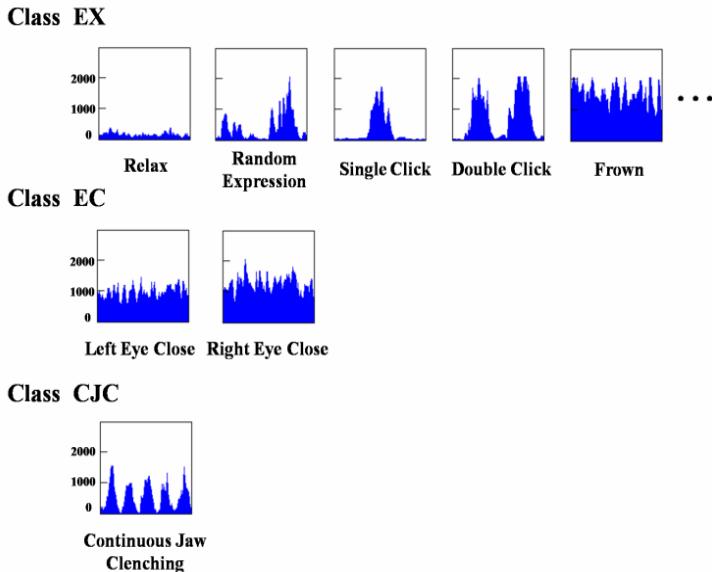
As shown in Fig. 3 and Fig. 4, CJC movement can cause a series of short and regular EMG waveforms while SJC arouse a sharp and high waveform. DJC is made of two consecutive SJC movements. LEC and REC movements are one eye open and

one eye close movements, starting from eye closing movement of both left and right eyes. The movement state can last by keeping the eye closed. Unlike gentle eye closing movements during blinking or sleeping, LEC and REC are heavier muscle movements and can be described as eye squeezing movements. The finishing state of the movement is marked by opening the closed eye.

### 3.2 EMG Pattern Classifications

In order to recognize five face control movements described above, features from EMG signals and face image information are jointly extracted and classified. Fig. 3 shows EMG waveform patterns obtained during corresponding control movements. These EMG patterns are classified with two separate procedures i.e. SJC-DJC procedure and CJC-EC procedure.

In the SJC-DJC procedure, SJC and DJC are recognized by a threshold based strategy. SJC is detected by calculating uninterrupted EMG overshoot amplitudes when EMG amplitude is continuously above a preset threshold. If the length of time falls within a preset range, then SJC patterns are detected. This strategy is also implemented by Torsten and Bernd in [1], which uses SJC movements to control a powered wheelchair.



**Fig. 3.** Class division for EMG movement pattern classification, movement patterns are displayed in 50 samplings (500ms)

In the second procedure, support vector machine (SVMs) classifier are trained by using training data collected from a standard subject performing various face movements including CJC, LEC, REC, SJC, DJC, relaxing, talking, smiling etc. These movement patterns are assembled into three classes which are Class EC, Class

CJC and Class Exclude (EX). As shown in Fig. 3, Class EC contains two patterns which are EMG patterns produced during LEC and REC movements. Class CJC contains a single pattern which from CJC movements. And Class EX contains all the other patterns other than the patterns in EC and CJC classes. Class EX contains a range of EMG patterns such as SJC and DJC movement patterns, patterns from facial movements when subject is relaxing, talking, smiling etc.

Training data are sampled according to the movement patterns described in these three classes, and the recorded movements' data are therefore segmented into 200ms sections for feature extraction purpose. Eight waveform features, including five from time domain which are mean absolute value (MAV), root mean square (RMS), waveform length (WL), zero crossing (ZC) and slope sign changes (SSC) and three from frequency domains which are frequency mean (FMN), frequency median (FMD) and frequency ratio (FR), are extracted from each data segment. With extracted features from three classes, a RBF kernel SVM classifier is selected, which gives a best classification accuracy of 93.5% on a threefold cross validation procedure.

### 3.2 Face Tracking and Eye Close Recognition

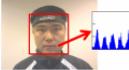
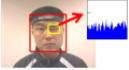
Alongside with classified EMG patterns, two face image patterns denoting LEC and REC movements are also classified with an adaptive Boosting learning approach on Haar-like features. As shown in Fig. 3, two classifiers for detecting left-closed eye and right-closed eye are trained. The training is based on sampling closed-eye images from 5 different subjects under various illumination conditions, changing subject face orientations and distance relative to the sampling camera.

As a result, a number of 1475 closed-eye images are collected for left eye training, and for the right eye, and the positive image set has 1318 closed eye images. Based on these images, a left-closed eye classifier with 20 cascades is trained, and the classification accuracy validated among sample training images is 91.75%. For the right-closed eye, the training process stops at 19 cascades with a classification validation accuracy of 92.2%.

### 3.3 Decision Fusion and Control Strategy

Fig. 4 describes the decision fusion process for classified EMG and visual patterns and wheelchair control commands. Six control commands namely “Go Forward” (GF), “Turn Left” (TL), “Turn Right” (TR), “Reduce Speed” (RS), “Stop” (ST) and “Go Backwards” (GB) are employed to mimic and replace an ordinary joystick control.

As depicted in Fig. 4, the wheelchair control logic can be described as follows: wheelchair will go forward and accelerate its speed when subject making CJC movement and will turn left or right when the subject is closing his right or left eye with LEC or REC movements. The wheelchair will keep on turning until the closed eye is opened, in this way subject can control how much angle the wheelchair has turned. If a SJC movement is detected, the wheelchair will drop its current speed into a safe speed and going forward. DJC movement will cause the wheelchair to stop and another DJC movement when the wheelchair has stopped will make wheelchair going backwards.

Facial Movements	EMG Patterns	Image Patterns	Control Commands
	Continuous Jaw Clench (CJC)	Any	 Go Forward (GF)
	Eye Close (EC)	Left Eye Close (LEC)	 Turn Right (TR)
	Eye Close (EC)	Right Eye Close (REC)	 Turn Left (TL)
	Single Jaw Clench (SJC)	Any	 Reduce Speed (RS)
	Double Jaw Clench (DJC)	Any	 Stop (ST)
	Double Jaw Clench (DJC)	Any	 Go Backwards (GB)

**Fig. 4.** Diagram showing sample face control movements and control strategy of EMG and visual based control methods

## 4 Real-World Experiment Setup and Results

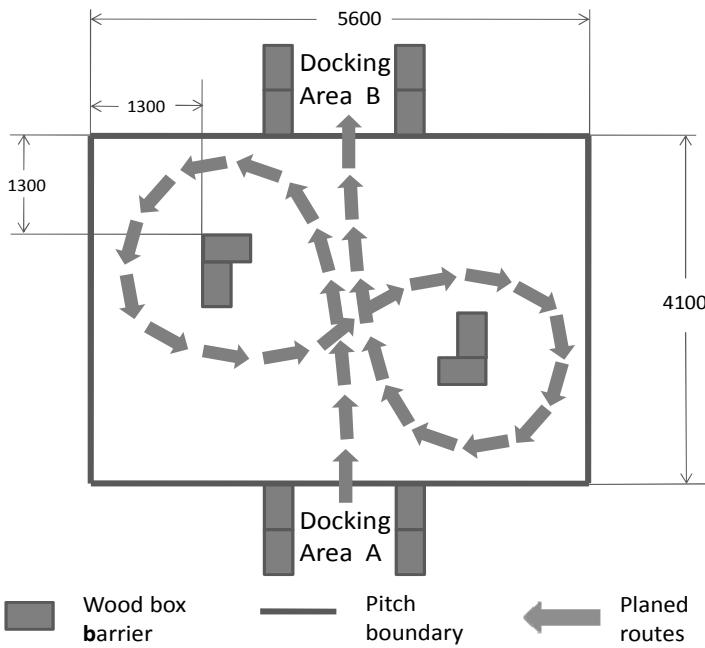
### 4.1 Experiment Setup

To test the performance of the hybrid (EMG and Visual based) control method, an indoor experiment environment is setup. Fig. 5 shows the schematic configuration of the map, as can be seen, the size of the map is 5600×4100mm and the actual size of the wheelchair is 1200×800mm. The environment is supplied with ample light intensity which refers to an indoor fluorescents illumination condition.

Five intact subjects are involved in the experiment and each subject are asked to follow a designed route (depicted in Fig. 5) from Docking area A to Docking area B using either joystick control or hybrid control method. There are ten tasks from A to B to be completed for each control method. For each designed task, the time consumption and the trajectory of the wheelchair are recorded. The position of the wheelchair is tracked via a VICON tracking system by attaching identifiable markers to the wheelchair (as shown in Fig.1 left).

During the experiment, we found that the layout and size of the map can have influence on the experiment results with respect to different control methods. In order to apply equal rules for both control methods, the map is designed to have adequate space for the user to maneuver , and each task the routes is designed for made up of

full control functions including two alternatively left and right turnings, following strait lines and docking into narrow places. For the hybrid control method, the wheel-chair speed (going forward and turning) is tuned for each subject to an optimized condition so that the subject can have a better performance with the control method on the map. As the joystick control interface is a mature and universal, the acceleration, angular and linear velocities are adjustable from zero to the maximized wheel-chair carrying ability.



**Fig. 5.** Map designed for indoor experiment and the route for subjects to follow during wheelchair control tasks

## 4.2 Experiment Results and Analysis

Fig. 6 to Fig. 10 in appendix show experimental results for five subjects, namely from subject A to subject E. Each diagram shows twenty recorded time durations corresponding to twenty tasks in which ten tasks are finished by the hybrid control and other ten by the joystick control. As can be seen, for all the five subjects completing the same tasks, joystick control took less time than facial movement based control. we can see that compared with the new hybrid control method, joystick control is more smooth and steady while the new hybrid control method costs more time and the performance between tasks varies considerably with the same subject.

Fig.6 and Fig. 10 show that for some subjects, the new control method presents a relatively inferior performance. The individual difference on EMG features and image

features can affect the performance and make the new control method unsteady and difficult to handle. A control experience survey from five subjects also show that the new method is more difficult to manipulate compared with traditional joystick control especially during persistent turning actions; subjects feel fatigue when making eye closing and squeezing movement for up to ten seconds.

## 5 Conclusion and Future Work

This paper provides a solution for a face movement based wheelchair control interface that combines features from both forehead EMG data and facial images. Five subjects participated in the test experiments in order to evaluate the performance of the new interface.

From the experiment results, we conclude that: (1) Computer vision and facial EMG based control can be effectively used in real-world real-time control due to their fast respond speed, but its control complexity, error rate and muscle fatigue should be improved; (2) Although the trajectories resulted from the face movement based control is more sparse and irregular compared with the ones resulted from the traditional joystick control, the two systems are comparable.; (3) By combining multi-modality sensor information from both forehead EMG signals and facial images, facial motion noises such as eye blinks and random facial expression movements including laughing, smiling and talking are effectively avoided during control.

Our future work will improve the current control method in two ways. One way is to keep the current sensor modalities and investigate more robust and extictive features and movement patterns on existing EMG signal and image information. The other way is to include more sensory modalities and develop more complex control system with manifold sensor combinations.

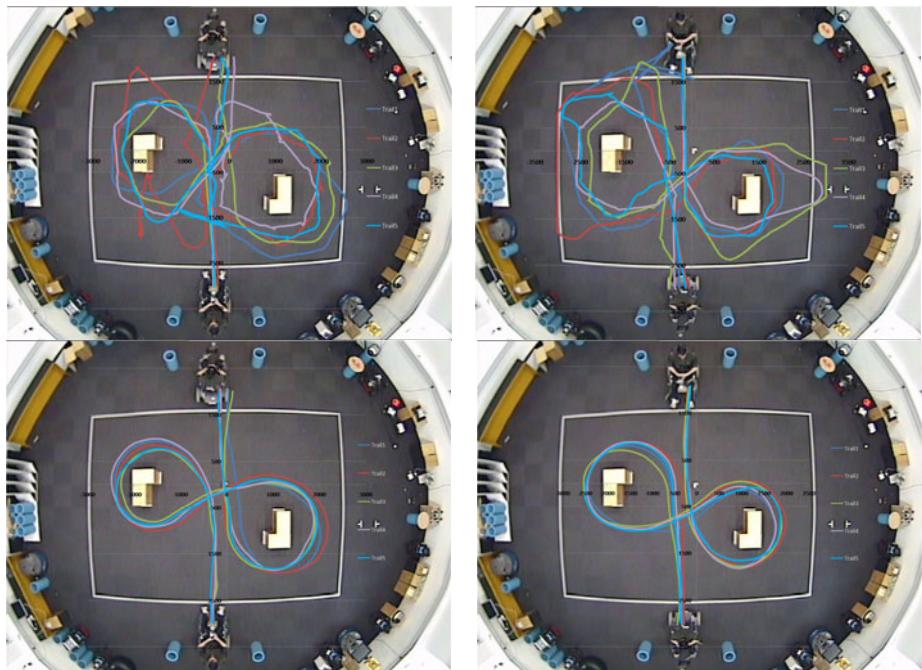
## References

- [1] Felzer, T., Freisleben, B.: HaWCoS: The ‘Hands-free’ Wheelchair Control System. In: Proceedings of International ACM SIGACCESS Conference on Computers and Accessibility, pp. 127–134. ACM Press, New York (2002)
- [2] Moon, I., Lee, M., Chu, J., Mun, M.: Wearable EMG-based HCI for electric-powered wheelchair users with motor disabilities. In: Proceedings of IEEE International Conference on Robotics and Automation, pp. 2649–2654 (2005)
- [3] Barea, R., Boquete, L., Mazo, M., Lopez, E.: System for assisted mobility using eye movements based on electrooculography. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 10, 209–218 (2002)
- [4] Rebsamen, B., Burdet, E., Guan, C., Zhang, H., Teo, C.L., Zeng, Q., Laugier Jr., C., Marcelo, A.H.: Controlling a Wheelchair Indoors Using Thought. *IEEE Intelligent Systems* 22(2), 18–24 (2007)
- [5] Tsui, C.S.L., Jia, P., Gan, J.Q., Hu, H., Yuan, K.: EMG-based hands-free wheelchair control with EOG attention shift detection. In: Proceedings of IEEE International Conference on Robotics and Biomimetics, pp. 1266–1271 (December 2007)

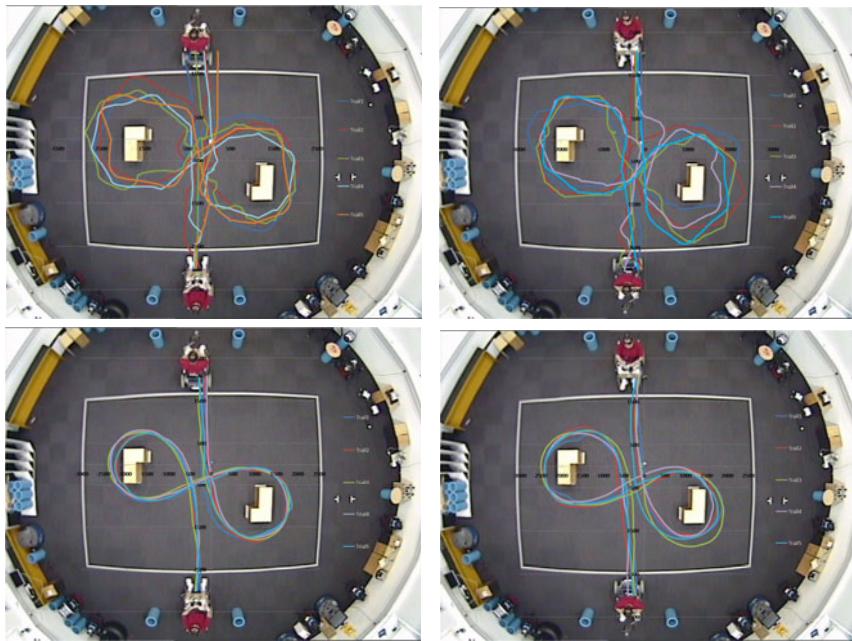
- [6] Ju, J.S., Shin, Y., Kim, E.Y.: Vision based interface system for hands free control of an Intelligent Wheelchair. *Journal of Neuroengineering and Rehabilitation* 6 (2009)
- [7] Li, X., Tan, T., Zhao, X.: Multi-modal Navigation for Interactive Wheelchair. In: Tan, T., Shi, Y., Gao, W. (eds.) ICMI 2000. LNCS, vol. 1948, pp. 590–598. Springer, Heidelberg (2000)
- [8] Matsumoto, Y., Ino, T., Ogasawara, T.: Development of Intelligent Wheelchair System with Face and Gaze Based Interface. In: Proceedings of 10th IEEE Int. Workshop on Robot and Human Communication (ROMAN 2001), pp. 262–267 (2001)
- [9] Ferreira, A., Silva, R.L., Celeste, W.C., Bastos Filho, T.F., Sarcinelli Filho, M.: Human-machine interface based on muscular and brain signals applied to a robotic wheelchair. *Journal of Physics: Conference Series* 90, 012094 (2007)
- [10] Cyberlink Brainfingers, Hands-free Computer Access Solution,  
<http://www.brainfingers.com/>
- [11] Jia, P., Hu, H., Lu, T., Yuan, K.: Head gesture recognition for hands-free control of an intelligent wheelchair. *Journal of Industrial Robot* 34(1), 60–68 (2007)

## Appendix

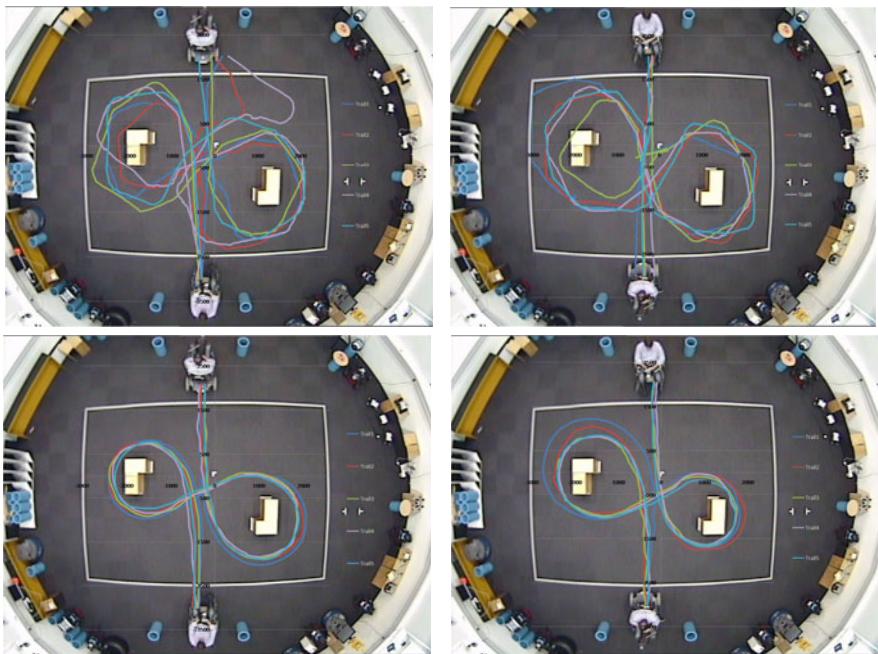
The following figures show the trajectories of real-world experiment results for five subjects, i.e. alphabetically from Subject A to E.



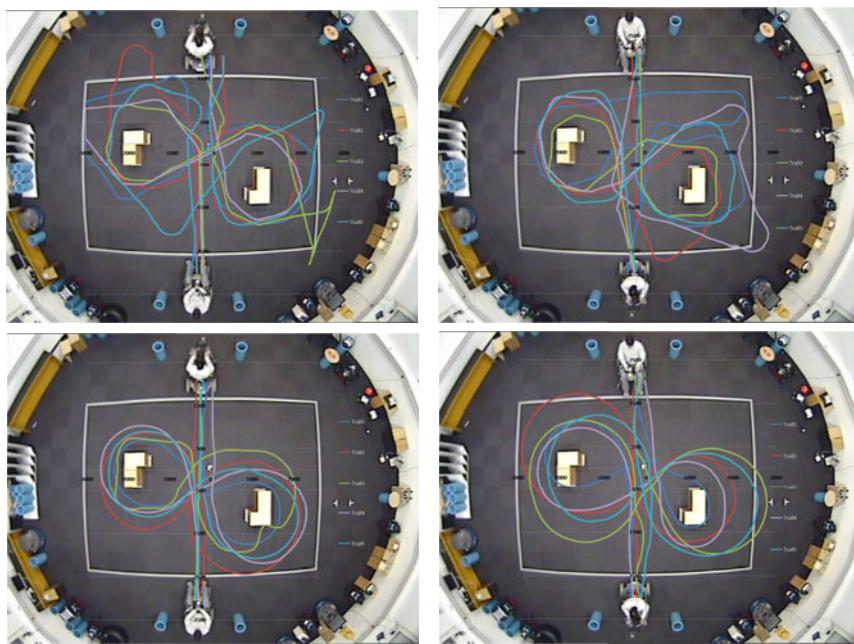
**Fig. 6.** Trajectories recorded on Subject A during completing four designed tasks: upper left and upper right (EMG and Visual based control); lower left and lower right (joystick control)



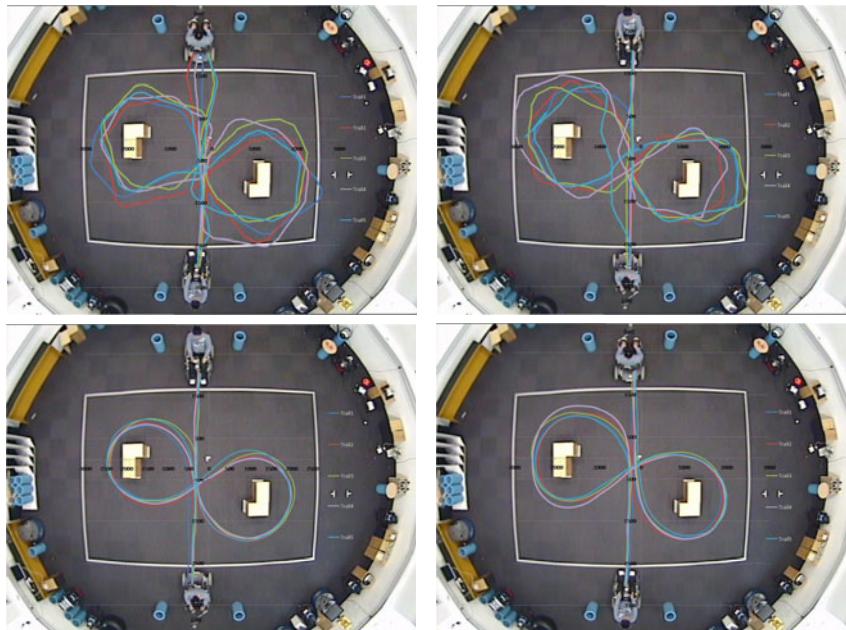
**Fig. 7.** Trajectories recorded on Subject B during completing four designed tasks: upper left and upper right (EMG and Visual based control); lower left and lower right (joystick control)



**Fig. 8.** Trajectories recorded on Subject C during completing four designed tasks: upper left and upper right (EMG and Visual based control); lower left and lower right (joystick control)



**Fig. 9.** Trajectories recorded on Subject D during completing four designed tasks: upper left and upper right (EMG and Visual based control); lower left and lower right (joystick control)



**Fig. 10.** Trajectories recorded on Subject E during completing four designed tasks: upper left and upper right (EMG and Visual based control); lower left and lower right (joystick control)

# A Novel Hand-Gesture Recognition Method Based on Finger State Projection for Control of Robotic Hands

Wenzhen Yuan and Wenzeng Zhang

Key Laboratory for Advanced Materials Processing Technology, Ministry of Education;  
Dept. of Mechanical Engineering, Tsinghua University, Beijing 100084, China  
yuan\_wz08@126.com, wenzeng@tsinghua.edu.cn

**Abstract.** This paper proposes a novel method for hand-gesture recognition based on finger state projection, called FSP Method, which is used to control robotic hands. The control system using FSP Method can simplify the control process of robotic hand considerably while avoiding the drawbacks of traditional hand-gesture recognition methods. The FSP Method measures the projection length of fingers through a monocular vision to infer the state of fingers, and therefore the angle of each joints in fingers. The information is used to control the motors of a robotic hand and make it to pose as is wished. The experimental results show that the FSP Method is effective. The FSP Method does not need lots of study work which is almost required by traditional methods, and the FSP Method has higher adaptability and bigger recognition range than traditional methods.

**Keyword:** Control of robotic hands, hand gesture recognition, human-computer interaction, machine vision.

## 1 Introduction

The robotic hand is an important part of the design of humanoid robot, and to a certain extent it is considered as the most complex part of the designing. With the rapid development and popularization of humanoid robot, the humanoid robotic hand is bound to play a more important part in industrial fields and daily life. Compared to the other part of a robot, the robotic hands gather much more joints in a comparatively confined space. A fully-formed humanoid robotic hand contains no less than 15 joints in a single hand. The numerous joints enable the high dexterity of robotic hand and make the manipulating more complex at the same time.[1]

Currently there are three dominant ways of controlling robotic hands: controlling by fixed program, controlling by operation panel, controlling by data gloves. All the three methods have their insufficiencies and thus hampered the popularization of humanoid robotic hands. The method of manipulating by fixed program requires lots of work and the manipulator's possessing adequate technique on programming and is hard to change the moving routine.[2] The operation panel manipulation uses an operation panel which contains lots of buttons or rotary knob corresponding to different joints.

This demands the manipulator being familiar with all the buttons, which requires lots of effort, and the rapid and dexterous manipulation is hard to achieve through this way. The data glove is a convenient way and can realize the dexterous real-time control and have no technical require to the operator. But the cost of a data glove is not desirable and different people with different size of hands may have difficulties use it, and it is not comfortable to wear the glove for long time. So up to now generally it is only used in laboratories.[3]

As the situation stated above, an avenue of controlling robotic hand with a good real-timing character and low cost is in need. It should be easy to manipulate as well. So the avenue of manipulating through vision based hand gesture recognition drew our attention. The Robotics Institute in Beihang University has made similar attempt by recognizing the symbols on the finger tips and analyzing their movement to decide the movement of the corresponding robotic finger.[4] However the method is not convenient enough and has too much restriction to the manipulating environment.

This paper brings up a novel method of robotic hand manipulation through monocular vision based recognition of hand gestures. This method is much more efficient, flexible, convenient and applied to robotic hand manipulation than the traditional hand gesture recognition methods. Completed experiments proved the method a much convenient and effective way to control robotic hands.

## 2 The Novel Method Based on Finger State Projection

### 2.1 The Drawbacks of Traditional Methods of Hand Gesture Recognition

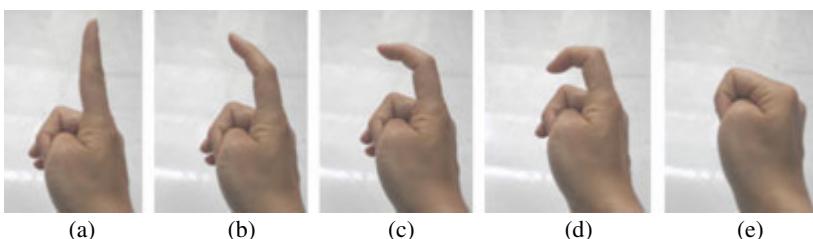
Though pattern recognition has developed rapidly over recent years, the current gesture recognition technique is not very satisfactory. The reason of the situation is the flexibility of human hands. Compared to the face identification, the hand gestures have much more varieties and the features are much harder to be recognized and fetched by machine intelligent.

Current gesture recognition methods build up a library of gestures, and recognize the specific gesture by matching the current gesture with the samples in the library. The matching process is mostly either inefficient or imprecise. And the method requires complicated study procedure for each operator and the gestures it can recognize is restricted to the gestures in the library through study. Those features restricted the wide use of the technique and its application in robotic hand manipulation.[5–7]

### 2.2 Model Simplification Based on the Finger Movement

According to the situation stated above, it could not be neglected that building up a method of gesture recognition which can possess efficiency, convenience, accuracy, generality is nearly impossible under current condition. Nevertheless, as the objective of the control system is to robotic hands, it is not necessary to recognize all the poses and details of human hands, only the information related to the control should be fetched. In this case the model can be simplified greatly.

The operation objective in this paper is for humanoid robotic hands, and under-actuated robotic hands in particular. A kind of under-actuated robotic hands use coupling technique to reduce the number of motors used in the robotic hands while keep the function of the robotic hands to grasping and gesturing in poses like human hand. A theoretical basis for this is that the rotating of human hand joints is mostly coupling instead of moving independently. For example, for most people it is hard and unnatural to move the distal finger joint without moving other joints, but when rotating the middle finger joint the distal joint would rotate together in a adequate angle. So the control system may use one motor and some coupling mechanism to drive the two joints together in the robotic hand. In this way the under-actuated robotic hands can make a simple mechanism as well as low cost. As the number of motors in the hand is reduced, the limit of its size is not very strict and the robotic hand is able to have a greater strength. So the under-actuated robotic hand is a trend for the popularization of robotic hands.[8–10]



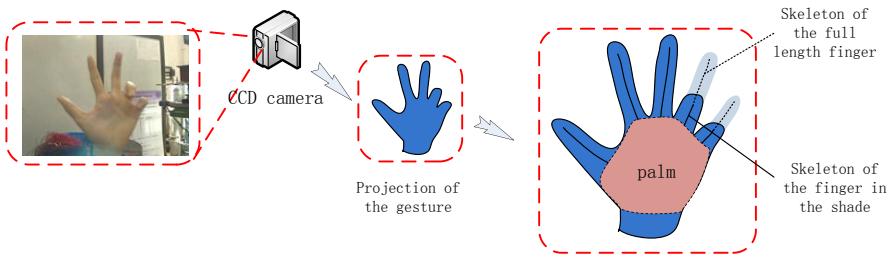
**Fig. 1.** The natural state of index finger bending at different degree. The pictures clearly showed that the distal joint and the middle joint are coupling when moving.

Having analyzing the manipulating objective, model of the manipulation may be simplified. The information needed to be received from the operator is the states of each joints on the hand. And as the theory of coupling movement of human hand, if the bending degree of the whole finger is known, the bending angles of each joints on the finger can be estimated. So if the bending state of each finger of the operator's gesture is get, it is possible to control the robotic hands.

The same theory and method are applied to dexterous hands.

### 2.3 Model Built Up

As is analyzed above, what is needed to be obtained from the operator's gesture is the bending state of each finger. Viewing from the front side, the bending degree of a single finger can be represented by the length of its shape. The shorter the shape is, the greater the finger is bent. So in the FSP Method, the operator is required to put his hand facing directly and ensure his fingers' shades are apart in the camera view. Then the shade of the finger is skeletonized and the length of the skeleton is measured. By comparing with the standard length the bending degree of the finger can be got.



**Fig. 2.** The sketch map of core FSP gesture recognition method. By comparing the length of finger skeleton in the frame and the length of full finger skeleton the finger's state is known.

The FSP Method dose not require a long procedure of study and can be easily used by different people. And when manipulating this system, no special gloves or background is required. The study procedure of this method is within only one frame to adapt to the hand of the manipulator.

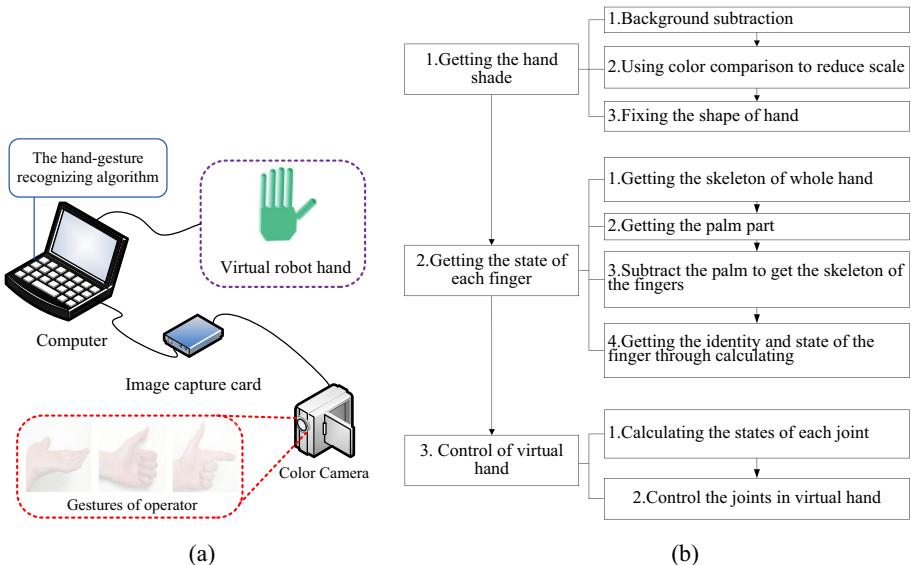
The completed experiments on this method proved this method an efficient and relatively reliable way to manipulate robotic hands. The first version of the program can achieve roughly 15 frames per second. The following table made a comparison between the traditional gesture recognizing algorithm and the one used in this method.

**Table 1.** The comparison between traditional gesture recognizing methods and FSP Method

items	Traditional method (simple)	Traditional method (refined)	FSP Method
<b>Program efficiency</b>	high	low	high
<b>Reliably (to control robotic hands)</b>	low	high	relatively high
<b>Generality</b>	bad	bad	good
<b>Kinds of gestures available</b>	limited (depend on the library built through study)	limited (depend on the library built through study)	theoretically unlimited
<b>Procedure of study</b>	long and complicated	long and complicated	nearly no study procedure
<b>Require for the hand</b>	some require colored gloves	some require colored gloves	no gloves are required
<b>Require for the manipulating environment</b>	require a background of single and even color	require a background of single and even color	applicable to relatively complex background
<b>Require for the camera</b>	single camera, no specific requirements	some require multi cameras and the specific adjustment of cameras	single camera, no specific requirements

### 3 The Realizing of FSP Method

The hardware system and procedure of the FSP Method is shown in the figures. As the restriction of the hardware, a virtual 3D robotic hand is made the manipulating objective instead of a real robotic hand.



**Fig. 3.** (a) The hardware system of FSP Method, (b) The flow path of the FSP Method

#### 3.1 Getting the Hand Shape from Complicated Multicolor Background

##### 3.1.1 Getting the Target Working Area through Background Subtraction Method

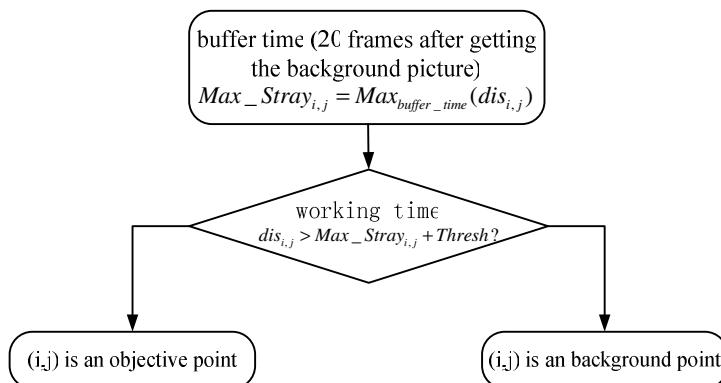
Separating the hand with the multicolor background is a hard job, especially when the background is unknown. As the color fluctuations of the hand itself is considerably as the different part of skin have different colors and the uneven illumination caused by the uneven surface of the hand when posing different gestures, identifying the hand area from the background is very hard and steady.

FSP Method uses background subtraction to separate the hand with the background as the first step of hand shape fetching. The completed experiments has proved this method a very efficient and relatively reliable way to separate the background.

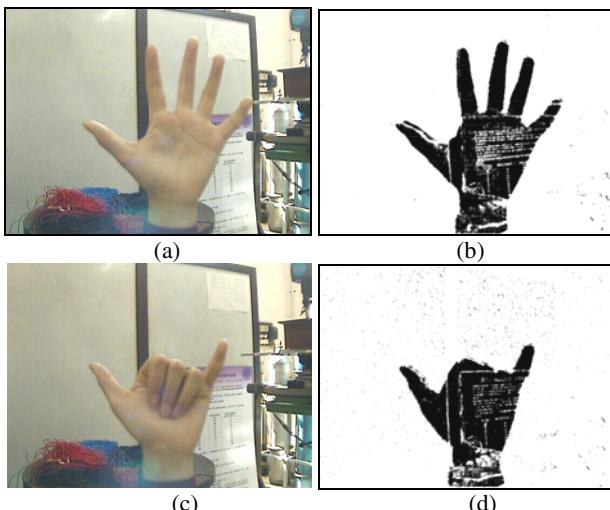
The subtraction uses multicolored separation and RGB as the color space. To measure the degree of one point deviation from the background, the method used the Euler distance. And in order to improve the efficiency, the rooting procedure is submitted:

$$dis_{i,j} = (R_{i,j} - R_{0,i,j})^2 + (G_{i,j} - G_{0,i,j})^2 + (B_{i,j} - B_{0,i,j})^2$$

Experiments showed that the subtraction result is not steady, as there is a great chance for the system to misjudge a background point as non-background point. This is resulted by the instability of the camera. When compare two pictures took by the same camera with the same parameters of the identical scene, manipulator can find some spots on it are totally differently colored in the two pictures. That may result from the camera's hypersensitivity to certain color or illumination. It may cause great error to the separate result obtained. To minish the error, FSP Method establish a buffer time (20 frames, no more than 2 seconds). During the buffer time, the system require that the background remain static and record the fluctuations of the value. Then during the work time, the system compare the stray value of each point in a frame with the max stray value recorded to decide whether the point is in the moving area.



**Fig. 4.** Flow chart showing how the improved background subtraction work in FSP Method



**Fig. 5.** The result of the improved background subtraction in FSP Method to get the approximate area of the hand

Experiments proved that this improvement minimizes the error greatly.

### 3.1.2 Getting the Hand Area by Color Threshold

In order to remove points of other moving objects but not the hand, the method uses color comparing on the shade got from the former step. The *distance* of the point in the non-background area to the standard color of hand is calculated, and it is used to measure how likely the point is located in the hand area. As the most desirable character to distinct the color of hand is tone but not the luminosity, here the method chose the Lab space as the color space and reduce the influence of luminosity by multiplying a factor. So the following expression is obtained:

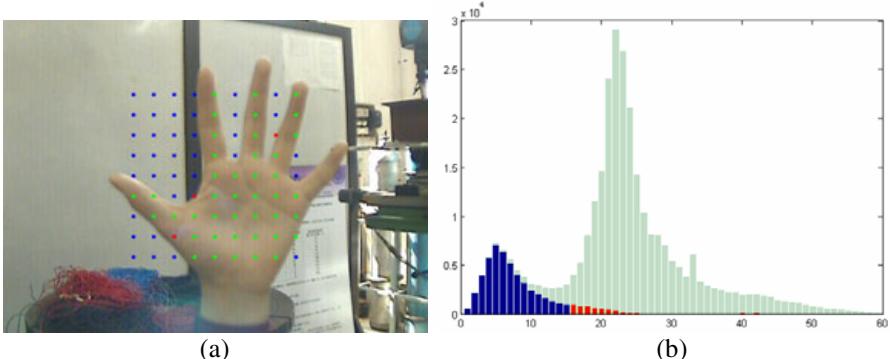
$$\text{Color\_dis}_{i,j} = (L_{i,j} - L_{0,i,j})^2 \times 0.1 + (a_{i,j} - a_{0,i,j})^2 + (b_{i,j} - b_{0,i,j})^2$$

And if the Color\_dis is smaller than the threshold, FSP Method decides the point is in the hand area.

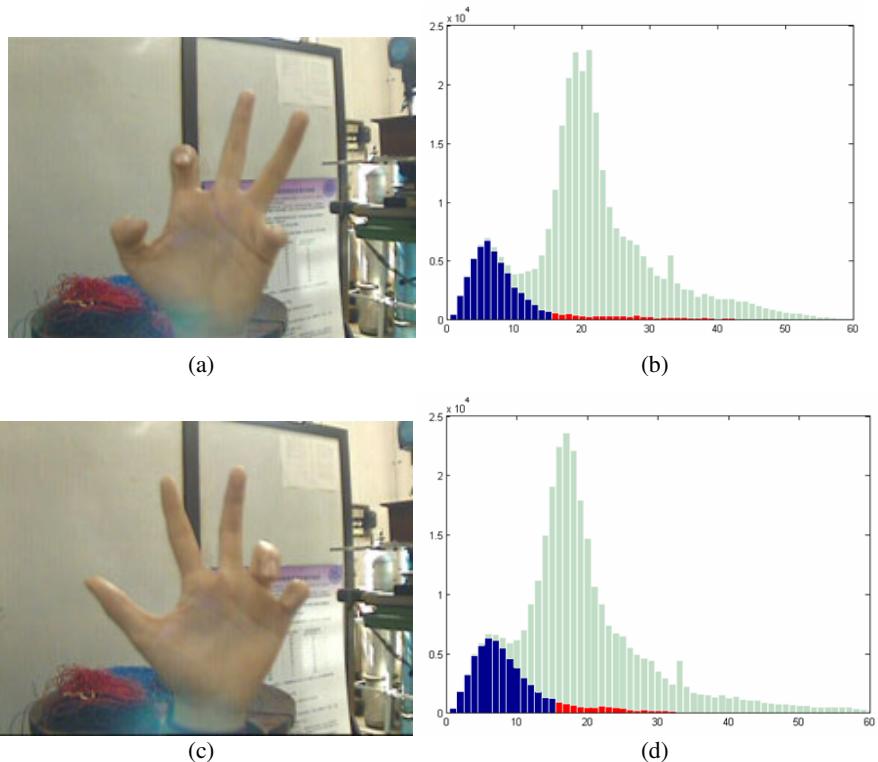
### 3.1.3 Getting the Standard Color

The getting of standard color is another problem. The traditional way is to scan every pixel in the shade and build up the range of the skin. But lots of work is required in this method and much of it is unnecessary. In FSP Method, a new algorithm is set up to get the sample point in the initialization in the area that is likely to be in the skin space, then outliers are eliminated. The mean color of the left points can be considered as the standard color. This algorithm improved the efficiency to a large extent, and the completed experiments have proved the method of high reliability.

Fig. 6. and Fig. 7. are the examples of the color fetching method. Fig. 6. (a) is an initial frame, and sample points are got from it. The blue points are points in the sample area but without the background subtracted area, the red points are within the background subtracted area but they were the outliers of the majority. The green points are points left and taken as samples. Their mean color of them is taken as the standard color. Fig. 6. (b) is the color test on the *distance* of the pixels from the standard color. The X axis is the color *distance*, the Y axis is the count of the pixels of the *distance*. The blue part on the graph is the points in the fixed hand shade got from the next step and



**Fig. 6.** Getting the color sample from an initial frame and the test

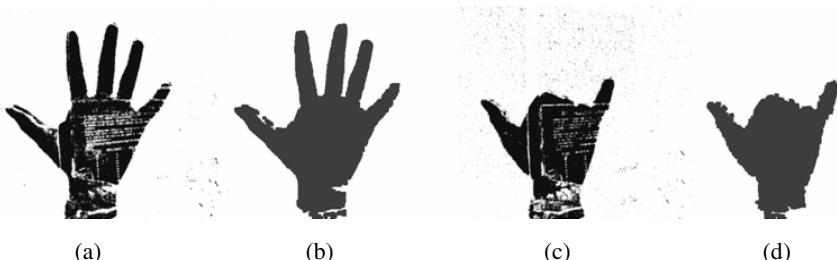


**Fig. 7.** Tests on ordinary gestures

within the color thresh, the red ones are the points in the final shade but without the color thresh, the light green ones are points in the background. Fig. 7. is the same color test of two other frames of ordinary gestures.

### 3.1.4 Fixing the Hand Shape

After the former two steps, the method approximately get the area of hand. But unavoidably some noise points are chosen as well. Substantially it is sure the overwhelming is the hand



**Fig. 8.** Comparison between the shades got from background subtraction and the fixed shade. It can be clearly seen that noise points are removed and the whole shape is satisfactory now.

area. So the first step of fixing is to find the largest connected domain. Then much of the noise or interferences are dislodged.

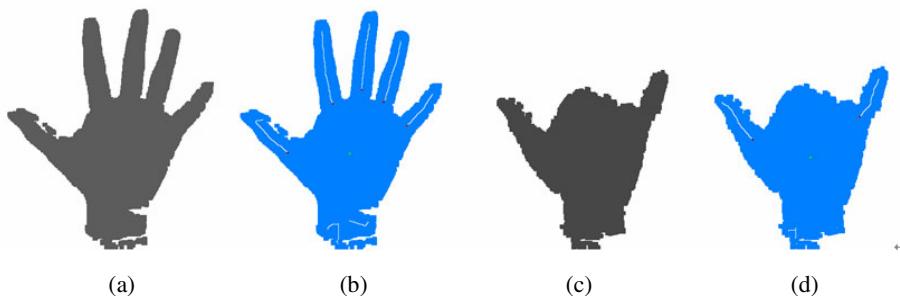
Then the method refines the shape to dislodge the noise further. FSP uses the opening operation first to get rid of the narrow trenches, then filled the holes in the shade. After these a fine shade of the hand is obtained.

### 3.2 Getting the Skeleton of the Fingers

To get the skeleton of the whole hand, FSP uses morphological erosion on the shade of hand to the least interconnected part. To make convenience for the subsequent steps, the method marks the points that are eroded with the round it was been eroded, which represent its distance to the background.

Then FSP gets the palm. It uses dilation from the most center point, which is marked with the max distance from the background and could be regarded as the center part of the palm. The method dilates from this point, round by round, and stop when it first reach the background point on whichever direction.

FSP then dislodge the part in the palm area from the hand skeleton, and approximately get the separated skeletons of each fingers.



**Fig. 9.** The hand shade and the finger skeleton marked with color white in the shade

### 3.3 Getting the Information of the Fingers

#### 3.3.1 To Identify the Finger

From the observation one could find that if he keep the palm fixed and move the fingers at random, the positions of finger roots would approximately remain constant. This character of hand could help us to identify the fingers.

Firstly, FSP locates the center point of the palm. By convenience the centroid of the palm area can be used (already in section 3.2). Experiments show that its position would approximately change with the palm and remains relatively stationary to the point. Then FSP gets the root of each finger. Then the method connects the finger root and the palm center point with a line, and calculated the angle between the line and the vertical direction. Then it matches the angle with the angles got in the standard graph obtained in the initialization to identify the finger.

### 3.3.2 Getting the State of the Fingers

The bending degree of the fingers can be reflected from the length of the shade of the fingers, and FSP measures the length of the finger skeleton as the length. To minish the effect of the distance of the hand and the camera which would result in the size of the hand shade, the method uses the relative length of the finger, which is the pixel length of the finger divided by the area of the palm. Then it calculates a ratio number  $s$  as following:

$$\text{length}_r[i] = \frac{\text{length}(\text{finger}[i])}{\text{Area}(\text{Palm})}, \quad s[i] = \frac{\text{length}_r[i]}{\text{length}_0[i]}.$$

Here the  $\text{length}_0[i]$  represents the finger- $i$ 's relative length in the standard graph. And after these the method use  $s$  to approximately determine the bending degree of the finger. As the method is not an accurate manipulating method, the value  $s$  is not an accurate value, but it is enough to determine the approximate state of the finger. In completed experiments, FSP set up 3 thresholds for  $s$  to identify 4 states of the finger.

### 3.4 The Manipulating Part

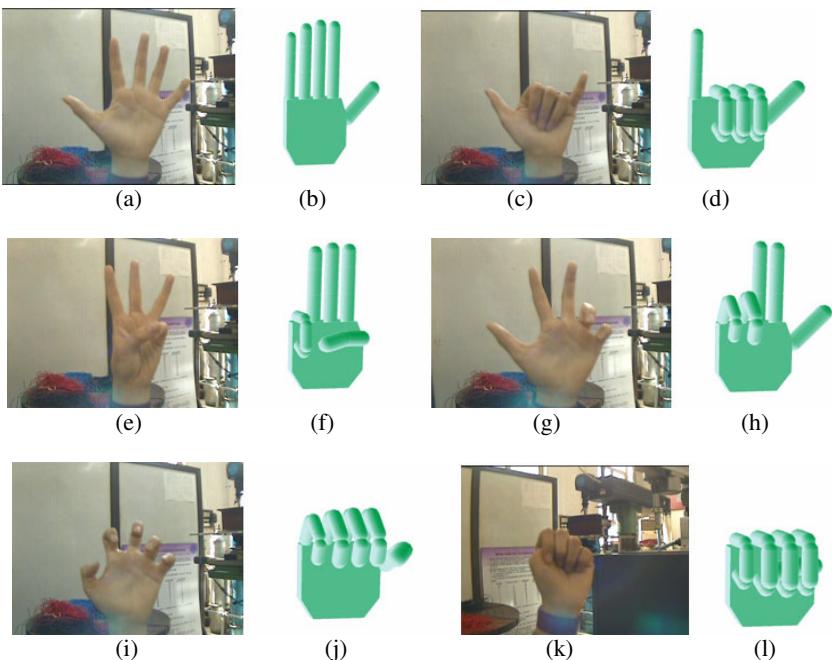
After the bending degree of each fingers is got, the program can estimate the angles of the motors on the fingers, and make the motors rotate to the specific position. Then the robotic hand would pose the expected pose. If manipulators expect the system to work satisfactory function, more mechanical improvement is required to improve its performance. For example, if the robotic hand is expected to do excellent grasping job, pressure transducers and automatic adaptive devices can be added to the mechanical section of the hand. And when manipulating different kinds of robotic hands, the program can be improved according to the different character of the hand.

As the hardware is limited, the author has not got the chance of installing the manipulating system on a real robotic hand up to now. To test the effect of the manipulation system, an 3D model of a robotic hand was built as the objective of the manipulation. As the virtual hand has 15 joints as most of the basic robotic hand, it can represent the real work state of a robotic hand well.

## 4 Manipulation Experiments

Completed experiments showed the FSP Method can successfully fetch gestures made by naked hand from a comparatively complicated background and order the virtual 3D hand to follow the gesture. The real-timing of the method is magnificent that it can reach roughly 15 frames per second when running on an ordinary lap-top.

In the experiments finger states were roughly divided into four statuses and each corresponds specific angels for the motors on the robotic fingers. The accuracy of the recognition can reach over 80 percent by estimate when manipulating by the regulation. The following pictures showed the system recognizing different gestures from relatively complicated backgrounds and manipulate the robotic hand to follow it.



**Fig. 10.** Some screen shots of experiments of the FSP Method

## 5 Conclusion

This paper brought forward a new method of manipulating robotic hand by monocular vision based recognition of hand gestures. Compared to traditional hand gesture recognition algorithm, the new algorithm of recognition used in this paper is more efficient, adaptive, unlimited, and it does not require long process of study or high operating skill for the operator. The experimental results show that the FSP Method is effective. The FSP Method does not need lots of study work which is almost required by traditional methods, and the FSP Method has higher adaptability and bigger recognition range than traditional methods.

**Acknowledgements.** This paper was supported by Hi-Tech R&D Program of China (No. 2007AA04Z258), National Natural Science Foundation of China (No. 50905093), Tsinghua Basic Research Foundation (No.JC2007009) and the foundation of Key Laboratory for Advanced Material Processing Technology, Ministry of Education, P.R. China (No.2008011).

## References

1. Okamura, A.M., Smaby, N., Cutkosky, M.R.: An overview of dexterous manipulation. In: IEEE inter. Conf. on Robotics & Automation, pp. 255–262. IEEE Computer Society Press, Los Alamitos (2000)

2. Zhang, A., Zhang, Q.: Programming method for robot teaching. *Machine Tool & Automatic Manufacturing Technique* (4), 47–49 (2003)
3. Hu, Y.: An overview of gesture recognition. *China Science and Technology Information* (2), 42 (2005)
4. Li, J., Su, W., Zhang, Y., et al.: Vision-Based Grasp Planning System for Dexterous Hands. In: *Inter. Conf. on Intelligent manipulation and Grasping*, pp. 555–558 (July 2004)
5. Sun, L., Zhang, L., Guo, C.: Technologies of Hand Gesture Recognition Based on Vision. *Computer Technology and Development* 18(10), 214–216 (2008)
6. Tang, Z., Feng, Z.: A Gesture Recognition System Used on Robot Vision. *Computer Engineering and Applications* 18(16), 51–54 (2005)
7. Li, R., Jia, J.: A Method of Extracting Gesture from Complex Background. *J. Huazhong Univ. of Sci. & Tech (Natural Science Edition)* 36(Suppl. I), 80–82 (2008)
8. Dollar, A.M., Howe, R.D.: Towards grasping in unstructured environments: grasper compliance and configuration optimization. *Advanced Robotics* 19(5), 523–543 (2005)
9. Zhang, W., Che, D., Liu, H., et al.: Super under-actuated multi-fingered mechanical hand with modular self-adaptive gear-rack mechanism. *Industrial Robot: An International Journal* 36(3), 255–262 (2009)
10. Liu, H., Gao, X., Shi, S.: Under-actuated self-adaptive artificial hand: Chinese Patent: CN 1292719C (2007)

# Adaptive Compliance Control for Collision-Tolerant Robot Arm with Viscoelastic Trunk

Chungwei Chin<sup>1</sup>, Yanan Li<sup>2</sup>, Shuzhi Sam Ge<sup>3</sup>, and John-John Cabibihan<sup>4</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, National University of Singapore,  
Singapore 117576, Singapore  
jowy\_1029@hotmail.com

<sup>2</sup> Social Robotics Lab, Interactive Digital Media Institute and NUS  
Graduate School for Integrative Sciences and Engineering, National University of Singapore,  
Singapore 119613, Singapore  
hit.li.yn@gmail.com

<sup>3</sup> Institute of Intelligent Systems and Information Technology (ISIT), University of Electronic  
Science and Technology of China, Chengdu 610054, China, and Social  
Robotics Lab, Interactive Digital Media Institute and the Department of Electrical and  
Computer Engineering, National University of Singapore, Singapore 117576, Singapore  
samge@nus.edu.sg

<sup>4</sup> Social Robotics Lab, Interactive Digital Media Institute and Department of Electrical and  
Computer Engineering, National University of Singapore, Singapore 117576, Singapore  
elecjj@nus.edu.sg

**Abstract.** In the present work, a collision-tolerant control method was proposed to achieve a stable task performance of the robot arm in an acceptable collision-safe margin. A robot arm covered with the soft capacitive sensing skin mounted on a passive mechanical viscoelastic trunk has been proposed to suppress the unexpected collisions which may happen from anywhere in the social context. The viscoelastic trunk contributed to the passive deformation in response to the collision forces and the joint configurations of the robot arm were directly calculated in accordance to the trunk deformation. An adaptive control scheme has been introduced to achieve a good end-effector trajectory performance upon collisions. Simulation results and actual assessment on a prototype model confirmed the effectiveness of proposed methods.

## 1 Introduction

To date, robots have been employed in many applications, in particular, nonindustrial areas such as services in homes and offices and for social welfare [1]. The chances of interactions between human and robot will become more frequent in an unstructured social workplace where their working zones are overlapped. It is believed that injuries could potentially occur during unexpected collisions between robot and human being. United Auto Workers [2] published a report which provides raw data on various injuries related to robot operations including bone fracture which could result from manipulator pinch points or direct crush loads. When a manipulator is operating in free-space (traditional industry environment), it is sufficient to specify and control the

position of the tool. In a social workplace which differs from industrial context, the number of tasks that demand contact between the human and robot is also increasing. There is a higher potential for unexpected collision impact to happen since it is not suitable to pre-program the social robots for dealing with potential contact that may happen around the manipulator like in industrial case. The danger criterion should be constructed from measures that contribute to reducing the impact force in case of unexpected human-robot impact. This requires the engineers to design the human-friendly robots have to keep the possibility of human-robot coexistence in mind.

In past work to deal with the collision, the compliance control has been extensively studied in the literature and can be achieved either by active means [3] or passive means. In the active compliance control, a force sensor feedback is used to modify the motion of the robot. The controller will make use of the feedback signals from force sensors to detect the applied external force and give a proper response. The active compliance approach generally has limitations in delayed control and incomplete safety in the case of electrical malfunctioning even though it can offer a high programming ability. In this regard, the active compliance approach is not enough to avoid either the damage to human or the failure to power transmission when unexpected collision occurs. In the passive compliance control, a passive mechanism is adopted to produce an appropriate reaction to the contact force [4-6]. A robot arm with passive compliant joints to deal with collision has been presented in [7]. However, the passive compliant joints possess heavy weight which is not suitable to be used in the social context. In [8], a collision-tolerant control law has been proposed which is able to recover the deviation of the end-effector caused by the passive trunk motion so that it can perform a certain task in a limited range of contact forces. However, it is found that it is not reasonable to be employed in practical situation due to its complicated computation. Also, it does not suit for real cases by considering a two-dimensional model and only a normal contact force. All these problems have to be tackled and integrated by an efficient solution. In order to construct a robot arm which can finish the pre-defined tasks while accommodating to the unexpected collisions, we propose a reliable and efficient control mechanism in the following sections.

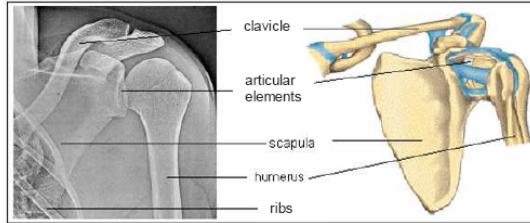
## 2 Modeling

In this section, the analogy of the human shoulder joint is investigated which allows passive compliance modifications on the mechanical design of the robot arms to take place. In Fig. 1, it is shown that three bones and the articulations between the bones make up the flexible and force-tolerable shoulder joints to perform a wide range of motions required in the arms and hands, such as pushing, pulling and lifting.

Characteristics of the shoulder joints can best be summarized in a rather simple Kelvin-Voigt viscoelastic model [9] with the combined attributes of viscosity and elasticity. The  $n$  degrees-of-freedom (DOFs) Kelvin-Voigt model is described as

$$(\dot{x}_s - \dot{x}_r) = \eta^{-1}[F_k - E(x_s - x_r)] \quad (1)$$

where  $F_k \in R^n$  is the force developed by the spring and damper,  $E \in R^{n \times n}$  and  $\eta \in R^{n \times n}$  are the stiffness and the damping parameters, respectively, and  $x_s \in R^n$  and  $x_r \in R^n$  denote the deformed and rest spring displacement vector, respectively.



**Fig. 1.** Anatomy of shoulder

Consider a robot arm mounted on a Kelvin-Voigt viscoelastic trunk which exhibits the free moving ability while adding more DOFs to the robot arm. With more DOFs involved, the flexibility and possibility are introduced to tolerate unexpected collision forces from various directions. The Jacobian-based inverse kinematics for is described as

$$\theta_i = \emptyset^{-1}(x_i) \quad (2)$$

where  $\emptyset$  denotes the forward kinematics function,  $x_i \in R^n$  denotes the end-effector's position vector, and  $\theta_i \in R^n$  denotes the angle for the  $i$ -th joint with  $i \leq n$ . The dynamics equation for the robot arm is described by

$$D(\theta)\ddot{\theta} + C(\theta, \dot{\theta})\dot{\theta} + G(\theta) = \tau + J_c^T F_c \quad (3)$$

where  $\theta$  denotes the generalized coordinate,  $F_c \in R^n$  denotes the external contact force on the robot arm,  $\tau \in R^n$  represents the control input,  $J_c \in R^{n \times n}$  denotes the Jacobian transformation from a random contact point to the joint space of the robot arm,  $D(\theta) \in R^{n \times n}$  is the symmetric bounded positive definite inertia matrix,  $C(\theta, \dot{\theta})\dot{\theta} \in R^n$  denotes the Coriolis and Centrifugal force vector, and  $G(\theta) \in R^n$  is the gravitational force vector.

### 3 Control Design

The following control design is aimed to achieve two objectives, which include accommodating to the collision force and at the same time, maintaining stable end-effector trajectory.

#### 3.1 Collision-Tolerant Control Scheme

A random collision  $F_c \in R^n$  that happens to collide with the robot arm can be separated into two components: a trunk dissipative force vector  $F_{td} \in R^n$  causing viscoelastic trunk deformations and torque-generating tangent force vector  $F_{tg} \in R^n$  exerting external torques effect on individual actuators after the trunk suppression. The former is being suppressed and canceled out by the viscoelastic trunk deformation  $(x_s - x_r) \in R^n$  while the latter is to be addressed by the individual

actuator  $\tau_i$ , with  $i = 1, 2, \dots, n$ .  $T_c^T \in R^{n \times n}$  denotes the transformation matrix of the trunk with respect to the collision point. The trunk dissipative force  $F_{td} \in R^n$  is then given as

$$F_{td} = T_c^T F_c \quad (4)$$

The deformed passive viscoelastic trunk will exhibit  $-F_{td}$  to cancel out the dissipative tangent force effect. The torque-generating force vector  $F_{tg}$  is described as

$$F_{tg} = F'_c = (1 - T_c^T) F_c \quad (5)$$

The control law is then described as

$$\tau = K_p(\theta_d - \theta) - K_d(\dot{\theta}_d - \dot{\theta}) + G(\theta) - J_c^T F_{tg} \quad (6)$$

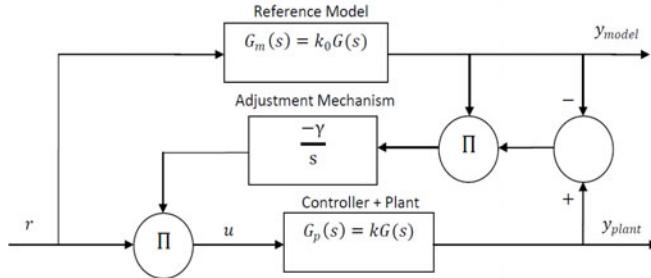
where  $\theta_d$  is the reference trajectory while  $K_p \in R^{n \times n}$  and  $K_d \in R^{n \times n}$  are proportional and derivative (PD) control gain matrices, respectively. The closed-loop equation becomes

$$D(\theta)\ddot{\theta} + C(\theta, \dot{\theta})\dot{\theta} + K_d(\dot{\theta}_d - \dot{\theta}) - K_p(\theta_d - \theta) = 0 \quad (7)$$

which has been proven to be stable in the control literature.

### 3.2 Adaptive Control Scheme

Some drawbacks of PD controller have been discovered. The proposed controller is required to be adaptive which involves modifying the control law to cope with the fact that the parameters of the system are time-varying.



**Fig. 2.** Model reference adaptive control block diagram

As shown in Fig. 2, a reference model is formed with a desired value of  $k_0$ , and through adaptation of a feedforward gain, the response of the plant is made to match this model. The update law is derived by calculating the sensitivity derivative and restated in terms of the model output:

$$\frac{\delta e}{\delta \theta} = kGU_c = \frac{k}{k_0}y_m \quad (8)$$

The constants  $k$  and  $k_0$  are combined into  $\gamma$ :

$$\frac{d\theta}{dt} = -\gamma' \frac{k}{k_0} y_m e = -\gamma y_m e \quad (9)$$

To tune this system, the values of  $k_0$  and  $\gamma$  are to be varied while generally  $\gamma$  is kept in a small value. The tuning of  $\gamma$  is crucial to the adaptation rate and stability of the controller.

The deformations of spring and damper are constantly fed into the adaptive controller to update the reference trajectory:

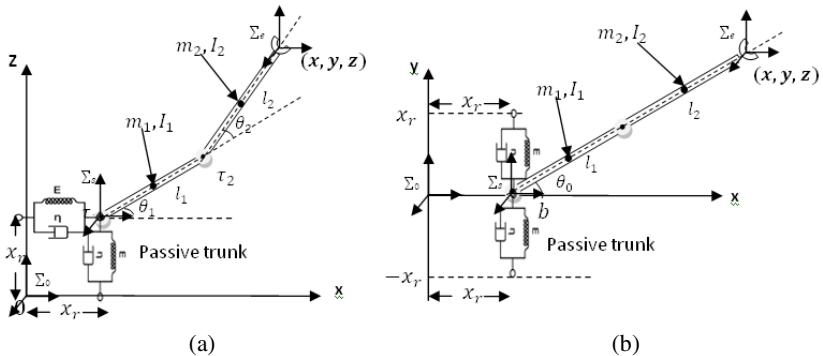
$$P'_d = P_d + (x_s - x_r) \quad (10)$$

where  $P_d \in R^n$  denotes the current reference trajectory while  $P'_d \in R^n$  denotes the updated reference followed by the trunk deformation for the time-being. After the algorithm has been derived it is now ready for the verification of the effectiveness through a case study on a specific robot arm structure.

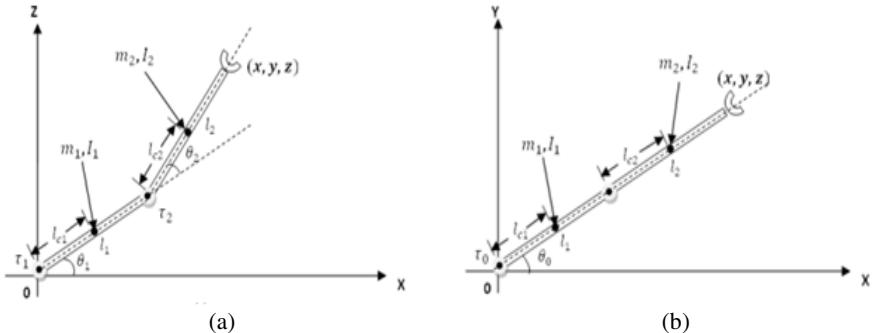
## 4 Simulation

A two-link robot arm with 3 revolute joints mounted on a three-dimensional viscoelastic trunk is considered to align with the paper's objective, which includes the safety issue and the task performance. Dynamics equation of the model is described in (6) and (7) with  $n = 3$ .

The arm collision simulation with the viscoelastic trunk, as shown in Fig. 3, is compared with that with a rigid trunk, as shown in Fig. 4. Adaptive controller with appropriate control gains has been employed in both systems. An object collides with the arm link located in each arm workspace. The common parameters of both robot arms in the simulations are given in Table 1.



**Fig. 3.** Three-dimensional robot arm with viscoelastic trunk (a) side view (b) top view



**Fig. 4.** Three-dimensional robot arm with rigid trunk (a) side view (b) top view

**Table 1.** Parameters of the Simulation Models

Link Number	1	2
Length [m]	0.3	0.25
Mass [kg]	1.3	1.1
Moment of Inertia [ $\text{kgm}^2$ ]	0.25	0.25
Model Inertia w.r.t. Joint 0 [ $\text{kg.m}^2$ ]		0.25
Viscoelastic Stiffness Matrix	diag(80,80,80)	
Viscoelastic Damping Matrix	diag(80,80,80)	
x, y, z coordinates desired [m]	(0.1, 0.42, -0.35)	

A collision situation has been considered in the current studies with the following procedures. The collision with magnitude  $F_c = 30N$ , angle w.r.t. z-axis  $\theta_{c1} = 0.8rad$  and angle w.r.t. x-axis  $\theta_{c2} = 0.7rad$  collides with link 2 at a distance  $h = 0.1m$  from joint 2 at  $t = 1s$ , and the contact force is exerted at the same point for a period of  $1s$  and finally discarded at  $t = 2s$ . For collision-tolerant model, the fine-tuned *PD* control gains are  $K_p = \text{diag}(23,8,3)$ ,  $K_d = \text{diag}(25,10,1)$  while for rigid model the control gains are  $K_p = \text{diag}(13,1,3)$ ,  $K_d = \text{diag}(11,6,1)$ . It was assumed that no friction exists between the surface of the arm link and contacting object.

By observing the simulation results, the end-effector positions have been found tracking the reference position closely in both types of models due to the fine-tuned control gains. The individual torque developed by each joint of collision-tolerant manipulator has been found below  $20Ncm$ . On the other hand, the robot arm with the rigid base is observed having large actuator torques during the collision. A large value of torques could reflect a significant burden to the robot arm. It can be concluded that for the robot arm with the rigid base, the contact force between the arm and the contacting object is greater than that for the robot arm with the viscoelastic trunk. As such, the three-dimensional redundant DOF exhibited by the passive trunk can accommodate to the collision forces with various orientations.

## 5 Practical Assessment

### 5.1 Hardware Implementation

Ideally the viscoelastic model characteristics are described as in (1) to relate the force and trunk displacement. However, it is difficult to employ a material which exhibits both the spring-damper analogy perfectly. For simplicity, the damping property is excluded while only the stiffness property  $E \in R^{n \times n}$  will be considered. The trunk model dynamics is then simplified as

$$(x_s - x_r) = E^{-1}F_t \quad (11)$$

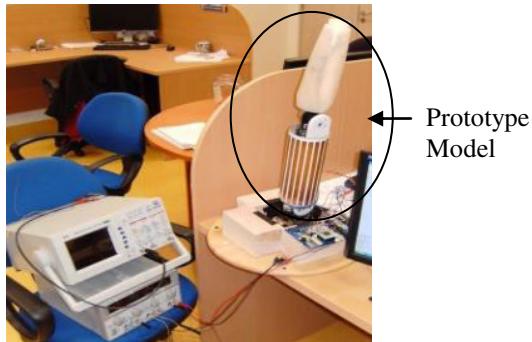
Five compression springs each with stiffness constant  $26N/cm$  that are sensitive enough to sense the exerted forces are connected together to form the viscoelastic trunk. A two-link 3 DOF robot arm is fabricated with Dynamixel RX28 being the joint actuator. Three RX28 actuators are linked to one bus for the controller to communicate with.

Simulation studies were performed under prior knowledge of collision which is not practical in real life where the information is usually unknown beforehand. To derive the contact location from a large area of contact surface effectively, CapSense CSD algorithm [10] is adopted to the robot arm. A slider sensor with 200 resolutions is modeled as a copper tape layout form whereas the sensing result is fed to the main controller in the double integer format.

### 5.2 Assessment Procedures

The complete prototype system has been set up, as shown in Fig. 5. Due to the simplification done on viscoelastic structure, in this current assessment, the purpose of the algorithm is to maintain the end-effector position within a certain margin instead of performing a full tracking. As mentioned above, by controlling the trajectory of three joints during a collision, the controller will try to make the end-effector coordinates move along the vertical  $z$  axis. In other words, the desired coordinates for  $x$  and  $y$  will always be the same while  $z$  coordinate will be a varying number which inversely proportional to the collision force magnitude.

Three collision tests have been performed in this section each with collision force with different contact location and magnitude. The collision sequences have been analyzed and described by sequence of pictures as shown in Fig. 6. Two cases are performed under unexpected situations where the controller does not have prior knowledge of where the collisions came from as well as how large they are. Test 1 considers forces with single direction in one move while case 2 encounters force with changing magnitude as well as direction.



**Fig. 5.** Experimental setup of prototype model

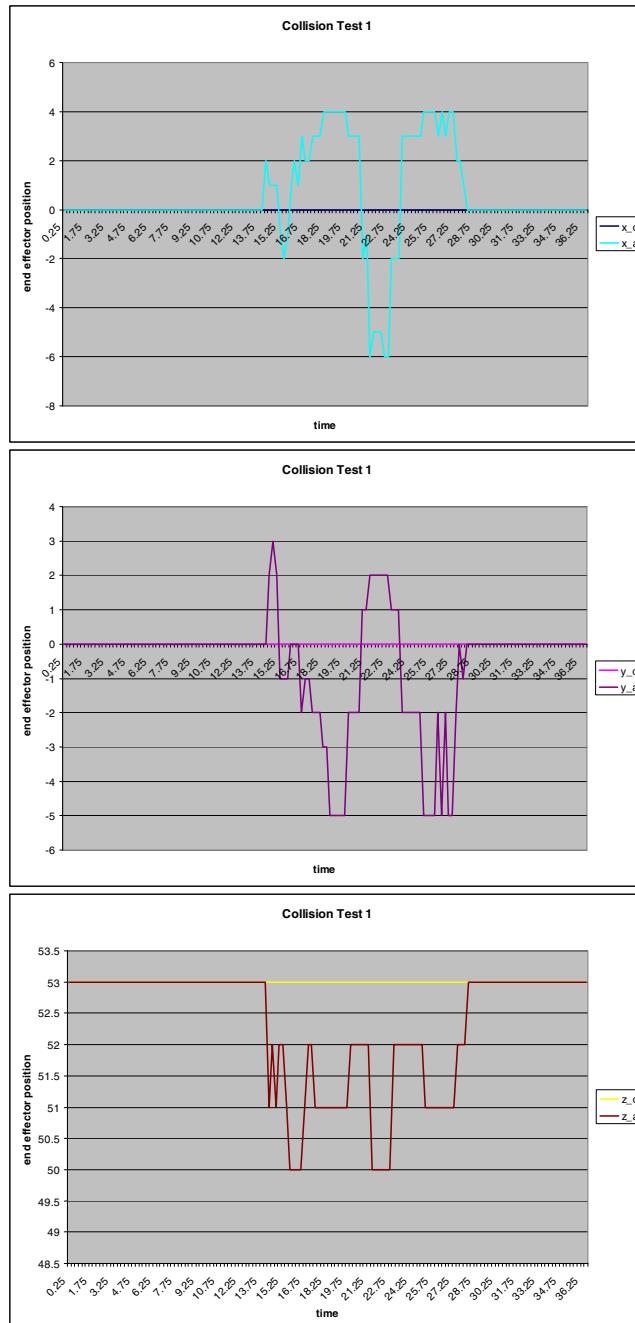


**Fig. 6.** Assessment on two collision situations (upper and lower)

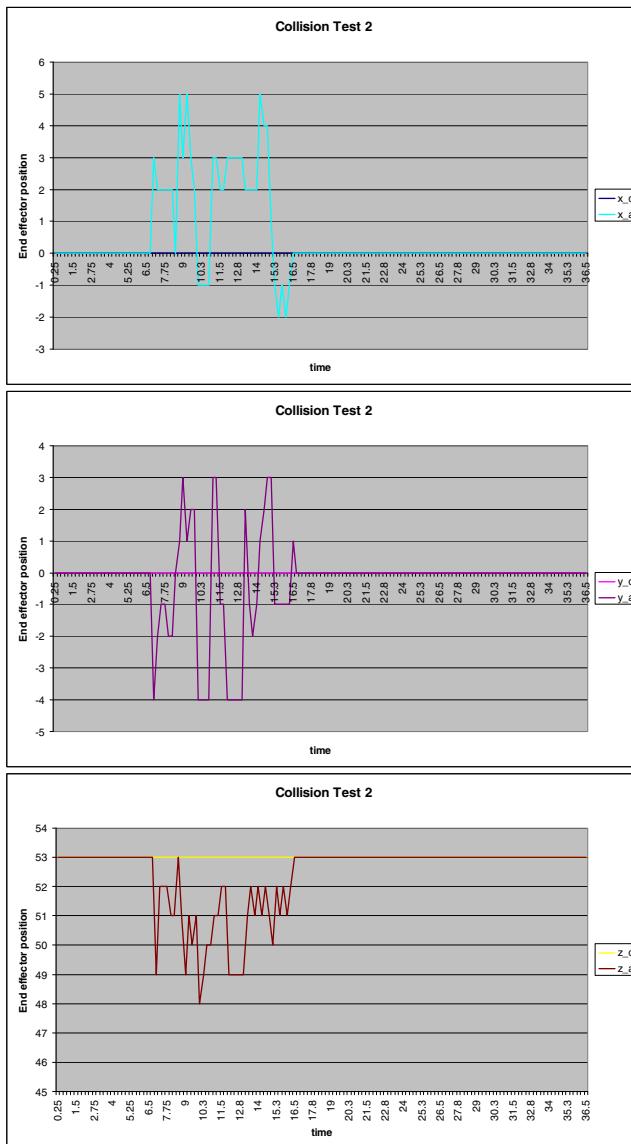
### 5.3 Assessment Results and Discussions

An 8-bit AVR ATMEGA128 microcontroller is used to store data and perform the proposed algorithm. SPI slave configuration port clocked at the CapSense board clock frequency will receive the collision location data from Master-Out-Slave-In (MOSI) pin sent by CapSense SPI master. These data can be collaborated with the viscoelastic trunk deformation data measured by two potentiometers on one side of the trunk in inverse kinematics described in (1) to obtain the end-effectors trajectory in Cartesian space.

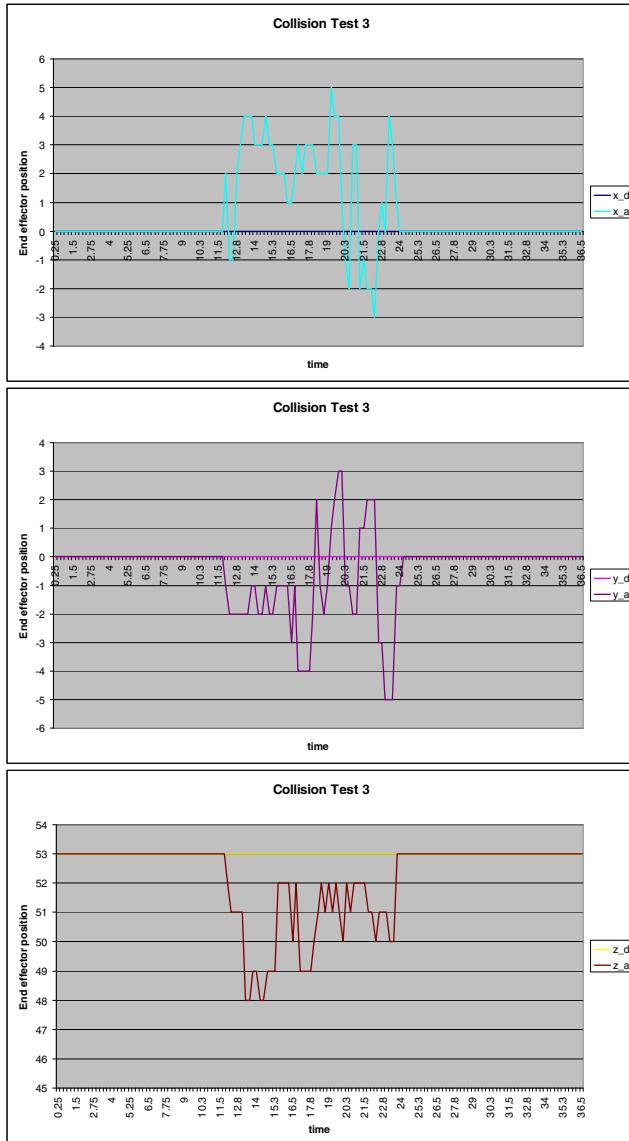
By observing Fig. 7, 8 and 9, where the desired Cartesian coordinates and the actual experimental data are denoted as  $x_d, y_d, z_d$  and  $x_a, y_a, z_a$ , respectively, it is found that during the collisions, the robot arm is able to respond to them using the proposed control algorithm. The end-effector deviations for three cases are found less than 10% which is within acceptable precision range in the social context. The contact forces between the robot arm and the human hand are claimed small due to the passively deformed trunk that has suppressed the big portions of the collision forces.



**Fig. 7.** Comparison of end-effector trajectory in situation 1



**Fig. 8.** Comparison of end-effector trajectory in situation 2



**Fig. 9.** Comparison of end-effector trajectory in situation 3

## 6 Conclusion

For the human-robot coexistence, a collision-tolerant control algorithm is introduced in this paper through the use of a passively deformed viscoelastic trunk, which is able to suppress the unexpected collision force. The hardware implementation of a 3 DOF prototype model integrated with the proposed algorithm has been realized and fabricated

on the basis of human-safety issues. Acquisition of the collision information has been associated with the implementation of a soft capacitive sensing contact model which is capable to sense the contact location by using CMOS technologies. The validity of the proposed system and control approach has been verified through a series of simulation studies and an aggressive assessment. The results have shown the effectiveness of the system in rejecting the unexpected collision while maintaining the end-effector trajectory in the desired range.

## Acknowledgement

The authors would like to thank Mr. Alex Li Ruomu from Social Robotics Laboratory for the previous technical supports. This study was supported by Singapore National Research Foundation, Interactive Digital Media R&D Program, under research grant R-705-000-017-279.

## References

1. Kawamura, K., Iskarous, M.: Trends in Service Robots for the Disabled and the Elderly. In: IEEE/RSJ Intl. Conf. Intelligent Robots and Systems, Germany, pp. 1647–1653 (1994)
2. UAW Health and Safety Department: Review of Robot Injuries - One of the Best Kept Secrets. In: National Robot Conference, Ypsilanti, Michigan, U.S.A, pp. 1647–1653 (2004)
3. Nevins, J.L., Whitney, D.E.: Assembly Research. *Automatica* 16(6), 24–38 (1980)
4. Yoon, S.S., Kang, S.: Safe Arm with MR-based Passive Compliant Joints and Viscoelastic Covering for Service Robot Applications. In: Intl. Conf. Intelligent Robots and Systems, Las Vegas, Nevada, U.S.A, pp. 2191–2196 (2003)
5. Sim, T.S., Lim Jr., K.B., Ang, M.H.: Compliant end-effector coupling (CEEC) - A tool for robotic assembly operations. In: The First Chinese World Congress on Intelligent Control and Intelligent Automation, China, pp. 1078–1084 (1993)
6. Ang Jr., M.H., Andeen, G.B.: Specifying and Achieving Passive Compliance Based on Manipulator Structure. *IEEE Trans. Robotics and Automation* 11(4), 504–515 (1995)
7. McCrum, N.G., Buckley, C.P.: *Principles of Polymer Engineering*, pp. 117–176. Bucknell (2003)
8. Lim, H.: Collision-Tolerant Control of Human-Friendly Robot with Viscoelastic Trunk. *IEEE Trans. Mechatronics* 4(4), 417–427 (1999)
9. Meyers, C.: *Mechanical Behavior of Materials*, pp. 570–580. Prentice Hall Inc., Englewood Cliffs (1999)
10. Lewis, F.L., Abdallah, C.T., Dawson, D.M.: *Control of Robot Manipulators*. Macmillan, New York (1993)

# CAN-Bus Based Distributed Control System for Hydraulic Turbine Blade Repairing Robot

Kangtai Xiang, Zhenguo Sun, Hongjun Dai, Qiang Chen, and Jiajun Liu

Key Laboratory for Advanced Materials Processing Technology, Ministry of Education;  
Dept. of Mechanical Engineering, Tsinghua University, Beijing 100084, China  
xkt08@mails.tsinghua.edu.cn, sunzhg@tsinghua.edu.cn

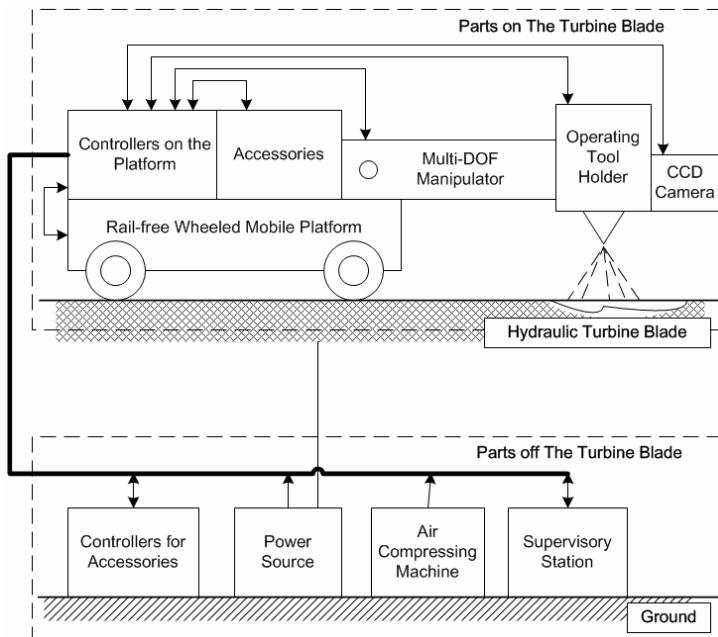
**Abstract.** A CAN-bus (Controller Area Network) based distributed control system was designed for a mobile on-site hydraulic turbine blade repairing robot (HTBRR). The system is composed of one supervisory station, seven joint servo nodes, and four accessory nodes. Visual image sensing, virtual reality and graphic based path planning are implemented on the supervisory station, to improve tele-operation. Each joint servo node is composed of a DSP based control unit and a motor driving module. These nodes are coordinated to realize the motion control of the platform and the multi-DOF manipulator. The accessory nodes are used to control and monitor the apparatus such as welding power source, gouging power source and so on. With a specially designed state machine and the CANopen communication protocol, the designed control system enables HTBRR to carry out various operations, such as air gouging, deposit welding and grinding.

**Keywords:** Distributed control system, mobile robot, virtual reality.

## 1 Introduction

As a part of China's roadmap for a low-carbon energy economy, lots of hydropower stations are constructed and put into operation recently. Due to the serious erosion caused by heavy sediment-laden water However, hydraulic turbine blades of those stations must be maintained regularly to sustain safe and stable operation. Off-site repairing can guarantee optimal blade profile with the help of specially designed automation equipment in the plant [1-3]. But the whole hydraulic turbine runners must be hanged out of the foundation pit, disassembled, and refurbished after repairing. This procedure is really expensive and time-consuming. In the case of manually on-site repairing, scaffolds should be set up at first, and then workers crawl into the runner area to grind the damaged areas and carry out deposit welding. For this reason, welding robots are considered to be the best tools to solve the problems of high labor cost, low efficiency and poor repairing quality. So far, Scompi series robot 345, developed by the research institute of Hydro-Quebec and Invensys of Canada, is the only commercial equipment that can be used to repair the hydraulic turbine blades on-site. It is a track-based six-axis remote-controlled, multi-function robot, which can carry out precision grinding and deposit welding [4, 5]. But the main deficiency of this robot is

that the guiding rails have to be installed and demounted repeatedly according to various damage positions. Furthermore, in the positions with small curvature or constrained space, the rail cannot be installed at all.



**Fig. 1.** Architecture of hydraulic turbine blade repairing robot system

Recently, Tsinghua University and China Yangtze Power Co., Ltd. have cooperated in a research project, aiming to develop a mobile on-site hydraulic turbine blade repairing robot (HTBRR), as shown in Fig.1. HTBRR is composed of a rail-free wheeled mobile platform, a six-axis manipulator arm and an operating tool holder [6]. The operating tools such as welding torch, grinding wheel or air gouging gun may be equipped alternatively onto the tool holder to carry out different repairing processes. An industrial CCD camera is mounted in front of the tool holder for real-time monitoring. With the patented non-contacted magnetic absorption technology, both the mobile platform and the tool holder can self-adaptively absorb and flexibly move on the surface of blade, so as to make the repairing procedures more efficient and flexible [7].

The typical repairing procedure involves several steps: (1) inspecting the damaged areas; (2) air gouging; (3) grinding before welding; (4) deposit welding; (5) grinding the repaired area to their original streamlining profile. HTBRR is expected to carry out all the tasks with the instructions from the operator. Therefore, design of the control system becomes more complex and difficult.

In this paper, CAN bus is introduced to set up the distributed multi-node control system. The master node (Supervisory Station) and all slave nodes (joint servo nodes

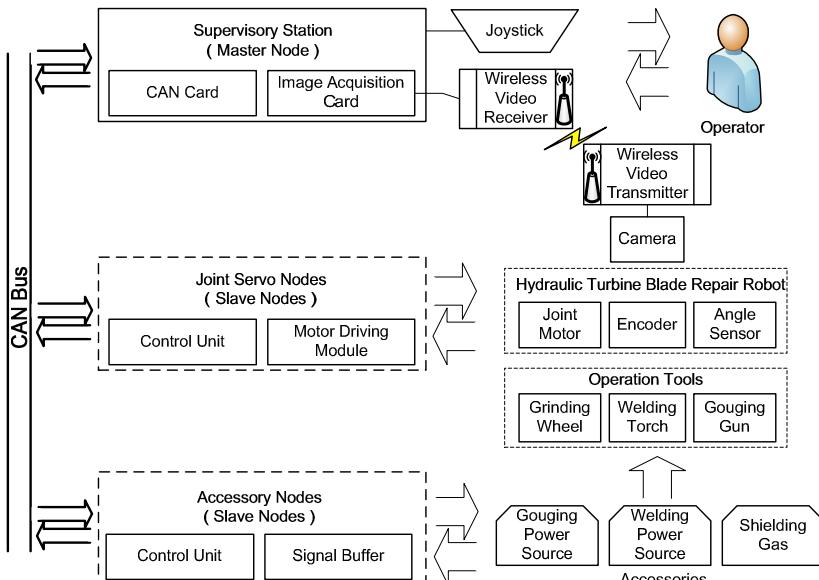
and accessory control nodes) communicate and act synergically under the CANopen protocol [8]. This allows the operator to change HTBRR's function through simply replacing the working tools and corresponding slave nodes.

## 2 Architecture of the Distributed Control System

The distributed control system is composed of a master node and several slave nodes, as illustrated in Fig.2. All nodes are connected together via CAN bus. Installation of more slave nodes is allowed for function extension.

Supervisory Station is used for all human-robot interactions. Image captured by the CCD camera is streamed back to the master node via a wireless module. The captured image is rendered or reserved for further use by Supervisory Station. The Operator can monitor the whole repairing procedure on Supervisory Station, and input manual instructions with a joystick interface. The slave nodes will coordinate to accomplish the tasks under different instructions.

Slave nodes are classified into joint servo nodes and accessory control nodes. Joint servo nodes are mounted on the mobile platform and the manipulator arm to drive actuators to fulfill repairing actions. Accessory control nodes are placed close to the related apparatuses, to control and monitor gouging power source, welding power source, shielding gas supply in this application.



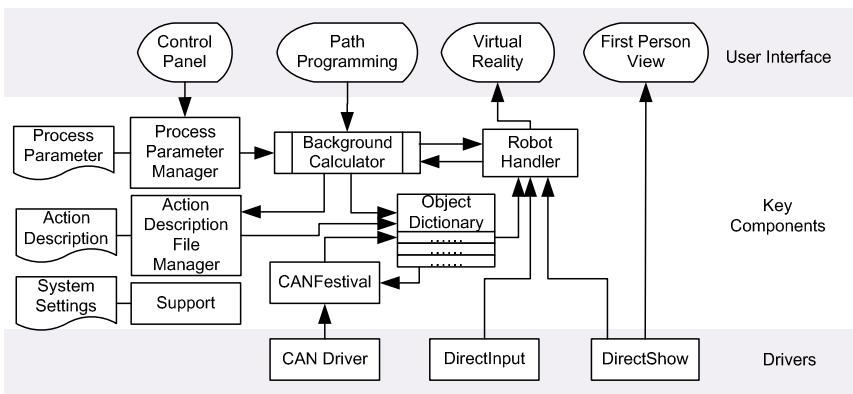
**Fig. 2.** Schematic illustration of the distributed control system

### 3 Hardware and Software Design

#### 3.1 Supervisory Station

Supervisory Station is an industrial computer outside the foundation pit, with an Intel Core Due E7400 2.8GHz CPU, 2GB DDR2 memory, and Microsoft Windows XP. A PCI7841 dual-port CAN card and a PCI image acquisition card with a 640 x 480 resolution are installed for CAN bus communication and image grabbing respectively. A Betop joystick is selected for instruction input. A Partom FOX-R02 wireless video receiver, which works on the 2.4GHz ISM band at 1000mW, is chosen to receive the images from operating field.

A custom single-document interface (SDI) application based on Microsoft Fundamental Class (MFC) serves as the sole human-robot interface for HTBRR. Fig.3 illustrates the framework of this supervisory application.



**Fig. 3.** Framework of the human-robot interface

DirectShow from Windows SDK is introduced to obtain image data from the image acquisition card. Working at buffer mode, SampleGrabber interface of DirectShow makes a copy of the video frame before rendering it. The copy is reserved for tool path planning and 3D simulation. It will be updated when necessary.

DirectInput from DirectX SDK is introduced to map common instructions to buttons and pivots of the joystick. For example, speeds of the two driving wheels of the mobile platform are mapped to the pivots in velocity control mode. This scheme makes position control of the mobile platform and the manipulator arm more flexible. More complex operation, such as moving the end effector referenced in the vision field, involves forward kinematics and backward kinematics calculation of the robot manipulator. That is why a Robot Handler is required.

Robot Handler plays the role of an instruction buffer. All the control instructions are applied to a shadow robot maintained inside the Robot Handler. Based on the current configuration and the pending configuration of the shadow robot, PVTs (position, velocity and time) for joint servo nodes are calculated and distributed. Subsequent

instructions of the same type will be ignored before the error of the pending configuration and the actual one reaches a maximum tolerance. The application reads actual configuration from Object Dictionary, which is maintained by an open source CANopen framework called CANFestival.

A textual file based process parameter database is introduced to save related parameters for different repairing tasks. The operator may modify those parameters through a textual editor or the process parameter manager of the application. These parameters are provided to the action description file to generate tool path.

Action description file manager generates action description files during tool path planning and interprets them during operation. Other components of the application include state machine, system settings manager etc. The state machine is used to coordinate behavior of the entire system (more details are introduced in section 3.4).

### 3.2 Slave Nodes

Each joint servo node is composed of a DSP based control unit and a motor driving module. TMS320F2812 is selected as the core processor of the control unit. External memory spaces are extended with a 256K x 16bit RAM and a 512K x 16bit Flash to meet the requirements of code and data storage. Both brushed dc motors and brushless dc motors are employed as joint actuators. Brushed motor driving module is designed based on LMD18200 as illustrated in Fig.4.a. It is able to support a 3A maximum continuous current at 48V. Brushless motor driver module is designed based on MC33039, as illustrated in Fig.4.b. It can support a 4A maximum continuous current at 48V. Motor driving modules are selected according to the requirements of different joint nodes. These slave joint nodes are coordinated to realize the motion control of the platform and the multi-DOF manipulator.

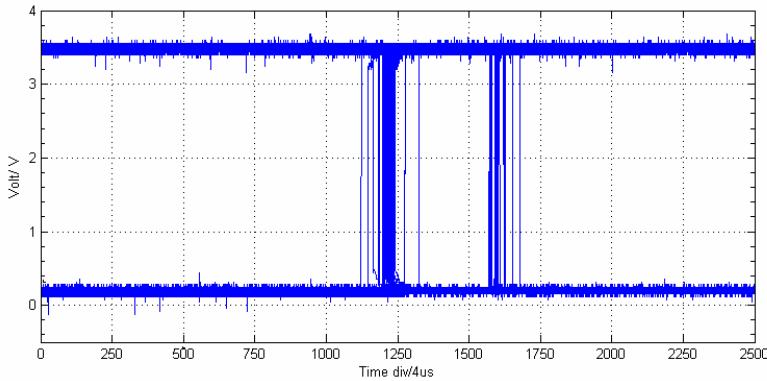


**Fig. 4.** Motor driving modules for joint actuators. (a) Driving module for brushed motor. (b) Driving module for brushless motor.

The accessory nodes are used to control and monitor the apparatus such as welding power source, gouging power source and so on.

DSP/BIOS operation system is used to handle multiple tasks dispatch, data storage, peripherals interrupts, and communication protocols [9]. CANFestival framework has been imported to DSP/BIOS. Synchronization of the slave nodes are guaranteed with a SYNC-Object transmission mechanism [8]. At the slave ends, most received commands

are updated after a succeeding SYNC-Object. Although the master may send this object at irregular intervals, responds of the slaves are synchronized. Fig.5. illustrates the offset in response time of 4 slave nodes within the same bus. The nodes are programmed to toggle an IO after the SYNC object. 20 random samples within 9 minutes are collected and overlapped. In this case the maximum response time difference is 2.2ms which is quite acceptable for control of HTBRR.

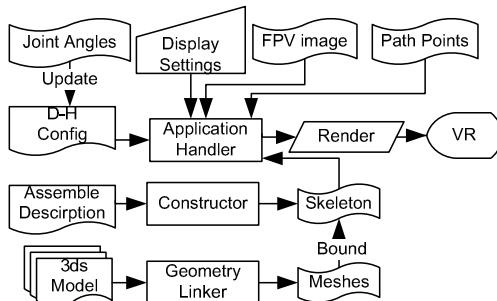


**Fig. 5.** Offset in response time of slave nodes

### 3.3 Virtual Reality Based Tele-presence

The graphical user interface of the supervisory application can be divided into three regions: control panel, first person view (FPV) of the robot and virtual reality (VR) view.

FPV displays the real time vision obtained from the image acquisition card during operation. In this way, the operator is able to observe the exact damage of the blade and decide whether the repairing work is necessary. However, the operator is unable to see the end effector or the posture of the manipulator, because the vision field of the mounted camera is limited to a rectangle in front of the operating tool holder, as illustrated in Fig.1. Therefore, OpenGL based virtual reality is employed, to enhance the tele-presence.

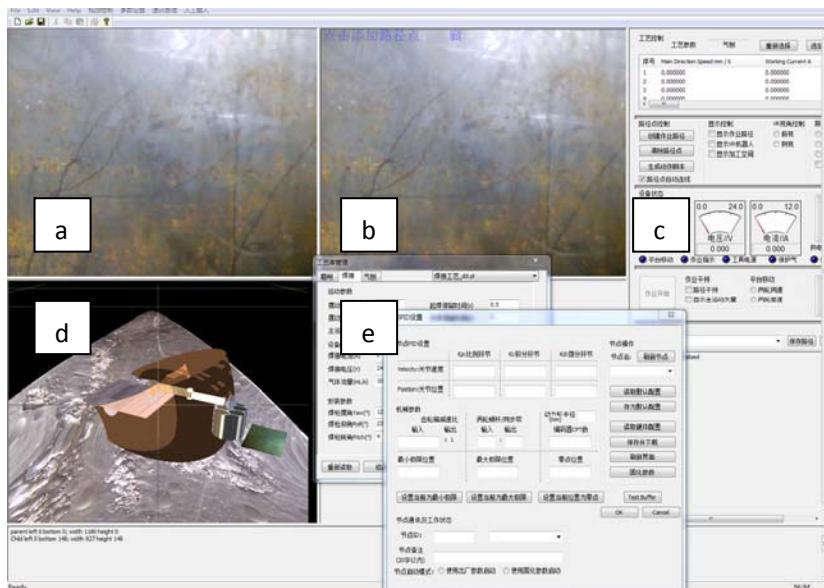


**Fig. 6.** Schematic illustration of virtual reality

Motion of HTBRR on the blade is simulated in VR view. Implement codes are divided into three parts, as shown in Fig.6. Geometry Linker reads ‘3ds’ format model files created by third party CAD software. Constructor reads a configure file containing Denavit-Hartenberg coordinates and initial pose of the robot. During system initialization, all objects in the model files are loaded into the heap, and the skeleton of the robot is constructed. Those objects are further reassembled and bounded to the skeleton correspondingly to form a virtual robot. Application Handler receives instructions from other components of the application and adjusts the model’s posture accordingly, which leads to an interactive simulation.

After camera calibration, the position and size of the vision field relative to the operating tool holder are determined. Before repairing procedure starts, the application associates the texture of the vision field with the FPV image copy on which tool path is planned. The texture remains unchanged throughout the operation. As visibility of any part of the virtual robot is changeable in VR, the operator is able to adjust the perspective, hide the operating tool holder and watch the tool tip moving as programmed on the corroded surface.

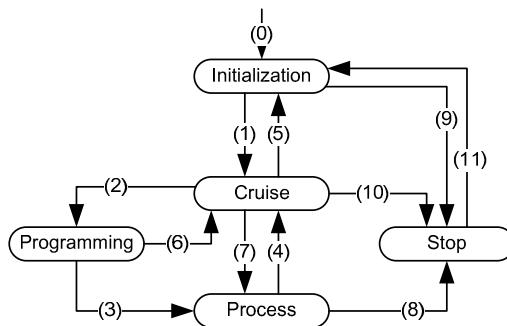
A limitation of this scheme arises from the localization of the mobile platform. Its performance on curved surface is far less satisfactory than on plane surface, if no enough posture feedback is given. Fig.7. is a screen capture of the described supervisory application. In this beta version, a textured blade model is loaded as the working environment (model of the mobile robot is substituted with a prototype copy due to patent consideration). As one can notice that the vision field is floating above the surface of the blade. Improper localization can lead to such weird behavior. Localization of this mobile robot is one of the research topics of the team.



**Fig. 7.** User interface of the supervisory application. (a)First person view of the robot. (b)Tool path planning view. (c)Control panel. (d) Virtual reality view. (e)Configuration dialogs.

### 3.4 State Machine Based Workflow

After booting up, HTBRR will work at one of the five states, as shown in Fig.8. The state machine is maintained by the master node. States of slave nodes are updated simultaneously through communication. Table1 lists all the conversion trigger events and effects. During initialization, when self configuration and system self-test have been completed, the master node downloads initialization parameters to joint servo nodes. Then state machine shifts to cruise mode, which allows the operator to guide the robot to search damaged areas on the blade surface. Joystick through DirectInput eases the manual navigation and posture adjustment. When repairing is necessary, the operator will select a set of process parameters from the database and draw key points of the pending tool path. During this procedure, CRegion from GDI+ is used to indicate boundary of the workspace and to check path points at each step.



**Fig. 8.** State machine based workflow

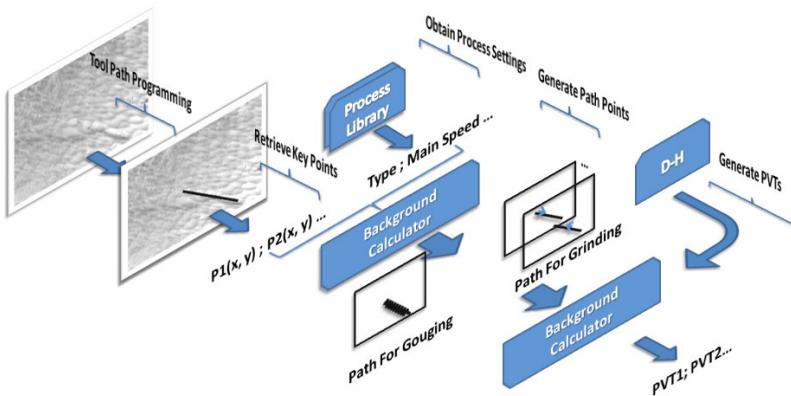
**Table 1.** State conversion events and effects

No	Trigger Event	Effect
0	boot up	lock all
1	initialization complete	unlock all
2	begin programming	lock all
3,7	start process	unlock arm
4	process complete	unlock carriage
5	settings changed	lock all
6	abort programming	unlock all
8,9,10	wrong setting or emergency stop	lock all
11	err caught and reboot	lock all

After all settings are ready, an action description file is created. This file integrates original tool path points, generated joints' PVTs, commands such as 'power on', 'gas on' and so on, together with their sequences and intervals. The way to generate PVTs mainly depends on current process type, as shown in Fig.9.

At first, the application retrieves key points of the pending path. Then these key points together with the process parameters obtained from the process parameter database are sent to the background calculator. A refined tool path for the particular

process type is generated. In the illustrated example, the same key point set results in two different tool paths. If swing is required, the refined path for gouging is like saw tooth along the original direction. The refined path for grinding will be a repeated path with decreasing height and alternate direction. Denavit-Hartenberg coordinates are required for further calculation of PVTs and all mechanical parts must be locked prior to path planning.



**Fig. 9.** PVT generation process

Commands generation also depends on process type. Required commands are pushed to the action description file in the same order of their executing series. The refined tool path stored is used for validation aim. When the operator starts the process, the file is reloaded and all action commands will be executed in series.

## 4 Conclusions

Distributed control system for a mobile robot is designed to fulfill on-site hydraulic turbine blade repairing. Serving as the master node in a CANopen based communication network, the supervisory station for human-robot interaction consists the functions of visual image sensing, virtual reality and graphic based path planning. Two kinds of slave nodes are designed respectively to realize the motion control of mobile robot and monitor the apparatus during repairing procedure. With the assistance of the CAN bus based control system, the HTBRR can carry out repairing process as air gouging, deposit welding and grinding.

## Acknowledgements

The authors gratefully acknowledge the financial support from the National Natural Science Foundation of China (No. 50875147) and the National High Technology Research and Development Program of China (No. 2007AA04Z258).

## References

1. Zichnski, C., Mianowski, K., Narzarczuk, K., et al.: A Prototype Robot for Polishing and Milling Large Objects. *Industrial Robot: An International Journal* 130(1), 67–76 (2003)
2. Huang, H., Gong, Z.M., Chen, X.Q., et al.: Smart Robotic System for 3D Profile Turbine Vane Airfoil Repair. *The International Journal of Advanced Manufacturing Technology* 21, 275–283 (2003)
3. Nonaka, Y., Sakaue, S., Yanagihira, Y., et al.: Development of an Impeller Grinding Robot System and a Gyro-Moment Compensated Compliance Control. In: *IEEE International Conference on Robotics and Automation*, Nagoya, vol. 2, pp. 2084–2089 (May 1995)
4. Fihey, J.L., Hazel, B., Laroche, Y.: The Scompi Technology. In: *International Forum on in Situ Robotic Repair of Hydraulic Turbines*, Harbin, China (2000)
5. SCOMPI Portable Multi-Process Fabrication Robot Datasheet,  
<http://www.solutions2accelerate.com/>
6. Chen, Q., Gui, Z.: A Robot for Welding Repair of Hydraulic Turbine Blade. In: *IEEE International Conference on Robotics, Automation and Mechatronics*, pp. 155–159. IEEE Press, Chengdu (2008)
7. Gui, Z., Chen, Q., Du, D., et al.: A Non-Contact Wheeled Wall-Climbing Robot with Magnetic Adhesion Devices, Chinese Patent: CN 1739925C (March 2006)
8. CiA. CiA DS-301 V4.0.2 – CANopen Application Layer and Communication Profile,  
<http://www.can-cia.org>
9. TMS320 DSP/BIOS User's Guide, <http://focus.ti.com/>

# Visual Assistance to an Advanced Mechatronic Platform for Pick and Place Tasks

Rigas Kouskouridas and Antonios Gasteratos

Democritus University of Thrace,  
Department of Production and Management Engineering,  
Vas. Sofias 12, 67100 Xanthi, Greece  
[{rkouskou,agaster}@pme.duth.gr](mailto:{rkouskou,agaster}@pme.duth.gr)

**Abstract.** Recent advances in mechanical and electronic engineering led to the building of more sophisticated mechatronic systems excelling in simplicity, reliability and versatility. On the contrary, the complexity of their parts necessitate integrated control systems along with advanced visual feedback. Generally, a vision system aims at bridging the gap between humans and machines in terms of providing to the latter information about what is perceived visually. This paper shows how the vision system of an advanced mechatronic framework named ACROBOTER is used for the localization of objects. ACROBOTER develops a new locomotion technology that can effectively be utilized in a workplace environment for manipulating small objects simultaneously. Its vision system is based on a multi-camera framework that is responsible for both finding patterns and providing their location in the 3D working space. Moreover, this work presents a novel method for recognizing objects in a scene and providing their spatial information.

**Keywords:** Intelligent Systems, Multi-camera Systems, Object Recognition, 3D Position Estimation.

## 1 Introduction

In the last few years, robotics researchers aimed at building autonomous vehicles in order to substitute humans in time demanding tasks. Furthermore, industries address all their efforts to developing machines capable of assisting people to everyday life. The need of robotic platforms working closely to humans necessitates the usage of intelligent mechatronics and vision systems. Towards this end, ACROBOTER (Autonomous Collaborative Robots to Swing and Work in Everyday Environment) constitutes one of the most advanced integrated sensorial system. The main concept of the ACROBOTER project is the elaboration of a new locomotion technique that extends the workspace of existing robots in the vertical direction. As a result, it is able to work on the top of tables, wardrobes and can be used for manipulating objects placed on shelves, tables, work surfaces or on the floor. It is apparent that such a robotic platform is capable of moving fast in any 3D direction in an interior working space.

This concept is highly innovative, in the sense that it combines the planar stepping motion of an arm on the ceiling and a 3D controlled pendulum like motion of the working unit. In addition, the ACROBOTER is designed to work autonomously or in close cooperation with humans and to be capable of collaborating with other robotic devices. The ACROBOTER platform consists of mechanical, navigational and control of several subsystems, including a grid of anchor points, a climber unit that moves on the grid of anchor points, and a swinging unit, which is fixed onto the climber unit. The ACROBOTER platform through the swinging unit might carry other service robot or a tool to perform the assigned tasks. The embedded controllers of these main components are integrated with the global control system, while the cameras built onto the four corners of the room provide up to date, navigational information on a dynamically changing working environment. In addition, as part of the global controller, a human machine interface enables the user to interact with and/or to control the ACROBOTER.

The vision system of the ACROBOTER adopts a multi-camera framework in order to provide the visual information needed for the accurate navigation of the robot. Moreover, it includes four individual cameras, installed in the four corners of the room that collaborate to achieve optimum results and cover the 97% of the working volume. In detail, the vision system consists of four Grasshopper cameras manufactured by Point Grey Research that are able to capture images up to 1280 X 960 pixels resolution and are connected to the PC via a firewire port, using the IEEE 1394b transfer protocol. Moreover, the ultimate goal of the vision system is to recognize objects found in the scene and estimating their location in the 3D working space. The aforementioned demanding task is adequately fulfilled by developing and implementing techniques beyond the state-of-the-art as presented in here. The exhibited methodology and results comprise the physical extension of the works presented in the last ICIRA conference [1], [2]. The remainder of this paper is structured as follows: In Section 2, we give an overview about the related work in the areas of multi-camera systems, integrated mechatronic systems and object recognition. The overall system architecture, both hardware and software, of the ACROBOTER is exhibited in Section 3, where every part of the project is presented. Furthermore, in Section 4 a detailed description of the proposed object recognition and position estimation technique is given along with the respective experimental results. Finally, the work concludes with some final notes and discussion in Section 5.

## 2 Related Work

A major challenge in the development of a robot for indoors applications is to equip it with the ability of moving freely and quickly in any direction. In addition, the obstacles found in buildings, specifically in a domestic environment where objects may not be static over time, is a drawback. Examples include stairs, doorsteps, chairs, tables and various surface changes including carpets. Past developments that address the general need for a robot to climb walls and ceiling

inside a building are the MATS robot [3], biped walking robots like the Honda Asimo robot [4] and mobile robot devices such as the Care-o-Bot from IPA [5]. The task to move freely inside a building includes many issues, which are not easily solved. The closest solution to the ACROBOTER concept found in the literature so far is the Flora ceiling based service [6] robot. Flora is suspended onto the ceiling and uses electromagnetic force to keep the mobile cart on the ceiling and telescopic arms to navigate the working unit into the 3D workspace. The permanent magnet traction method and the rigid telescopic structure make this service robot concept less affordable for multipurpose domestic use.

In turn, the ACROBOTER is distinguished by its innovative locomotion concept, which allows to run the entire volume of a room. Additionally it is designed to be able to work autonomously or in close cooperation with humans, and to be capable of collaborating with other robotic devices. In particular, there are two main issues to work with, which are in favor for the ACROBOTER system: autonomous operation / navigation and power. The working principles of biped robots mean that they walk on the floor and thus the navigation in a domestic environment is crucial for successful operation. The power of the system is a limitation compared with the ACROBOTER that can be fed with power in a practically unlimited way.

In the field of multi-camera systems, the Stanford Multi-Camera Array [7] dominated for several years as the ultimate multi-view framework. It consists of 128 cameras that can be arranged and used in a variety of ways. The final camera architecture corresponds to an arc of separate panels aiming at a point in the middle of the room. In turn, the multi-camera system, developed by the UD Graphics Lab [8], consists of 10 Flea2 1394b cameras, where a subset of them are mounted on movable plates attached to the ceiling racks and the others are placed at different positions using tripods. The target applications of the aforementioned framework are multi-view surveillance, capturing of dynamic fluid surfaces or appearance modeling. In addition, the Mitsubishi Electric Research Laboratories (MERL) has designed a framework where the object-wise semantics are extracted from a non-overlapping field-of-view multi-camera system [9]. Its main goal is the construction of an automatic object tracking and video summarization method via background subtraction and mean-shift analysis.

In turn, with a view to the efficient accomplishment of navigation or manipulation tasks, robots must enjoy the perception of depth, for both static and dynamic working environments. As a result, over the past decade, significant research efforts were devoted to the building of autonomous vision systems, capable of providing a sense of location and direction to robots. In manipulation applications the task is to determine objects' position and orientation relative to a specific co-ordinate system. The procedures of obstacle avoidance or objects' manipulation assistance can be accomplished by integrating vital visual information derived from pose estimation techniques. Furthermore, algorithms that were recently proposed are completed by using either visual sensors [10], [11] or the later combined with inertial ones [12]. Another open research topic in computer vision is the depth estimation. In [13] the discrete Fourier Transform (FT) is

utilized to estimate both local and global transformation of the spatial information. On the other hand, one of the most efficient ways to calculate scene's depth is the adoption of a disparity estimation method. Stereo vision frameworks invoke correspondences derived from two slightly different images to extract sufficient spatial information. A recent survey of existing disparity estimation methods is presented in [14].

During the past few years, remarkable efforts were made to build new vision frameworks for robust object recognition in cluttered environments. To this end, researchers emphasized in creating recognition schemes based on local range features [15], [16]. Algorithms of this field extract features with local extent that are invariant to possible illumination, viewpoint, rotation and scale changes. During the past decade, several techniques that enforce the essential role of local features in demanding pattern recognition tasks were presented [17], [18]. The indispensable visual distinctiveness of an object in a scene is ensured by locally sampled descriptions. Furthermore, the two main sub-mechanisms of such frameworks are the detectors and descriptors of areas of interest. The efficiency of the two sub-mechanisms is investigated in [18], where detectors and descriptors are evaluated for object recognition purposes. Currently, the two most widely used object recognition frameworks based on local appearance features, are the Scale Invariant Feature Transform (SIFT) [19] and Speeded-Up Robust Features (SURF) [20]. Both of them implement a detector and a descriptor for fast feature extraction and, as a result, they are adopted in almost any newly appeared object recognition technique. The common issue in both of them is the fact that their detector depicts significant efficiency [18], [21], [22] since it is not affected by possible image alterations. On the contrary, the most important drawback of SIFT's and SURF's descriptor, constitutes the fact that its performance alters significantly under possible rotation, scale, viewpoint and illumination changes. In [23] and [24], two different techniques for object recognition and relative pose coherence are proposed. Both of them utilize subroutines to minimize mismatches between testing images (unknown scenes) and known sets (objects' images contained in databases).

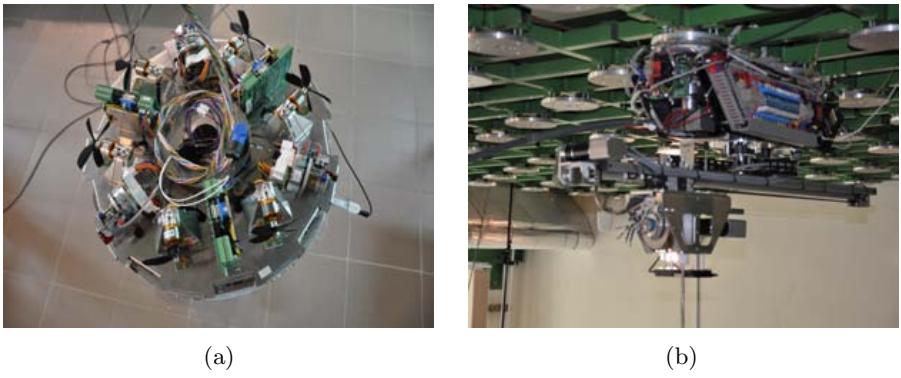
### 3 System Architecture

The ACROBOTER consists of several sub-systems namely the swinging unit, the climber unit, the anchor points and the human-machine interface (HMI). In the following section we briefly describe the main components of the aforementioned frameworks since in this work emphasis is given to ACROBOTER's vision system.

#### 3.1 Swinging Unit

The swinging unit, shown in Figure 1(a), integrates the functions of 3D motion of ACROBOTER (including positioning, orienting and stabilization), tool changing and interfacing. It can serve as a platform for moving other (existing) service

robots or it can be equipped with different kinds of end-effectors in order to accomplish specific tasks. Furthermore, the swinging unit incorporates additional mechanical and control related principles. Three windable cables are used to keep the unit in horizontal position. These cables are also used to position the swinging unit vertically, which result in a redundant actuation with improved manoeuvrability and obstacle avoidance. Moreover, the required thrust is provided by ducted fans placed on the top surface of the platform. In addition, the vectoring and reversing of the thrust are provided by two ducted fans with adjustable blade pitch impellers.



**Fig. 1.** (a) The swinging unit of the ACROBOTER serving as a platform. (b) The climber unit which is responsible for the overall movement of the robot.

### 3.2 Climber Unit

The climber unit (see Figure 1(b)) consists of the following sub-systems: Anchor arm with anchor points, rotation arm and winding mechanism. These components give the climbing unit the ability to position the swinging unit, both horizontally and vertically in the environment. The climber unit is also responsible to lead electric power from the energy source in the ceiling to the swinging unit. The climber unit itself is also powered by the same energy source. The main control tasks performed by the climber unit involve the active tracking control of the 3D trajectories along with the contribution to the compensation of real-time errors caused by simplified idealized motion planning.

### 3.3 Anchor Points

The anchor points system is the first system in a chain of mechanically interconnected systems and it is illustrated in Figure 2(a). The tasks of this subsystem are to support, rotate and swing the bulk of all subsystems on the anchor points that are attached on the roof of a specified structure. The subsystem is structurally divided into two parts: static anchors attached to the ceiling that provide

the necessary support to the system (Anchor Points), and a mobile part (Anchor Arm) that enables rotation around the anchor points and displacement over the ceiling. The Anchor points are passive units made of structural steel. They provide structural support and electrical power to the ACROBOTER system. The Anchor Arm consists of a static beam that provides the necessary support from the anchor points and a driving part that enables rotation about the anchor point.



**Fig. 2.** (a) The anchor points of the ACROBOTER. (b) (Left) User interacting with Samsung Q1 tablet device, (Right) HMI interface computer showing the simulator running to validate a user selected task.

### 3.4 Human Machine Interface

The HMI acts as a gateway between the user and the task scheduling system, necessitating close integration between the two systems. Whilst the interface implementation will be specific to the candidate scenario of the ACROBOTER robot, several candidate HMI concepts have been developed which are applicable to a range of scenarios. A mobile interface device (Samsung Q1 ultra-portable PC), which is depicted in Figure 2(b), is utilized for controlling the robot since it allows the user to freely move around the robot's environment whilst issuing commands. This device is equipped with a 1GHz Intel processor, 1GB of RAM, and an Intel integrated graphics processor. This specification is capable of running the WPF based user interface, as well as the simulator. Communication with the rest of the control system is performed by an integrated WiFi link.

## 4 3D Object Recognition and Position Estimation

The overall goal of ACROBOTER's vision system is the simultaneously recognition of possible objects found in the working space of the robot and the 3D location estimation of the sought objects. It comprises of four cameras installed

in the respective corners of the room. The proposed position estimation process is motivated by the idea that the detected features are located on given geometric positions, thus they can be considered as the corners of a polyhedron, the centre of gravity of which is computed and it is associated to the actual centre of mass of the sought object. Once the features' center of mass is known and the object is recognized, the distance of the object from the camera is trivial, given at least one recorded position of the object. Practically, the proposed technique tries to estimate the geometrical proportion between an image of a known object (contained in the database) and another one of a scene containing the same object in different arrangements. By taking into account the distribution of object's features around their centre in the image space we can estimate spatial information about the object. Especially, we are able to calculate object's distance from the camera along with its location relative to the camera's coordinate system. Afterwards, since there is a pre-knowledge about the cameras' position in space (calculated through the calibration matrix - extrinsic parameters of the camera) we estimate objects' location relative to the global coordinate system of the robot. This is accomplished by multiplying a transformation matrix to the one extracted previously. As a result, by comparing object's image to that of a scene's containing it we are able to transform information from image space to the real world. The main idea underlying our algorithm is the maintenance of the properties of SIFT, whilst making an attempt to further exploit them in order to assign the distance of an object from the camera's frame. The main stages of the proposed algorithm are as follows:

- Stage I** *Apply SIFT to the scene's and object's image, in order to estimate the features position in each of them.*
- Stage II** *Obtain the  $N$  features that match in the two images by applying the matching sub-procedure of SIFT.*
- Stage III** *Find the features' centers of mass for both images. This is accomplished by estimating the mean values of the features positions in the two images.*
- Stage IV** *Calculate the mean Euclidian distance (in pixels) of each feature from the corresponding center of mass that is extracted in the previous stage. Set as  $E_S$  and  $E_O$  the mean Euclidian distances in the scene and object image, respectively.*
- Stage V** *Estimate  $d_S$  which corresponds to the ratio of the two mean distances  $E_S$  and  $E_O$ . Furthermore, we introduce the pre-computed depth  $d_O$ , which is obtained, during the training session and while the object is captured alone.*
- Stage VI** *The previously extracted spatial information relative to the camera's plane are transformed to the global X,Y and Z of the object.*

Stage 1 could be apprehended as the training session of our algorithm. In this phase, for each image in the database keypoint features are extracted using SIFT. Each object is captured at different distances from the camera and the pre-computed depth  $d_O$  is stored for further use. This process is performed while the system is offline, thus, executable time is not taken into account. The results are stored for further use at the next phases. In Stage 2, the matching sub-procedure of SIFT is performed. Especially, descriptors that are common in both images (scene and object) are extracted. It is apparent that, one image representing the scene is compared with several others, representing the object from different viewpoints. Furthermore, the locations of the common features are stored for further exploitation. In Stage 3, the position estimation sub-procedure takes place till the end of the algorithm. Moreover, at this phase, the features' centers of mass in both images are calculated. The last is obtained by estimating the mean values of features locations in both representations. In Stage 4, the distance of each keypoint from the center of mass is calculated. This is measured in pixels with the use of Euclidian Distance. By the end of this sub-routine, we are able to collect significant spatial information of an object in a scene. This is accomplished by simply estimating the distribution of trained features around their center of mass. Finally, in Stage 5, the object's distance from the camera is computed. The pre-computed depth  $d_O$  measured during the training session (Stage 1), is taken into account. The ratio  $d_S$  is used to measure the proportion of object's features to those found in the scene. After the necessary training session and the database construction at the initial stages of the method, an online search engine follows. This is responsible for querying in the scene for objects contained in the trained database. When an object is found, the scene's image is compared to this object, which provides the majority of common matches. Finally, features' information from both images is interpolated with a view to object's position allocation.

The proposed method was exhaustively tested at the final integration meeting of the project. The scene captured corresponds to the real working space conditions under which the ACROBOTER will operate. The cameras have been successfully installed in the four corners of the room and calibrated in order to calculate the extrinsic and intrinsic parameters. Furthermore, the multi-camera network has been adequately established through the respective IEEE 1394b firewire ports and the necessary data related bandwidth adjustment has been accomplished. The dimensions of the cameras' array are: length 6.2 m, width 4.4 m and the height of each camera from the ground is 3.3 m. In Figure 3 the four viewpoints of the multi-camera system are illustrated. The tests were executed on a typical PC with a core2duo 2.2 GHz processor, 2 GB RAM and Windows XP operating system. Furthermore, the procedure executes in approximately 4 sec per camera. It is apparent that the overall computational cost of the position estimation process is mainly depended on the trained images per object. An over-enhanced database results into high execution times. In detail, the proposed method requires approximately 0.8 sec per training image in order to estimate the necessary features extracted through the recognition process.



**Fig. 3.** The same scene captured from four different viewpoints. The 3(a), 3(b), 3(c) and 3(d) correspond to camera no. 1 viewpoint, camera no. 2 etc.

The proposed method initially aims at estimating object's distances from the cameras. This estimation provide to the robot raw information concerning the relative position of the object to the camera. The results of this first phase are shown in Figure 4. The efficiency of the proposed technique remains at very high standards and especially the mean value of the distance estimation is 93.4 % and the standard deviation oscillates around 5.2. In turn, as a next step the initial spatial information derived is taken into account for the relative and absolute location assignment of the recognized object. Figure 5 depicts the estimations derived by the second phase of the proposed method. The X, Y and Z positions of the object relative to the camera are estimated through the calculated distance of the first phase. In turn, the absolute object's position are extracted by multiplying the transformation matrix T of each camera with the relative coordinates. Once more the efficiency concerning the relative to camera coordinates remains high with mean value 86.7% and with standard deviation around 9.34. As far as the absolute coordinates are concerned, the proposed method's yield has a mean value 84.4 % and a standard deviation 10.7. It is obvious that concerning the fourth camera a decrease of the algorithms efficiency is noted. This is due to the fact that scene as captured from camera 4, contains several textured objects that are incorrectly matched to the sought object.

		camera 1			camera 2		
		measured	groundtruth	efficiency	measured	groundtruth	efficiency
<b>Distance</b>		618	635	97.3%	485	510	95.1%
			camera 3			camera 4	
		measured	groundtruth	efficiency	measured	groundtruth	efficiency
<b>Distance</b>		365	350	95.7%	455	530	85.8%

**Fig. 4.** The proposed method is able to estimate object's distance from the camera with remarkable efficiency

<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #ADD8E6;">Coordinates relative to camera 1</th> <th>measured</th> <th>groundtruth</th> <th>efficiency</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>360</td> <td>380</td> <td>94.7%</td> </tr> <tr> <td>Y</td> <td>445</td> <td>480</td> <td>92.7%</td> </tr> <tr> <td>Z</td> <td>248</td> <td>260</td> <td>95.4%</td> </tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #ADD8E6;">Coordinates relative to robot</th> <th>measured</th> <th>groundtruth</th> <th>efficiency</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>355</td> <td>380</td> <td>93.4%</td> </tr> <tr> <td>Y</td> <td>175</td> <td>200</td> <td>87.5%</td> </tr> <tr> <td>Z</td> <td>52</td> <td>70</td> <td>74.3%</td> </tr> </tbody> </table>	Coordinates relative to camera 1	measured	groundtruth	efficiency	X	360	380	94.7%	Y	445	480	92.7%	Z	248	260	95.4%	Coordinates relative to robot	measured	groundtruth	efficiency	X	355	380	93.4%	Y	175	200	87.5%	Z	52	70	74.3%	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #ADD8E6;">Coordinates relative to camera 2</th> <th>measured</th> <th>groundtruth</th> <th>efficiency</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>67</td> <td>60</td> <td>88.3%</td> </tr> <tr> <td>Y</td> <td>425</td> <td>480</td> <td>88.5%</td> </tr> <tr> <td>Z</td> <td>299</td> <td>260</td> <td>85.0%</td> </tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #ADD8E6;">Coordinates relative to robot</th> <th>measured</th> <th>groundtruth</th> <th>efficiency</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>410</td> <td>380</td> <td>92.1%</td> </tr> <tr> <td>Y</td> <td>190</td> <td>200</td> <td>95.0%</td> </tr> <tr> <td>Z</td> <td>80</td> <td>70</td> <td>85.7%</td> </tr> </tbody> </table>	Coordinates relative to camera 2	measured	groundtruth	efficiency	X	67	60	88.3%	Y	425	480	88.5%	Z	299	260	85.0%	Coordinates relative to robot	measured	groundtruth	efficiency	X	410	380	92.1%	Y	190	200	95.0%	Z	80	70	85.7%
Coordinates relative to camera 1	measured	groundtruth	efficiency																																																														
X	360	380	94.7%																																																														
Y	445	480	92.7%																																																														
Z	248	260	95.4%																																																														
Coordinates relative to robot	measured	groundtruth	efficiency																																																														
X	355	380	93.4%																																																														
Y	175	200	87.5%																																																														
Z	52	70	74.3%																																																														
Coordinates relative to camera 2	measured	groundtruth	efficiency																																																														
X	67	60	88.3%																																																														
Y	425	480	88.5%																																																														
Z	299	260	85.0%																																																														
Coordinates relative to robot	measured	groundtruth	efficiency																																																														
X	410	380	92.1%																																																														
Y	190	200	95.0%																																																														
Z	80	70	85.7%																																																														
<b>(a)</b>	<b>(b)</b>																																																																
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #ADD8E6;">Coordinates relative to camera 3</th> <th>measured</th> <th>groundtruth</th> <th>efficiency</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>70</td> <td>60</td> <td>83.3%</td> </tr> <tr> <td>Y</td> <td>150</td> <td>140</td> <td>92.9%</td> </tr> <tr> <td>Z</td> <td>259</td> <td>260</td> <td>99.6%</td> </tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #ADD8E6;">Coordinates relative to robot</th> <th>measured</th> <th>groundtruth</th> <th>efficiency</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>396</td> <td>380</td> <td>95.8%</td> </tr> <tr> <td>Y</td> <td>176</td> <td>200</td> <td>88.0%</td> </tr> <tr> <td>Z</td> <td>74</td> <td>70</td> <td>94.3%</td> </tr> </tbody> </table>	Coordinates relative to camera 3	measured	groundtruth	efficiency	X	70	60	83.3%	Y	150	140	92.9%	Z	259	260	99.6%	Coordinates relative to robot	measured	groundtruth	efficiency	X	396	380	95.8%	Y	176	200	88.0%	Z	74	70	94.3%	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #ADD8E6;">Coordinates relative to camera 4</th> <th>measured</th> <th>groundtruth</th> <th>efficiency</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>341</td> <td>380</td> <td>89.7%</td> </tr> <tr> <td>Y</td> <td>205</td> <td>140</td> <td>53.6%</td> </tr> <tr> <td>Z</td> <td>200</td> <td>260</td> <td>76.9%</td> </tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #ADD8E6;">Coordinates relative to robot</th> <th>measured</th> <th>groundtruth</th> <th>efficiency</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>295</td> <td>380</td> <td>77.6%</td> </tr> <tr> <td>Y</td> <td>255</td> <td>200</td> <td>72.5%</td> </tr> <tr> <td>Z</td> <td>100</td> <td>70</td> <td>57.1%</td> </tr> </tbody> </table>	Coordinates relative to camera 4	measured	groundtruth	efficiency	X	341	380	89.7%	Y	205	140	53.6%	Z	200	260	76.9%	Coordinates relative to robot	measured	groundtruth	efficiency	X	295	380	77.6%	Y	255	200	72.5%	Z	100	70	57.1%
Coordinates relative to camera 3	measured	groundtruth	efficiency																																																														
X	70	60	83.3%																																																														
Y	150	140	92.9%																																																														
Z	259	260	99.6%																																																														
Coordinates relative to robot	measured	groundtruth	efficiency																																																														
X	396	380	95.8%																																																														
Y	176	200	88.0%																																																														
Z	74	70	94.3%																																																														
Coordinates relative to camera 4	measured	groundtruth	efficiency																																																														
X	341	380	89.7%																																																														
Y	205	140	53.6%																																																														
Z	200	260	76.9%																																																														
Coordinates relative to robot	measured	groundtruth	efficiency																																																														
X	295	380	77.6%																																																														
Y	255	200	72.5%																																																														
Z	100	70	57.1%																																																														
<b>(c)</b>	<b>(d)</b>																																																																

**Fig. 5.** The absolute and relative positions of the found object along with the proposed method's attributes

## 5 Conclusions

In this paper an intelligent multi-camera agent responsible for providing the necessary visual information to an advanced mechatronic system is presented. The vision system of the ACROBOTER project comprises of four cameras installed in the four corners of a room and it is responsible for two tasks that affect the overall efficiency of the system. Initially, it aims at the adequate recognition of objects found in the scene and are already contained in a database. The ultimate goal is to provide to the robot the position of the sought objects in its 3D working space. Experimental results representing those obtained during the

final integration meeting prove our claim that once the features' center of mass is known and the object is recognized, the distance of the object from the camera is trivial, given at least one recorded position of the object. This in turn yields to an accurate location assignment of the object and allows the swinging unit to fly over the object and manipulate it. Moreover, the proposed multi-camera framework can be exploited in several other indoor applications where simultaneous recognition and localization of objects is prerequisite. Finally, we firmly believe that due to the fact that all the sub-systems utilize methods, which are beyond the state-of-the-art, and its anthropocentric concept, ACROBOTER will make a great breakthrough at the field of autonomous mobile assistant robots.

**Acknowledgements.** This work is supported by the E.C. under the FP6 research project for Autonomous Collaborative Robots to Swing and Work in Everyday EnviRonment ACROBOTER, FP6-IST-2006-045530.

<http://www.acroboter-project.org>

## References

1. Kouskouridas, R., Kyriakoulis, N., Chrysostomou, D., Belagiannis, V., Mouroutsos, S., Gasteratos, A.: The Vision System of the ACROBOTER Project. In: International Conference on Intelligent Robotics and Applications, pp. 957–966 (2009)
2. Kouskouridas, R., Badekas, E., Gasteratos, A.: Simultaneous Visual Object Recognition and Position Estimation Using SIFT. In: Xie, M., Xiong, Y., Xiong, C., Liu, H., Hu, Z. (eds.) ICIRA 2009. LNCS, vol. 5928, pp. 866–875. Springer, Heidelberg (2009)
3. Balaguer, C., Gimenez, A., Huete, A., Sabatini, A., Topping, M., Bolmsjo, G.: The MATS robot: service climbing robot for personal assistance. IEEE Robotics & Automation Magazine 13(1), 51–58 (2006)
4. <http://www.honda.co.jp/ASIMO/>
5. <http://www.care-o-bot.de/english/>
6. Sato, T., Fukui, R., Morishita, H., Mori, T.: Construction of ceiling adsorbed mobile robots platform utilizing permanent magnet inductive traction method. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2004, vol. 1 (2004)
7. Vaish, V., Levoy, M., Szeliski, R., Zitnick, C., Kang, S.: Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2 (2006)
8. Zhang, J., McMillan, L., Yu, J., Hill, U.: Robust tracking and stereo matching under variable illumination. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1 (2006)
9. Zwicker, M., Vetro, A., Yea, S., Matusik, W., Pfister, H., Durand, F.: Resampling, antialiasing, and compression in multiview 3-D displays. IEEE Signal Processing Magazine 24(6), 88–96 (2007)
10. Schweigofer, G.: Robust pose estimation from a planar target. IEEE Trans. Pattern Anal. Mach. Intell. 28(12), 2024–2030 (2006)
11. Chandraker, M., Stock, C., Pinz, A.: Real-time camera pose in a room. In: Crowley, J.L., et al. (eds.) ICVS 2003. LNCS, vol. 2626, pp. 98–110. Springer, Heidelberg (2003)

12. Xu, D., Li, Y.F.: A new pose estimation method based on inertial and visual sensors for autonomous robots. In: IEEE International Conference on Robotics and Biomimetics, Sanya, China, pp. 405–410 (2007)
13. Torralba, A., Oliva, A.: Depth estimation from image structure. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1226–1238 (2002)
14. Naplapntidis, L., Sirakoulis, G., Gasteratos, A.: Review of stereo vision algorithms: from software to hardware. *International Journal of Optomechatronics* 2(4), 435–462 (2008)
15. Nister, D., Stewenius, H.: Scalable recognition with a vocabulary tree. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, USA, pp. 2161–2168 (2006)
16. Sivic, J., Zisserman, A.: Video google: a text retrieval approach to object matching in videos. In: Proceedings of the Ninth IEEE International Conference on Computer Vision, Nice, France, pp. 1470–1477 (2003)
17. Liao, M., Wei, L., Chen, W.: A novel affine invariant feature extraction for optical recognition, vol. 3, pp. 1769–1773 (August 2007)
18. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* 27(10), 1615–1630 (2005)
19. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60(2), 91–110 (2004)
20. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (surf). *Comput. Vis. Image Underst.* 110(3), 346–359 (2008)
21. Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Gool, L.V.: A comparison of affine region detectors. *Int. J. Comput. Vision* 65(1-2), 43–72 (2005)
22. Mikolajczyk, K., Schmid, C.: Scale & affine invariant interest point detectors. *Int. J. Comput. Vision* 60(1), 63–86 (2004)
23. Ferrari, V., Tuytelaars, T., Van Gool, L.: Simultaneous object recognition and segmentation by image exploration. In: Proceedings of the 8th European Conference on Computer Vision, Prague, Czech Republic, pp. 40–54 (2004)
24. Moreels, P., Perona, P.: A probabilistic cascade of detectors for individual object recognition. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) *ECCV 2008, Part III*. LNCS, vol. 5304, pp. 426–439. Springer, Heidelberg (2008)

# An Approach to Multi-robot Site Exploration Based on Principles of Self-organisation

Lyuba Alboul<sup>1</sup>, Hussein S. Abdul-Rahman<sup>1</sup>, Paul S. Haynes<sup>1</sup>,  
Jacques Penders<sup>1</sup>, and Julien Tharin<sup>2</sup>

<sup>1</sup> Centre for Automation and Robotics Research, Sheffield Hallam University, Sheffield, UK  
`{L.Alboul, H.Abdul-Rahman, P.Haynes, J.Penders}@shu.ac.uk`

<sup>2</sup> K-TEAM, Yverdon-les-Bains, Switzerland  
`jtharin@k-team.com`

**Abstract.** This paper presents a novel approach to multi-robot site exploration and map building considering the robot team as a self-organising system. The approach has been developed within the framework of the project GUARDIANS. The Map Building process represents not a separate activity, but an inherent by-product of self-organisation. The system consists of an (heterogeneous) robot swarm, a mobile ad-hoc network and an (evolving) topological map of the environment. The proposed map building approach takes advantage of a cooperating robot team (as opposed to a single robot) allowing accurate deployment and localisation in a structured, yet dynamic manner. A topological graph representation of the environment is formed, from which an initial metric representation is elicitable as edges are assigned lengths. This reasonable sketch of the environment can be further developed to a full metric map and be used as the basis of building ad-hoc mobile wireless communication and sensor networks. The presented algorithms also take into consideration sensor limitation and are tested on a group of Khepera III robots, specially upgraded to fulfil the needs of our approach.

**Multi-robot team:** Self-organising system, topological representation, relative positioning, cooperative map building, ad-hoc communication network.

## 1 Introduction

In the GUARDIANS project<sup>1</sup>, autonomous robots are designed to assist fire-fighters in search and rescue operations in an industrial warehouse in the event or danger of fire [14]. Robot tasks can be roughly split into two partially overlapping categories. The tasks of the first category are to directly assist the fire-fighters by a robot team, for example by guiding or accompanying a fire fighter and indicating possible obstacles and locations of danger. The second category comprises the tasks for a robot team acting without a human squad-leader, such as on site deployment, positioning as beacons and maintaining communication. Robots are also required to gather relevant environmental information. This paper focuses on the tasks for GUARDIANS robots in the second

<sup>1</sup> GUARDIANS, Group of Unmanned Assistant Robots Deployed in Aggregative Navigation supported by Scent Detection, EU FP6 ICT 045269.

category, and proposes a new approach for on site robot deployment for setting up a communication network and building a map of the environment.

Maintaining communication faces two major problems. First, metal cages generally present in a warehouse render radio reception problematic. The second problem concerns position detection or localisation. For indoor environments GPS is not generally available and simultaneous localisation and mapping (SLAM) has to be based on other sensors. However, the presence of smoke hinders conventional light based sensors which may not produce useful data (e.g. cameras). The radio signal for wireless communication is not disturbed by smoke and serves as a coarse fall back. For communication various wireless technologies are available including Wireless LAN, Bluetooth and ZigBee.

The presented approach assumes LRF is still possible, which has sense even in the presence of smoke since smoke develops close to the ceiling and the robots involved (Erratics or Khepera III) are (very) small in height. The main purpose of the LRF is to facilitate robot detection and distance estimation, which eventually yields accurate robot localisation and positioning; detection is backed up by wireless communication. However, the idea behind the approach is generic, and the LRF can be substituted by other sensors or sensor system when necessary, such as the system based on infrared developed in [15]. Besides the LRG and WiFi, odometry is also taken into consideration. The supplementary and more conventional use of the LRF is for obstacle avoidance and metric map building. However, the main result of our approach is a topological representation of the environment; the robots play active roles in constructing such a representation, being ‘dynamic nodes’ in the representation as well as acting as landmarks. One novel feature is that the topological map is the first to be built, that can later be upgraded to a more detailed metric representation. In the robotics literature topological approaches are in general not viewed as independent, as topological maps are built on top of grid-based map or feature-based maps by partitioning them into coherent regions [10].

Algorithms have been implemented in C++ with Player/Stage middleware<sup>2</sup>. Player can control both real and ‘simulated’ robots thus allowing direct application of simulation algorithms to real-life scenarios. Khepera III robots, are used, produced by the partner K-Team, and upgraded specifically by installing the LRF Hokuyo; both models URG-04LX and URG-04LX-UG0 can be used [6]. The set-up for a real-life implementation is in the final stage of development. A bespoke TCP/IP data communications suite was also developed, which is independent of Player.

## 2 Related Work

The problem of global self-localisation with no *a priori* environmental information, is considered one of the most difficult in robotics. It is related to the famous SLAM problem of a robot simultaneously localising and building a map of the environment. Most existing works on the subject focus on a single robot only, and the approaches are mainly probabilistic in their nature due to environmental uncertainty.

---

<sup>2</sup> <http://playerstage.cvs.sourceforge.net/viewvc/playerstage/papers/>

In the last decade, several works appeared tackling the problem of cooperative multi-robot localisation [4] [16] [8], and which fall mostly into three categories. The first category, related to the multi-SLAM problem, builds separate maps which are later fused to form a global map [5]. The second category relates to on site robot distribution by using, for example, artificial potential fields [2], or radio signal strength [3]. When the number of robots is sufficiently large, these approaches can achieve a large coverage area, but links between robots can be lost and distributions may be irregular thus preventing correct estimation of distances and robust localisation. The third category, which partially overlaps the first two, takes full advantage of a multi-robot system. Robots are used as temporary landmarks for mutual recognition and localisation, resulting in a more accurate representation of the environment.

Our approach belongs to the third category. Robots represent a self-organising system consisting of both mobile and permanent or temporary beacon robots. On site distribution is achieved by way of nodes connected by virtual edges, thus describing the environment by means of a graph. Graph nodes are either virtual or real (robots) and their positions are calculated by applying a well-defined schema of robot movement and detection. Depending on the number of robots some robots can become static beacons used for building a robust ad-hoc communication network.

From a theoretical point of view, our method, to a certain extent, represents a fusion and further development of strategies proposed in [13] and [16]: rules for robot movement are similar to those in [13], but more uniform in that they allow further error decrease in robot self-localisation; our strategy for ‘sweeping’ the environment has some similarity to that in [16], but more elaborated. It is based on an initial virtual grid that allows building a graph representation of the environment simultaneously with robot movement.

Our method has been implemented on real robots with real sensors, using the Player middleware. Three robots are required for reliable robot position detection and topological representation. Four or five robots can provide a more robust system, as these robots can be stationary beacons and used as points of reference, thus increasing accuracy of robot localisation.

We also propose a robot detection method, based on the robot shape and size, and taking into account that robots are ‘dynamic’ objects. This detection system can be further improved by labelling each robot with a special retro-reflective tag that can be detected by a laser range finder based on the intensity of the reflected laser beam. A similar approach was used in the works of Howard *et al* on multirobot simultaneous localisation and mapping [8].

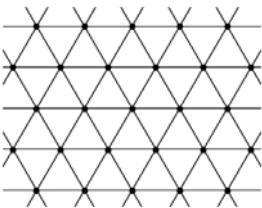
### 3 Topological Map Building Framework

Our approach is a combination of guided positioning and distribution strategies. Each robot has two fields of view, one related to the LRF and the second to WiFi. Their ranges differ considerably: LRF range is either 4 or 5.6m depending on the model, whereas the transmission range of a reliable radio signal (IEEE 802.11) is 30m or more.

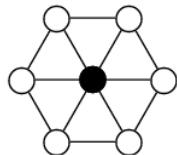
### 3.1 Main Concepts behind the Schema

A discrete graph-theoretic approach is adopted, which considers robots as (dynamic) graph nodes and their movements along edges of the graph. LRF is used as the main sensor for position detection with radio signal as backup. Coordination of robots whilst correcting for odometry errors then becomes more manageable and a cooperative exploration algorithm has been developed.

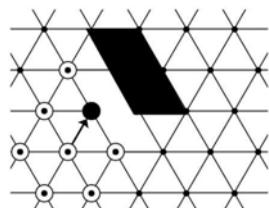
The site is initially covered by a virtual triangular grid (triangular tiling), depicted in Fig. 1. The grid spans infinitely in all directions, and as robots explore the site local parts of the grid become actualized. Each robot forms an attainable visibility domain (*AVD*), determined *a priori*. Clearly, it cannot be larger than the robot sensor range. An example AVD domain is depicted in Fig. 2. The domain is hexagonal, where black nodes indicate robots and white nodes areas the robot can ‘sense’. The sensed nodes are also the grid nodes to which the robot can move. It is not necessary that the robot visits all sensed nodes, since by sensing alone, nodes are effectively explored. The virtual AVD eventually becomes actualized following robot on site deployment. The robot moves along the edges of the grid to available nodes, transforming visited nodes into a (topological) graph referred to as the *initial topological sub-graph of the environment* or *ITSG*. The ITSG expands as the robot moves to the AVD periphery. If the robot encounters an obstacle, all incident edges of the node whose location ‘falls’ on the obstacle are removed. However, the node itself is kept as it may happen that its position is ‘behind’ the obstacle and therefore can be explored later on. However, if the node as it will turn out later during exploration, is really ‘within’ the obstacle it will be removed from the ITSG (Fig. 3).



**Fig. 1.** Virtual triangular grid



**Fig. 2.** AVD of a robot



**Fig. 3.** Actualized sub-graph

Surprisingly, the result of the presented approach bears some similarity to that of [18] in which a Kohonen Self-Organizing Network (SOM) is used to obtain a topological graph representation of the environment. The SOM node positions change during network convergence, but the graph itself does not, i.e. edges are not deleted. Our approach represents the environment better in the sense that unnecessary edges and nodes are removed and obstacles are represented as cycles in the graph. A further advantage is a lower computational cost; neural network approaches can take a long time to converge.

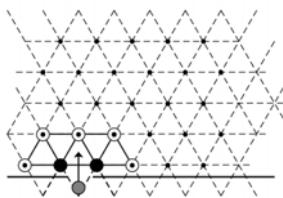
### 3.2 Guided Distribution and Positioning

In order to provide accurate robot positioning and localisation we employ the strategy that robots alternate roles between moving entities and temporary beacons.

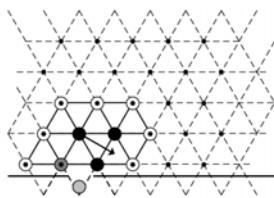
The first two robots take initial position by the site entrance. These are (initially) stationary and any two such stationary robots are referred to as *sentinels* (see Fig. 4). The distance between sentinels is predefined (the AVD radius). We put the AVD radius equal to a (at most) half of the LRF sensor range to reduce the number of unnecessary movements as well as to ensure better mutual recognition of robots. The third (anonymous) robot enters the site and moves to the apex of the equilateral triangle, the base of which is formed by the sentinels. Recognition of an anonymous robot is always carried out by two sentinels, following which the new robot is informed of its position (calculated relative to the sentinels), and instructed to manoeuvre to its new position, periodically requesting laser data from the sentinels to correct for odometry. The whole process operates over wireless TCP/IP communication.

The robots finally form a triangular cell, and their AVD are fused. The next step is similar in that one of the (old) sentinels move, initially on command of the new sentinel, and then autonomously on its own (requesting sentinel laser data when necessary as before). Decision over which sentinel will move next depends on collision information gathered by the sentinels and relies on the general underlying algorithm (see section 4.4).

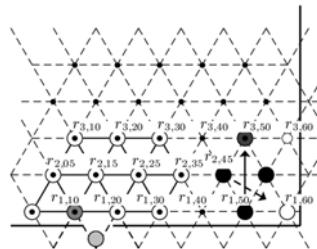
The robots now explore the environment by propagating the triangular cell formed by three robots either to the right (or left), or up (down), depending on the structure of the environment. When the possibility to move to the right is exhausted, the robots first move down if possible, otherwise the move ‘up’ and then again, depending on the sensed environment, move to the right (or to the left). The fourth robot, if available, stays as the beacon for maintaining the wireless communication. As demonstrated, three robots can move in the described manner thus ‘sweeping’ the entire site, while our set up assists localisation accuracy (see Fig. 5).



**Fig. 4.** Robot guided distribution. Initial step with resulting ITSG.



**Fig. 5.** Robot guided distribution. Moving around, with resulting ITSG.



**Fig. 6.** Robot guided distribution. Corner strategy.

The environment map is formed by the boundary nodes of the fused robot AVDs. If there is no obstacle, the map will represent a simple polygon. Each node can be addressed by the row and column number of the corresponding point in the triangular grid.

Fig. 6 depicts the strategy when robots encounter a corner. An obstruction arises when a robot laser range finder identifies a collision in a direction and at a distance not corresponding to a robot within the graph. The edge connecting the robots real node with the virtual node is then removed. Thus, that area of the graph is no longer navigable. On encountering a corner, the robot collective co-operates to propagate upwards, until they are able to move left or right again. This results in a ‘snaking’ movement of the robots dependent on the obstacles within the environment.

## 4 Algorithms and Testing on Simulation Data

The system was written in C++ on Debian Linux using the Eclipse Galileo IDE, with some later usage of Windows 7 and Microsoft Visual Studio 2010 IDE.

### 4.1 TCP/IP Communications Suite

Communications are achieved via a TCP/IP class (based on<sup>3</sup>) written specifically for the system. This is a threaded TCP/IP server/client which listens out for and establishes new client connections and listens for incoming messages, which it (i) pushes to a general message queue, or (ii) acts on directly before autonomously processing trivial requests. The latter function of the listener thread is concerned with non-intensive trivial processing such as position and laser range finder requests. This enables simple requests to be processed quietly thus leaving the main program to act upon the more processor intensive processing tasks instigated by requests at the message queue. Non-trivial messages include, for example, a ‘move robot to new position using sentinels’ sub-function which would give the non-trivial function full control. Further non-trivial processing would be placed on hold, but trivial processing is maintained.

### 4.2 Inauguration of Robots into the Team/Collective

In accordance with section 3.2, pseudo code for recognizing the identity of an anonymous robot is as follows:

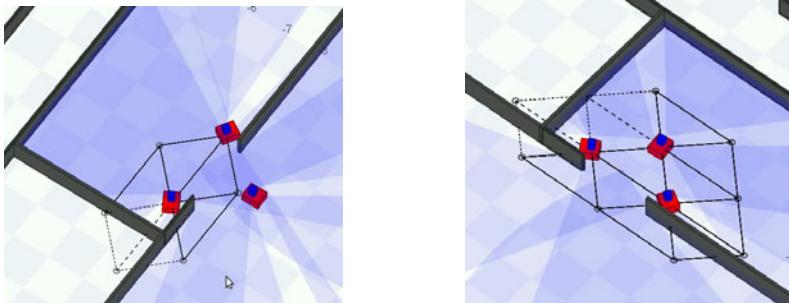
- From the higher-level algorithm (presented in 4.4), the collective recognises it must inaugurate a novel anonymous robot into the collective.
- Sentinels rotate such that their laser ranges begin at a 90 degree angle to the chord joining the two. Sentinels then expect any collision within a, say, 70 degree range of the laser starting indexes to be a robot. This may be thought of as a detection region in which anything is classed as a robot.
- Given the known sentinel positions, orientations, and laser data it is possible to compute the position of the object within the detection region.
- The distances to the detected object and the reverse angles of incidence are broadcast to all anonymous robots not in the collective.
- The anonymous robot matching these parameters is that within the detection region and the anonymous robot is identified.

---

<sup>3</sup> <http://beej.us/guide/bgnet/>

- Sentinels then transmit a message to the new robot informing it to move to its new position (the new position will be a node within the graph) whilst using laser to correct itself. At this point the new robot takes control of its own movement, requesting laser information from the sentinels to correct for odometry error.

The aforementioned steps are illustrated by two snapshots of the Player/Stage implementation, presented in Fig. 7.



**Fig. 7.** Inauguration of a new robot. Left - Initial step, Right - final step, with resulting ITSGs.

### 4.3 Robot Manoeuvre with Error Correction

Robot manoeuvre with error correction follows robot inauguration, but is actually independent of inauguration. This is because manoeuvrability is also required of robots within the graph as the team explores the graph further, hence the separate treatment in this subsection. However, the pseudo code which follows may be read as an extension to that in the previous subsection.

- On receiving the ‘manoeuvre into position’ command a robot orientates itself to the destination and travels to the new position. It should be noted that speed accelerates and decelerates according to a parabola, thus minimizing sharp changes in speed.
- Every half distance travelled, the new robot stops and requests laser data from the sentinels between which it travels to compute its own and expected positions. If the two positions do not match, then the new robot compensates by re-orientating before moving off. This step is repeated until:
- The robot reaches its destination within a pre-defined tolerance. It then transmits an arrival message back to the sentinel(s).

### 4.4 A General Algorithm

Three robots are allowed to move in one of six directions (depending on availability of movements within the graph). Directions are NORTH-EAST (NE), NORTH (N), NORTH-WEST (NW), SOUTH-WEST (SW), SOUTH (S), and SOUTH-EAST (SE).

A well-founded ordering of the directions allows for precedence of movement, e.g.  $S > SE > NE > SW > NW > N$ . Using this ordering a team of three robots can fully explore any triangular grid. The Algorithm pseudo code and a walk through now follows.

- [Start] Clear `clique_list`<sup>4</sup>
- Construct list of local cliques of size 2, `clique_list` (local to the current `_node`)
- Clear `expansion_directions_list`
- Consider expansion rules for each clique in `clique_list` (with respect to the graph) and store in `expansion_directions_list`.  
Each element of `expansion_directions_list` contains a direction, the ‘from’ node, and the ‘to’ node  $x = (dir, from, to)$
- for each  $x \in \text{expansion\_direction\_list}$ 
  - if  $x$  compromises next move then
    - \* mark `to` as visited
    - \* remove `to` from `junction_stack`
    - \* remove  $x$  from `expansion_direction_list`
  - endif
- next
- if `expansion_direction_list` is *not* empty,
  - sort `expansion_direction_list` according to well founded ordering
  - actualize the expansion direction ( $x$ )
  - `current_node = to`
  - remove `to` from `junction_stack`
- else
  - at this point a dead-end has been reached. Manoeuvre to the location at the top of the `junction_stack` (removing nodes from `junction_stack` as we go)
  - `current_node = juncture node (from top of junction_stack)`
- endif
- goto [Start]

The algorithm produces a list of cliques of size 2 surrounding `current_node`. A clique of size 2 consists of two nodes if those nodes contain a robot, so for a team of three robots there are a maximum of 3 cliques of size 2. It is then possible, along with the local graph, to determine the permitted expansion directions for each clique. There are six possible expansion directions to choose from corresponding to N, NE, etc. The first task is to deal with compromising expansion directions, i.e. dead ends. Such directions should be marked as visited and removed from the `expansion_direction_list`. The associated node is also removed from the `junction_stack` to avoid future return. If at this point the `expansion_direction_list` is not empty, then the prioritized expansion direction is actualized, otherwise we have reached a dead end and must return to the most recent junction using a shortest path algorithm.

---

<sup>4</sup> A clique is a complete subgraph.

## 5 Implementation on Real Robots

In order to test and validate our approach, miniature Khepera III [12] robots are used. The robots' functionality allows testing of algorithms in environments of varying complexity, but the robots are still compact enough to be used in the laboratory. Importantly, the robots are sufficiently powerful enough to apply to them sophisticated algorithms. Indeed, features available on the platform can match the performances of much bigger robots: velocity max of 5m/s for a robot diameter of 130mm. Khepera III has a weight of 690g with a payload of 2kg. Upgradable embedded computing power using the Korebot II system [11] can be added (800 MIPS 624MHz PXA270 processor, 128MB Ram, 32 MB Flash). The robot base includes an array of 9 Infrared Sensors for obstacle detection (0.5-25 cm) as well as 5 Ultrasonic Sensors for long range object detection (0.2-4.0 m). It also provides an optional front pair of ground Infrared Sensors for line following and table edge detection. The robot motor blocks are Swiss made quality mechanical parts, using very high quality DC motors for efficiency and accuracy. The replaceable battery pack system provides a unique solution for almost continuous experiments (1.5h), as an empty battery can be replaced in a couple of seconds. A Khepera III robot is depicted in Fig. 8. Through the KoreBot II, that uses an Arm processor, the robot is also able to host standard Compact Flash extension cards, supporting WiFi, Bluetooth, extra storage space, and many others. A Player/Stage driver for this robot has been developed previously during the GUARDIANS project [17].

### 5.1 LRF

Each robot is equipped with Hokuyo LRF to get information about the environment as well as for robot localisation and positioning. The Hokuyo company has two LRF models available for small mobile robotics (URG-04LX/URG-04LX-UG0).

It has a serial/USB connection for data transfer and even power supply by USB with the second model with an electrical consumption less than 0.5 Amperes at 5 Volts. It provides a measurement range of 0.6-4.1 [m] on a circular aperture of 240° with a measure point every 0.36° and a 0.1s refresh rate with a range error of 1% resp. 3% (see Fig. 9). Hokuyo provides a software programming guide and a library for this sensor [7]. Player/Stage provides a driver for the LRF sensor named 'urglaser'<sup>5</sup>. An electronic board was developed for integrating a LRF sensor on the Khepera III robot. Besides the communication data to the robot, it also contains a battery system for the sensor power supply with an autonomy of about 2.5 hours.

The robot and its sensors can be controlled remotely by a Player/Stage client. A Khepera III robot equipped with a Hokuyo LRF (situated on the top) is pictured in Fig. 8 and a laser scan example in Fig. 9. The ultrasonic sensors (range 0.2-4m) can be seen in the second from bottom 'ring'. The IR sensors, covering the nearest range around the robot are also displayed (next to the bottom).

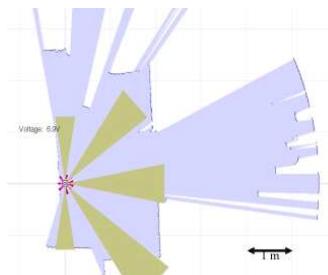
At the software level, several software modifications were needed to enable the robot to communicate with the new hardware. Firstly, the Player driver was updated to include the new sensor. An updated configuration file was also needed so that all laser

---

<sup>5</sup> [http://playerstage.sourceforge.net/doc/Player-cvs/player/group\\_driver\\_urglaser.html](http://playerstage.sourceforge.net/doc/Player-cvs/player/group_driver_urglaser.html)



**Fig. 8.** Khepera III with mounted Hokuyo's LRF



**Fig. 9.** Laser scan in Player/Stage



**Fig. 10.** Khepera III robot with extra shape for laser detection

commands could be interpreted and sent to the sensor. These modifications were important to enable full control of the new device using the Player project software only.

For direct control of the laser sensor using the KoreBot module, cross-compiling the Hokuyo libraries was necessary. Cross-compiling the device libraries and API provided by the manufacturer are carried out by using the ARM-Angstrom cross-compiler<sup>6</sup> and the bitbake software<sup>7</sup>.

## 5.2 Communication on Real Robots

In order to achieve efficient on site distribution the robots communicate by an ad-hoc wireless network. This network is built by the robots and represents a part of the self-organising GUARDIANS system. Each robot is equipped with a WiFi flash card (IEEE 802.11b/g, Ambicom WL500G-CF). A purpose built threaded TCP/IP class was developed to handle communication independent of Player/Stage (see subsection 4.1).

# 6 Experimental Set-Up

## 6.1 Robot Detection and Recognition

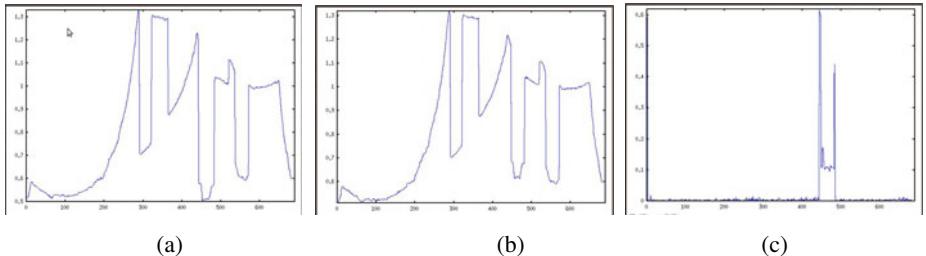
A robot can determine its position based on the positions of others using the triangulation method described earlier. However, in general, due to the small size of the LRF sensor, it is difficult to detect such a small shape. Therefore, an extra shape of a larger size is designed and fixed onto the robots as shown in Fig. 10.

Robot detection is carried out using the so-called ‘two scans’ method. A robot scans the surrounding environment twice using its LRF sensor with some time delay between scans. Because the robots have moved between the scans, the absolute difference between the scans will emphasise the location of the robots, thus leading to detection of

<sup>6</sup> [www.angstrom-distribution.org](http://www.angstrom-distribution.org)

<sup>7</sup> <http://bitbake.berlios.de/manual/>

the robots and illumination of all obstacles. The robot compares two scans, and broadcasts a message using the TCP/IP socket asking the other robots about their estimated positions. Based on this finding, each robot recognises other robots, estimates and, if necessary, corrects their positions. Figure 11 shows the laser measurements (distances taken) of the first and second scan respectively and the absolute difference between the two profiles shown in previous figures. Clearly, this difference eliminates all of the obstacles as they remain static between scans.

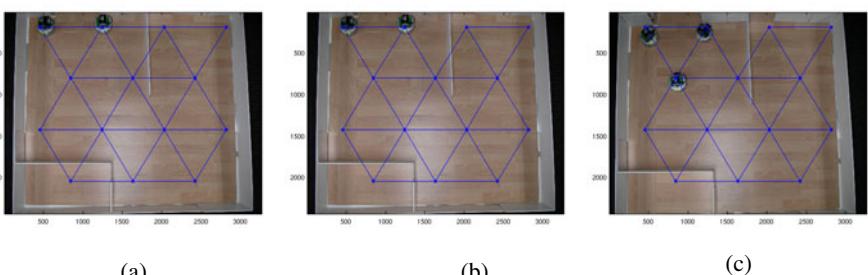


**Fig. 11.** Scan measurements (a) Initial scan graph. (b) Final scan graph. (c) Differences between two scan graphs.

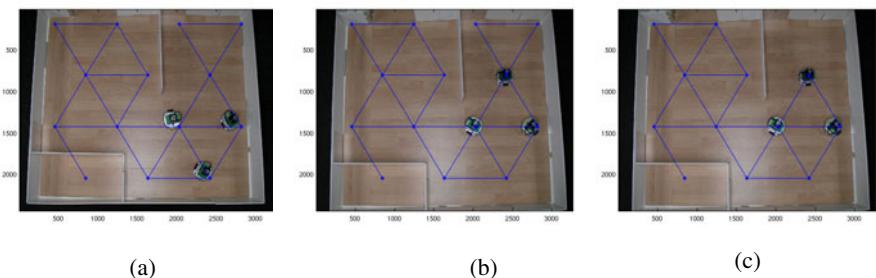
Recall that the strategy for robot movement requires that only one robot moves at a time, which simplifies the recognition process even further. However, our approach is more general, and allows recognition of several robots in the vicinity of the scanning robot. Actually, the method can also be used without communication, but communication provides robustness if there are moving obstacles as well. In the latter case not only one robot scans the environment, but the other robots as well. Based on the obtained differences in the scans and communicated approximate positions each robot can detect other robots in its AVD.

Fig. 12 and Fig. 13 depict initial and final steps of implementation on real robots.

Initially the grid covers the whole environment (arena). The final topological map represents an induced subgraph of the initial grid.



**Fig. 12.** Implementation: (a) Robots positioned as beacons. (b) Updated graph. (c) Third robot is introduced.



**Fig. 13.** (a),(b),(c) - Finals moves of the team and the resulting topological map

## 7 Conclusion

We have proposed an approach consisting of a cooperative positioning system (CPS) that accurately determines robot positions through cooperative control of individual robots in the group, construction of the topological map of the site and an ad-hoc wireless communication network. The topological map of the environment is obtained together with its *geometric realisation*, yielding in due course a topological representation and a ‘rough’ geometric representation of the environment, referred to as the basic two layers in the map building schema proposed in GUARDIANS. Further layers, such as local 2D maps, can be sufficiently easily built on top of the first two, as the positions of grid nodes are established. Skeletons can also be constructed; shortest path algorithms to return to the previously visited positions represent examples of skeletons (the schema is reported in detail in [1]).

In future we are planning to generalise guided distribution for a group of three robots to a group of  $n$  robots. Here we envisage three sub-directions: (i) development of guiding strategies of moving the group in formation, (ii) improvement of guided strategy to reduce the number of revisited nodes (i.e. dynamic modification of `expansion_directions`), (iii) splitting the group in several subgroups such that each subgroup would explore only a part of the environment in order to make more efficient exploration.

We are also planning to generalise the method to a flexible grid. Depending on the environmental data obtained from robot sensors some parts of the grid can expand and others contract to accommodate better the geometry of the site.

Currently, our method takes into consideration possible odometry errors as well as estimation errors of robot localisation by means of LRF data acquisition and their processing. We are now working on improvement of the method by introducing probability measures.

## Acknowledgment

The authors would like to thank all the GUARDIANS project partners for their contribution and cooperation especially Veysel Gazi, Ulf Witkowski, Stefan Hebrechtsmeier and Mohammed El-Habbal for their valuable comments.

## References

- [1] Alboul, L.: Conceptual Design Document on hierarchical hybrid schema for global map building and localisation. SHU (2009)
- [2] Howard, A., Mataric, M.J., Sukhatme, G.S.: Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem. In: 6th International Symposium on Distributed Autonomous Robotics Systems (DARS 2002), Fukuoka, Japan, June 25-27 (2002)
- [3] Ludwig, L., Gini, M.: Robotic Swarm Dispersion Using Wireless Intensity Signals. In: Distributed Autonomoumous Robotic Systems, vol. 7, pp. 135–144. Springer, Japan (2007)
- [4] Fox, D., Burgard, W., Kruppa, H., Thrun, S.: A Probabilistic Approach to Collaborative Multi-Robot Localization. *Autonomous Robots* 8(3), 325–344 (2000)
- [5] Fox, D., Ko, J., Konolige, K., Limketkai, B., Schulz, D., Stewart, B.: Distributed Multirobot Exploration and Mapping. *Proceedings of the IEEE* 94(7), 1325–1339 (2006)
- [6] Hokuyo laser range finder,  
[http://www.hokuyo-aut.jp/02sensor/07scanner/urg\\_04lx\\_ug01.html](http://www.hokuyo-aut.jp/02sensor/07scanner/urg_04lx_ug01.html)
- [7] Hokuyo's URG programming guide,  
[http://www.hokuyo-aut.jp/cgi-bin/urg\\_programs\\_en/](http://www.hokuyo-aut.jp/cgi-bin/urg_programs_en/)
- [8] Mataric, M.J., Sukhatme, G.S.: Localization for mobile robot teams: A distributed MLE approach. In: Experimental Robotics VIII. ser. Advanced Robotics Series, pp. 146–166 (2002)
- [9] Howard, A., Kitchen, L.: Cooperative localisation and mapping. In: International Conference on Field and Service Robotics (FSR 1999), p. 9297 (1999)
- [10] Huang, H., Beevers, K.R.: Topological map merging. *The International Journal Robotics Research* 24(8), 601–613 (2005)
- [11] K-Team, Switzerland, Korebot 2 board, <http://www.k-team.com/kteam/>
- [12] K-Team SA, Switzerland, Khepera III homepage,  
<http://www.k-team.com/kteam/>
- [13] Kurazume, K.R., Hirose, S.: An experimental study of a cooperative positioning system. *Autonomous Robots* 8(1), 4352 (2000)
- [14] Penders, J., Alboul, L., Roast, C., Cervera, E.: Robot swarming in the Guardians project. In: Proc. of ECCS 2007, vol. 6 (2007)
- [15] Pugh, J., Raemy, X., Favre, C., Falconi, R., Martinoli, A.: A Fast On-Board Relative Positioning Module for Multi-Robot Systems. *IEEE/ASME Transactions on Mechatronics, Focused Section on Mechatronics in Multi Robot Systems* (2009)
- [16] Rekleitis, I., Dudek, G., Milios, E.: Multi-robot collaboration for robust exploration. *Annals of Mathematics and Artificial Intelligence* 31, 7–40 (2001)
- [17] Tharin, J.: D4.4 Software package: Khepera III drivers for Player/Stage, User's/Programmer's Manual for the Software package. K-Team SA, Switzerland
- [18] Vlassis, N., Papakonstantinou, G., Tsanakas, P.: Robot Map Building by Kohonen's Self-Organizing Neural Networks. In: Proc. 1st Mobinet Symposium on Robotics for Health (1997)

# Analysis of Dynamic Characteristics of Single Piston Hydraulic Free Piton Engine

Haoling Ren, Haibo Xie, and Huayong Yang

State Key Laboratory of Fluid Power Transmission and Control, Zhejiang University,  
Hangzhou 310027, China  
hbzie@zju.edu.cn

**Abstract.** The operation theory of the single piston hydraulic free piston engine (SPHFPE) was introduced and was simplified as a gas spring-mass system. The mathematics model of the gas spring rate vs. displacement was set up. The equivalent spring rate and frequency was calculated. Emphasis was put on the asymmetry of the dynamic characteristics when there were damps. The larger the damps were, the more energy was needed, the longer the compression time was taken, the lower the velocity can get; otherwise, the shorter the expansion time was taken, the higher the velocity can get and the lower frequency could get. It was the damps that led to the asymmetry of the dynamic characteristics. The SPHFPE system was modeled with software package-AMESim. The movement characteristics and pressure characteristics were analyzed. The curves vs. displacement showed the operation mechanism of the SPHFPE. The simulation results followed the same law as the gas spring system and matched well with the test results of the reference.

**Keywords:** Hydraulic free piston engine (HFPE); spring-mass system; equivalent spring rate; AMESim; dynamic characteristic.

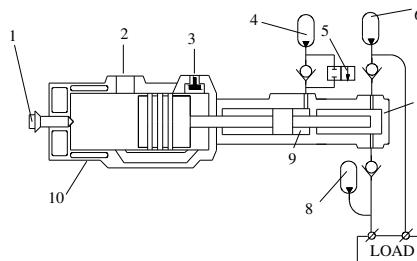
## 1 Introduction

Hydraulic free piston engine (HFPE) is a new type energy device which combines the internal combustion engine and hydraulic pump integrally to convert the rigid output into flexible output. Its free piston assembly (FPA) is constrained by no mechanical components and has an alterable dead points and compression ratio. The alterable compression ratio enables it to burn many type fuels and suits for many substitutable fuels developed to solve the energy shortage [1-7].

Single piston hydraulic free piston engine (SPHFPE) employs hydraulic oil as its working medium and has one combustion chamber and one hydraulic pump cylinder. Its compression stroke is implemented by the energy stored in the accumulator and the expansion stroke is completed by the released energy of the burning fuel and impels the pump to supply hydraulic energy to the loads. In some sense, the hydraulic pump is the load of the internal combustion engine. The SPHFPE is the simplest type of the three type HFPEs and has successful field application [1]. Through analyzing the frequency and movement characteristics of the SPHFPE, the essential experience and knowledge can be attained to develop the other two types of HFPEs and to grasp the operation mechanism of the HFPEs.

## 2 Operation Principle

Fig. 1 is the sketch of the SPHFPE developed by Innas BV, a Holland company. It includes: (1) Combustion components: injector, combustion chamber, scavenging system, intake system. (2) Hydraulic pump: high pressure accumulator and load, low pressure accumulator and supply lines, pump cylinder. (3) Compression device: compression accumulator, compression chamber, control valve. (4) Other accessorial devices.



1-injector; 2-exhaust ports; 3-intake ports; 4-compression accumulator; 5-control valve; 6-high pressure accumulator; 7-hydraulic pump; 8-low pressure accumulator; 9-compression chamber; 10-combustion chamber.

**Fig. 1.** Sketch of single piston hydraulic free piston engine [9]

The working process is: when the engine starts, the FPA positions at the most right point, called the right dead center (RDC) which is similar to the bottom dead center (BDC) in the internal combustion engine. When the signal triggers the control valve 5, the high pressure oil flows into the compression chamber 4 and pushes the FPA to left. At the same time the hydraulic pump sucks oil from the low pressure accumulator. The waste gases vent through the exhaust ports 2 from combustion chamber 10 until the exhaust ports close. The air in the combustion chamber was compressed till the FPA move to the left dead center(LDC) which is just like the top dead center (TDC) in the internal combustion engine. The diesel is injected into the chamber and pushes the FPA to right. In the same time the pump piston impels the oil into the high pressure accumulator and load and the compression piston impels the oil back into the compression accumulator for the next compression stroke. When the FPA reaches the RDC, a whole cycle is completed and the FPA stay still until the next signal to trigger the control valve.

## 3 Modeling of Key Units and Analysis of Frequency

In the SPHFPE, the combustion chamber and the compression accumulator are the key units of the SPHFPE and play an important part in the movement of the FPA. In some sense, the volumes of the two parts can be regarded as closed containers. When the gas in the chamber is compressed by the moving piston, it resists the movement of the piston; otherwise, when the gas is expanded, it pushes the piston to move. It is much more like a spring and so is called gas spring. But the gas can only be compressed, but not stretched [4]. The relationship between the force  $F$  on the gas and the gas spring rate  $K$  is,

$$K = -\frac{dF}{dy} \quad (1)$$

$$F = pA \quad (2)$$

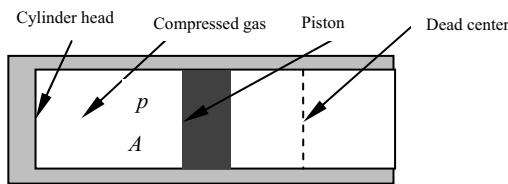
Where  $y$  is the compressed displacement of the gas spring,  $p$  is the gas pressure and  $A$  is the area of the piston.

We do some assumptions according to the operation principle of the SPHFPE.

- (1) Take the hydraulic pump as the load of the engine;
- (2) The high pressure oil in the compression accumulator is moved by the gas spring in the accumulator;
- (3) The gases in the combustion chamber can be divided into two parts: the fresh air which is taken as the ideal gas follows the adiabatic process, and the hot gases released through the fuel burning [3];
- (4) The injected fuel is to compensate the consumed energy by the damps and to impel the load. The fuel is injected in a short time and then can be considered as an instantaneous energy to compensate the energy loss.

### 3.1 Equivalent Spring Rate of Combustion Chamber

Fig.2. shows the schematic structure of the combustion chamber.



**Fig. 2.** Schematic structure of combustion chamber

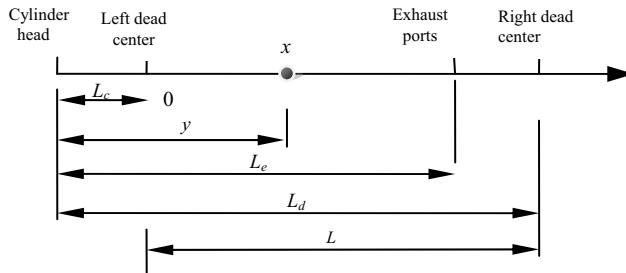
It is assumed that the gas is ideal and follows the adiabatic process while being compressed, then

$$p_0 V_0^k = p V^k \quad (3)$$

Where  $k$  is adiabatic index of the ideal gases;  $p_0$  and  $V_0$  are the pressure and volume of the gas in initial conditions, respectively;  $p$  and  $V$  are the pressure and volume of the gas in the compression process, respectively.

Fig.3. shows the structure parameters related to the spring rate,

- (1) The minimum distance between the cylinder head and the piston, i.e. the distance between the cylinder head and the LDC, is called dead length,  $L_d$ ;
- (2) The pressure in the combustion chamber is the same as that in the atmosphere when the exhaust ports open and this time the distance between the exhaust ports and cylinder head is  $L_e$ ;
- (3) The FPA decelerates to the RDC and the distance between the RDC and the cylinder head is  $L_d$ ;
- (4) The whole stroke is from the RDC to LDC, labeled as  $L$ ;
- (5) The distance between any point  $x$  in the stroke and cylinder head is  $y$ ;
- (6) The spring rate keeps zero when the exhaust ports open.



**Fig. 3.** Structure parameters of combustion chamber

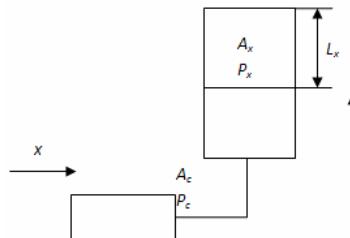
According to Eqs. (1~3) and Fig. 3, the gas spring rate of the gases in the combustion chamber can be described as,

$$K_g = \begin{cases} \frac{A_g p_0 L_e^k k}{(x + L_c)^{k+1}} & 0 \leq x \leq L_e - L_c \\ 0 & L_e - L_c \leq x \leq L \end{cases} \quad (4)$$

Where  $A_g$  is the area of the combustion chamber.

### 3.2 Equivalent Spring Rate of Compression Accumulator

Fig. 4 shows the schematic structure of the compression accumulator and compression chamber.



**Fig. 4.** Sketch of compression accumulator

According to Eqs. (1,2) and Fig. 4, then

$$F = p_x A_x \quad (5)$$

$$p_x V_x^k = p_1 V_1^k \quad (6)$$

$$V_x = A_x L_x \quad (7)$$

$$V_I = A_x L_x - A_c x \quad (8)$$

Where  $x$  is the displacement of the FPA;  $A_c$  is the area of the compression chamber;  $p_x$ ,  $V_x$  and  $L_x$  are the pressure, gas volume, gas length of the accumulator when the FPA is at the LDC, respectively;  $p_1$  and  $V_1$  are the pressure and gas volume in the accumulator when the displacement of FPA is  $x$ , respectively.

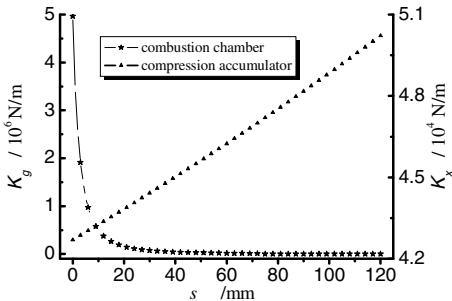
The spring rate of the compression accumulator is,

$$K_x = \frac{A_c p_x k}{L_x (1 - \frac{A_c}{A_x L_x} x)^{k+1}} \quad 0 \leq x \leq L \quad (9)$$

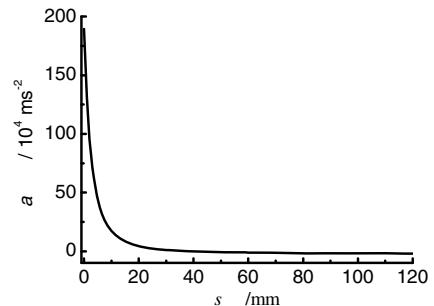
According to the design requirements, we set the parameters of the sample engine in table 1.

**Table 1.** Parameters used in the calculation

Parameter	Value	Unit
$A_g$	0.0095	$\text{mm}^2$
$A_c$	0.00095	$\text{mm}^2$
$A_x$	0.0285	$\text{mm}^2$
$p_x$	2	MPa
$r$	1.35	
$L_e$	80	mm
$L$	105	mm
$L_c$	6	mm
$L_x$	60	mm
$m$	2.6	kg



**Fig. 5.** Spring rate vs. displacement



**Fig. 6.** Acceleration vs. displacement of FPA

Fig. 5 is plotted with the above parameters and Eqs. (4,9). It shows that the gas spring rate of the combustion chamber and the compression accumulator changes with the FPA position. The gas spring rate of the combustion chamber  $K_g$  decreases dramatically when the displacement increases. And the spring rate keeps zero when the exhaust ports open. The gas spring rate of the compression accumulator  $K_x$  is linear with the displacement and increases gently with the increasing of the displacement.

The FPA acts like a mass between the two springs according to the operation principle, namely, these two springs parallel to the FPA. The piston receives the maximum forces at the LDC and has the largest acceleration up to  $1.89 \times 10^6 \text{ m/s}^2$ . But it lasts a very short distance before it reduces to  $1.93 \times 10^4 \text{ m/s}^2$ , illustrated in Fig. 6. This is a remarkable difference between the HFPE and the internal combustion engine. The piston of the internal combustion engine runs smoothly and has a low level acceleration due to

the flywheel [3,8]. There are different forces on the compression stroke and expansion stroke, thus it shows the different dynamic characteristics in these strokes. Generally, the dynamic characteristics depend on the forces acting on the FPA.

### 3.3 Equivalent Spring Rate of SPHFPE

The equivalent spring rate of the SPHFPE is calculated using an equivalent system with the same frequency and stroke. Then the spring rate of the equivalent system is the equivalent spring rate of the SPHFPE.

The two gas springs keep working in the whole process and so can be calculated with a half cycle.

The spring rate of the left combustion chamber is,

$$\int_0^L K_g(x) dx = \int_0^{L/2} K_{ge} x dx \quad (10)$$

With Eqs.(4),(10), the equivalent spring rate of the combustion chamber is,

$$K_{ge} = \frac{8A_g p_0 L_e^k}{L^2(k-1)} \left( \frac{1}{L_c^{k-1}} - \frac{1}{L_e^{k-1}} \right) + \frac{8A_g p_0 L_e^k L_c}{L^2} \left( \frac{1}{L_e^k} - \frac{1}{L_c^k} \right) \quad (11)$$

With the parameters listed above, then we can obtain,  $K_{ge}=1.32\times10^5 N/m$ .

The spring rate of the right compression accumulator is,

$$\int_0^L K_x(x) dx = \int_0^{L/2} K_{xe} x dx \quad (12)$$

With Eqs.(9&12), the equivalent spring rate of the compression accumulator is,

$$K_{xe} = \frac{8A_x p_x k}{L^2} \left[ \frac{1}{(k-1)(1-\frac{A_c L}{A_x L_x})^{k-1}} - \frac{1}{k(1-\frac{A_c L}{A_x L_x})^k} - \frac{1}{k(k-1)} \right] \quad (13)$$

With the parameters listed above, then we can obtain,  $K_{xe}=1.06\times10^5 N/m$ .

The total spring rate of the SPHFPE is,  $K_e = K_{ge} + K_{xe} = 2.38\times10^5 N/m$ .

### 3.4 Frequency Analysis of SPHFPE

For the spring-mass system, the relationship between the natural frequency and the spring rate is,  $f = \sqrt{K_e / m} / 2\pi$  and the calculated frequency of the system is 48Hz.

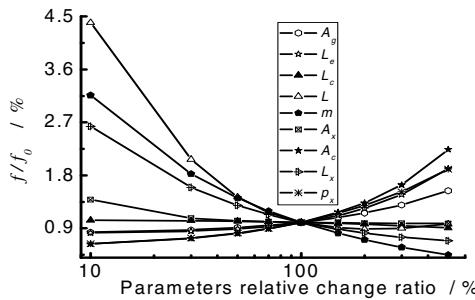
Eq. (11) and Eq. (13) show that the frequency of the SPHFPE is related to the following parameters:  $A_g$ ,  $A_x$ ,  $A_c$ ,  $L$ ,  $L_c$ ,  $L_e$ ,  $L_x$  and  $m$ .

Fig. 7 shows the frequency's relative change ratio with the parameters' relative change ratio. Table 2 shows how much the frequency in Fig. 7 depends on the various parameters.

**Table 2.** Parameters influence on frequency

$A_g$	$A_x$	$A_c$	$L$	$L_c$	$L_e$	$L_x$	$m$	$p_x$
+	-	+	--	-	+	--	--	+

Where '+' denotes the positive effect on the frequency, '-' denotes the negative effect on frequency and '--' denotes larger negative effect on frequency.



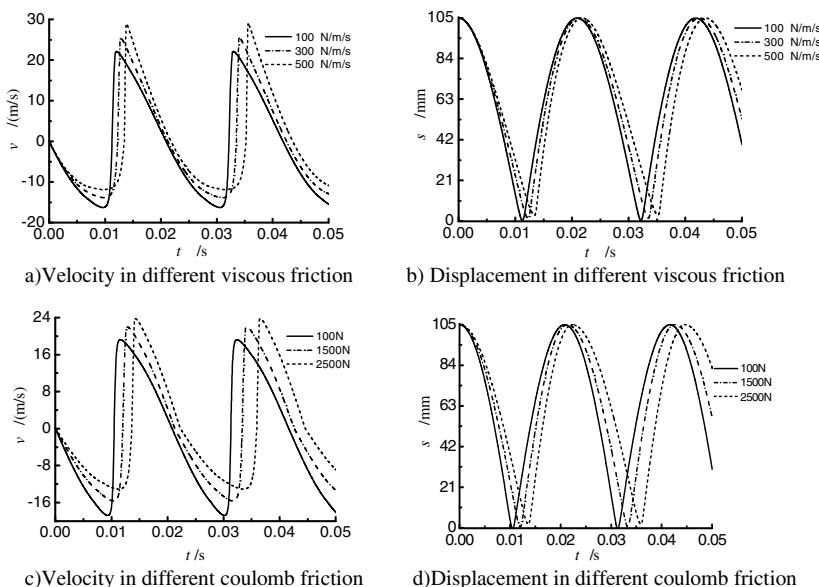
**Fig. 7.** Influence of parameters change on frequency

Improving the pressure in the compression accumulator, enlarging the area of the compression chamber, the combustion chamber; enlarging the distance between the cylinder head and the exhaust ports and lighten the mass of the FPA; shortening the stroke and the length of the accumulator can all help to improve the frequency and the output of the system.

### 3.5 Damps Effects on Frequency

The above analyses neglect all kinds of the damps. In the real systems, the damps such as the frictions and pump load consume the system energy and lead the system to stall. To keep the stable movement of the system, energy is input to the system to keep the energy balance.

Fig. 8 shows the curve of displacement and velocity in different viscous friction and coulomb friction.



**Fig. 8.** Damps effects on frequency

Fig. 8 shows that with the increasing of the friction, the frequency of the system is reduced. The displacement and velocity become more asymmetry. The velocity of the compression stroke, which is negative, is reduced, while the velocity of the expansion stroke, which is positive, is increased. The viscous friction has a greater influence on the frequency because it relates to the velocity. More information is listed in Table 3 and Table 4.

**Table 3.** Viscous friction influence on system

Coefficient of viscous friction	100 N/(m/s)	300 N/(m/s)	500 N/(m/s)
Input energy	15.6 KJ	39.2 KJ	84.2 KJ
Operation frequency	47.8 Hz	47.1 Hz	45.9 Hz
Positive velocity	22.09 m/s	25.42 m/s	28.92 m/s
Negative velocity	16.25 m/s	13.84 m/s	11.84 m/s

**Table 4.** Coulomb friction influence on system

Coulomb friction	100N	1500N	2500N
Input energy	0.93 KJ	16.8 KJ	36.8 KJ
Operation frequency	48.1 Hz	46.9 Hz	44.9 Hz
Positive velocity	19.23 m/s	21.98 m/s	23.77 m/s
Negative velocity	18.82 m/s	15.66 m/s	13.09 m/s

## 4 Dynamic Characteristic Analysis of SPHFPE

### 4.1 Simulation Model

The SPHFPE system was modeled based on Fig. 1, with the dynamical simulation software package-AMESim developed by LMS Imagine. Lab. It mainly uses the IFP engine, hydraulic, HCD, mechanical and signal libraries. The internal combustion engine is assembled using the two stroke engine components developed by LMS and IFP. It is an engine with loop scavenging systems controlled by ports, illustrated in Fig. 9. The parameters in Table 1 are used in this model.

### 4.2 Analysis of Dynamic Characteristics

In Fig. 9, the system is assumed to work at the maximum frequency, namely the FPA does not stay at RDC after the expansion stroke. Fig. 10 is the movement characteristics at continuous operation and the origin is LDC, the reference direction is from LDC to RDC which is opposite to Fig. 3. And Fig. 11 is the pressure characteristics of the SPHFPE.

Fig. 10 is the curves of displacement vs. time, velocity vs. time, acceleration vs. time, velocity vs. displacement, acceleration vs. displacement and acceleration vs. velocity, respectively.

The compression time is 15.3ms, expansion time is 10.8ms and the frequency is 38.3Hz, illustrated in Fig. 10-a. The compression stroke takes up to 58.6% of one cycle. So reducing the compression time can improve the operation frequency.

The maximum velocity of the compression stroke is 12.5m/s and that of expansion stroke is 17.7m/s, shown in Fig. 10-b. The displacement curve and velocity curve are

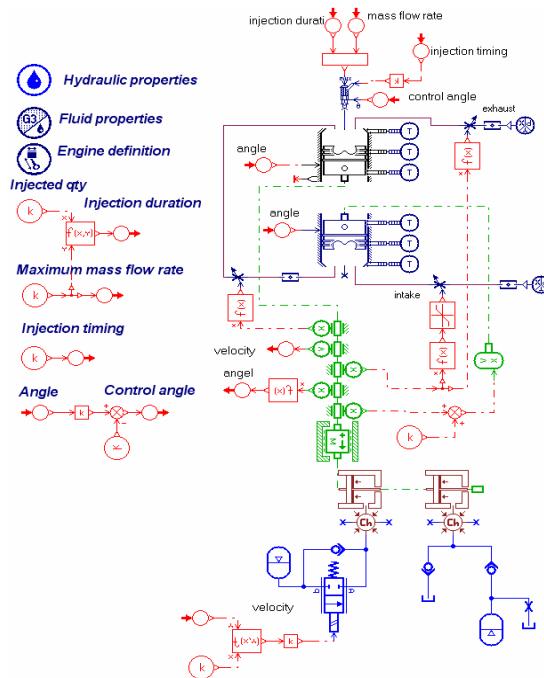


Fig. 9. Simulation model in AMESim

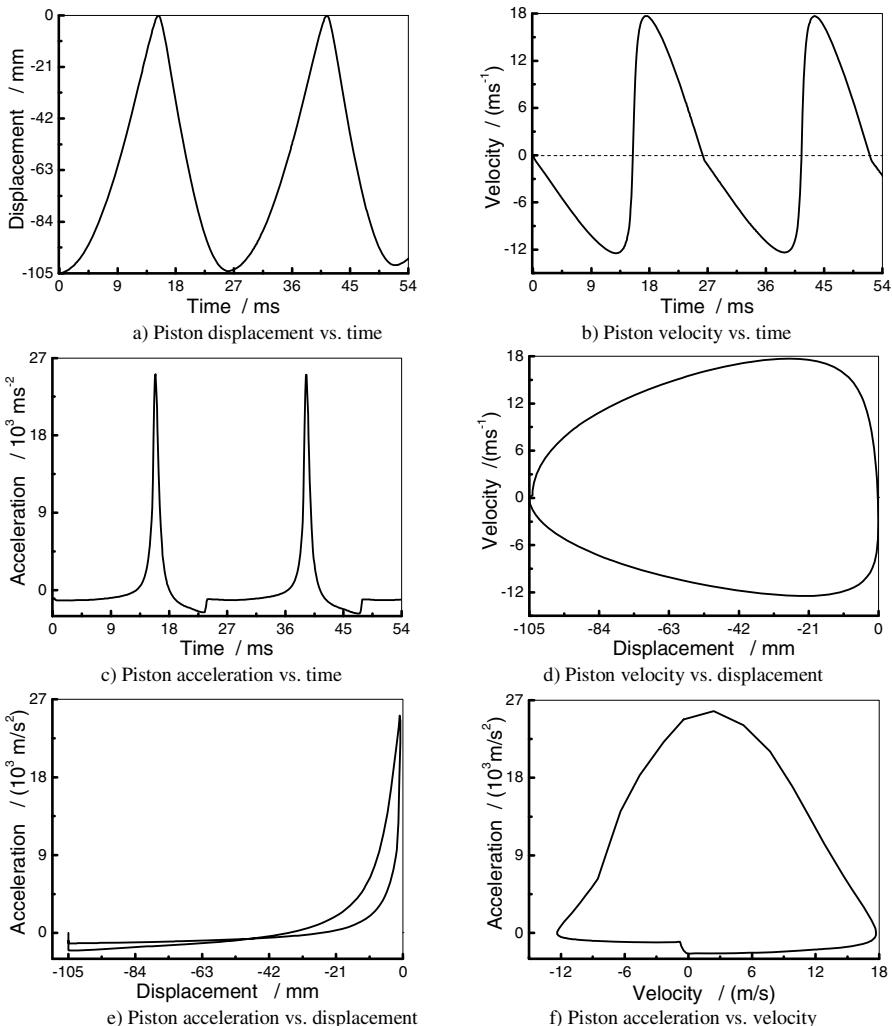
all asymmetry which meets the analysis of gas spring system in Fig. 8. In the analysis of the gas spring system, because the pump was simplified as a constant damp and the combustion process as a energy input unit, the system reaches a higher frequency than the simulation results. These two follow the same rule and the simulation results meet the test result well compared with the reference [1].

The FPA receives larger forces and produces larger acceleration near LDC because the SPHFPE has only one combustion chamber. The same conclusion can be drawn by comparing Fig. 6 with Fig. 10-c.

Fig. 10-d shows that the maximum velocity offsets from the LDC and the velocity reduces quickly to zero near LDC but slowly near RDC. And the maximum velocity in the compression stroke and the expansion stroke occur at different positions.

Shown in Fig. 10-e, the maximum acceleration occurs near LDC and reduces quickly to zero and then increases near RDC reversely. This change shows how the forces change with the displacement. And the acceleration curves of compression stroke and the expansion stroke are not superposed because the forces in the two strokes are not equal.

Fig. 10-f shows that the FPA reaches its peak value of the acceleration not at the point where the velocity is zero, but at 5m/s (i.e. in the expansion stroke). The acceleration is zero when the velocity reaches its peak value. The higher the velocity is, the smaller the acceleration can reach. The acceleration keeps higher when the velocity changes from negative to positive and keeps lower when the velocity changes reversely.



**Fig. 10.** Movement characteristics

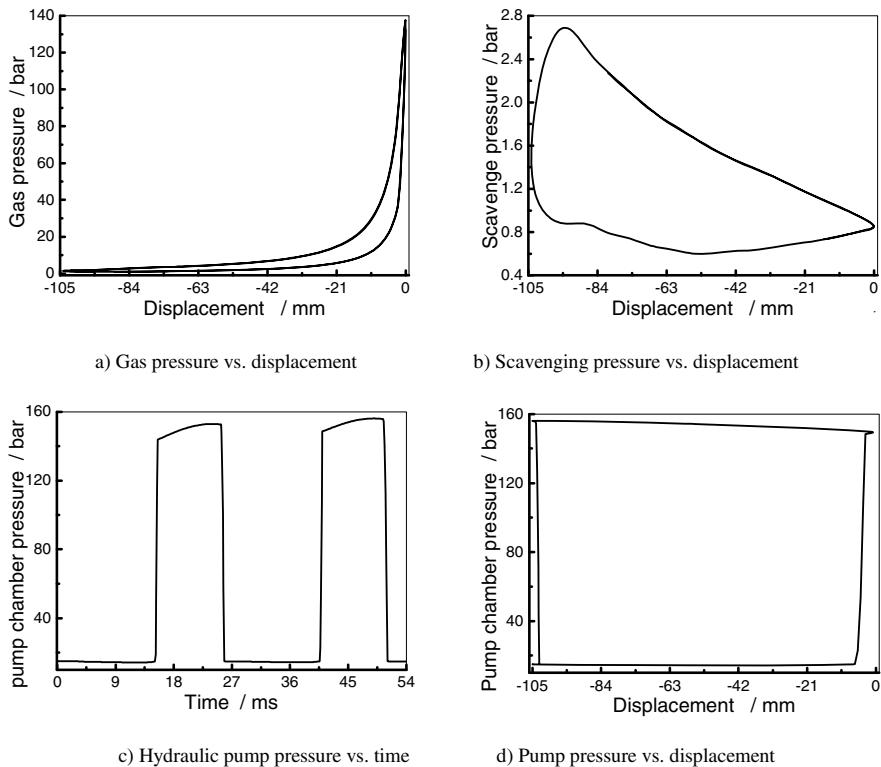
Fig. 11 is the curves of gas pressure vs. displacement, scavenging pressure vs. displacement, hydraulic pump pressure vs. time and pump pressure vs. displacement.

The gas pressure in combustion chamber increases when the FPA moves from RDC to LDC and reaches the peak value when fuels are injected into the chamber and burn. The FPA then moves to RDC and this leads the pressure reduce. The gas pressure vs. displacement is a loop because the expansion force is larger than the compression force, illustrated in Fig. 11-a.

Illustrated in Fig. 11-b, the swept volume increases when the FPA moves from RDC to LDC and makes some vacuum to intake more fresh air. The scavenging pressure increases in the expansion stroke and the pressured air is pushed to the combustion chamber to clean the burned gases.

Shown in Fig. 11-c, the pump pressure keeps low in the compression stroke because the pump sucks oil from low pressure accumulator while keeps high in the expansion stroke because the FPA impels the oil in the pump cylinder to the high pressure accumulator and load.

The pump pressure changes in the both dead centers and keeps constant in the other place, illustrated in Fig. 11-d.



**Fig. 11.** Pressure characteristics

## 5 Conclusion Remarks

This paper has simplified the SPHFPE system as a gas spring–mass system and analyzed the frequency characteristics and the influence parameters. The dynamic characteristics of the SPHFPE are studied by using the AMESim model.

- (1) The equivalent spring rates of the combustion chamber and the compression accumulator are not equal and both are the functions of displacement.
- (2) The movement characteristics become more asymmetry when damps are included. With the increasing of the damps, the time of the compression stroke increases while that of the expansion stroke decreases and the velocity of compression stroke is reduced, while that of the expansion stroke is increased. The frequency of the system is reduced.

(3)The FPA can reach a higher peak value of the acceleration than the traditional internal combustion engines near the LDC and thus the fuel injection must be controlled accurately to ensure the optimal combustion.

(4)The natural frequency and orifice area of the control valve have an effect on the pump pressure and flow rate. The higher the natural frequency is, the smaller the fluctuation of the pump pressure can obtain. The larger the orifice area is, the smaller the pressure drops, the quicker the FPA moves in the compression stroke and the higher the operation frequency can reach.

The gas spring system employed in this paper and the model built in AMESim can also be used to analyze the other free piston engines, such as free piston compressor, free piston gas generator and free piston generator. The only thing has to do is to change the operation medium.

## References

1. Achten, P.A.J., Oever, J.P.J., Potma, J., et al.: Horsepower with brains: the design of the Chiron free piston engine. SAE paper, 2000-01-2545, 3–50 (2000)
2. Hibi, A., Hu, Y.: A prime mover consists of a free piston internal combustion hydraulic power generator and a hydraulic motor. SAE paper, 930313, 1–10 (1993)
3. Tikkanen, S.: Evolution of engine-hydraulic free piston engine. Acta Polytechnica Scandinavica Mechanical Engineering Series N0.145, SPOO, Finland. 11–132 (2000)
4. Xia, B.Z.: Study on the mechanism of free piston engine hydraulic reciprocating pump. Thesis of PhD. Hangzhou: Zhejiang University, 1–31 (2003) (in Chinese)
5. Xia, B.Z., Wang, J.S., Fu, X., et al.: Development of a dual piston hydraulic free piston engine prototype and its compression ratio. Chinese journal of mechanical engineering 42(3), 117–123 (2006) (in Chinese)
6. Zhao, Z.F., Zhang, F.J., Zhao, C.L., et al.: Modeling and simulation for hydraulic free piston diesel engine. SAE paper, 2008-01-1528 (2008)
7. Zhao, Z.F., Zhang, F.J., Zhao, C.L., et al.: Research of dynamic characteristic of hydraulic free piston diesel based engine on Matlab/Simulink. Vehicle& Power Technology 110(2), 9–13 (2008) (in Chinese)
8. Tikkanen, S., Vilenius, M.: On the dynamic characteristics of the hydraulic free piston engine. In: The 2nd Tampere International Conference on Machine Automation, ICMA 1998, pp. 193–202 (1998)
9. Achten, P.A.J.: A review of free piston engine concepts. SAE paper, 941776, 1836–1847 (1994)

# Carbon-Based Nanostructured Coatings on NiTi Shape Memory Alloy for Biomedical Applications

Takanori Takeno<sup>1</sup>, Hiroyuki Shiota<sup>2</sup>, Hiroyuki Miki<sup>3</sup>,  
Toshiyuki Takagi<sup>3</sup>, and Yun Luo<sup>4</sup>

<sup>1</sup> Institute for International Advanced Interdisciplinary Research, Tohoku University Advanced Research and Education Organization, Japan

<sup>2</sup> Graduate School of Engineering, Tohoku University, Japan

<sup>3</sup> Institute of Fluid Science, Tohoku University, Japan

<sup>4</sup> State Key Laboratory Mechanical Systems and Vibration (Shanghai Jiao Tong University, China)

**Abstract.** Here, we propose carbon-based nanostructured coatings for nickel-titanium (NiTi) shape memory alloys (SMAs) for biomedical applications. NiTi SMAs are well-known biomedical materials; however, the elution of toxic Ni ions into the body has prevented SMAs from expanding their applications, particularly in the complete implantation of SMAs for human artificial muscles. One possibility for the suppression of leaching ions is to create a barrier coating, for which diamond-like carbon (DLC) may be a candidate. Due to the strong internal stress derived from the mismatch of the thermal expansion coefficient, however, the adhesive strength of DLCs is very low. In this study, we focused on the incorporation of tungsten into a DLC in order to reduce the internal stress of the coating. First, we present the definition of appropriate fabrication conditions. Uniaxial tensile tests were then performed, and we evaluated the adhesive strength of the coatings. We then determined the conditions that support the fabrication of a coating with the strongest adhesive strength. Thereafter, we deposited the coating onto the SMA with shape memory effects and conducted self-bending tests. We observed that the cracks initiated parallel to the longer direction of the specimens, most likely a result of a boundary slip between the parent phases of the SMA. The removed fraction area of the coating was estimated at less than 1 % after 105 cycles, indicating the potential usage of tungsten-containing DLC coatings for biomedical applications.

## 1 Introduction

Nickel-titanium (NiTi) shape memory alloys (SMAs) have attracted much attention because of their unique properties, such as shape memory effect (SME) due to the martensitic transformation and superelasticity (SE). The SMAs possess excellent mechanical properties like high recovery force to weight ratio and large deformation, which allows their use for various kinds of applications, including MEMS [1] and tribology [2]. In particular, SMAs have a potential for use

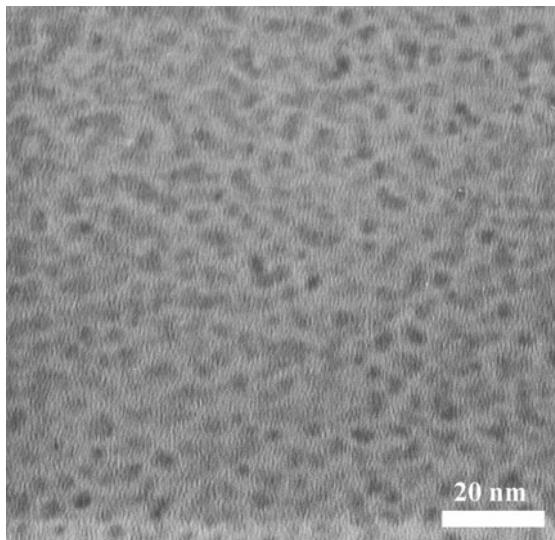
in biomedical fields since devices or actuators made of the SMAs are often very simple and powerful, leading to their successful applications [3][4]. Recent developments of the human implants utilizing the SMAs opens the possibility of using them in the medical sector [5][6]. However, there are concerns regarding elution of the toxic Ni ions that can lead to the allergenic reactions inside the body. To achieve the safe and reliable applications of the SMAs, it is very important to prevent this elution.

Diamond-like carbon (DLC) coatings have received much attention, because of their excellent properties such as high hardness, low wear, low friction coefficient, and biocompatibility [7]. Focusing on the gas barrier properties of DLC coatings [8][9], it may be possible to suppress toxic ion elution from SMAs. Kobayashi et al. demonstrated that the amount of Ni ions from SMA wires could be reduced greatly when the wires were coated with a DLC [10]. Specifically, they immersed the wires with and without DLC coatings in physiological saline for 6 months and then measured the amount of eluted ions. The released amount of ions was 933 ppb for the SMA, an amount that was greatly reduced down to 150 ppb for the DLC-coated SMA.

However, weak adhesive strength to the metallic substrate due to the high internal stress of the coating remains a concern. The weak mechanical property is likely due to a mismatch of the thermal expansion coefficient (TEC) between the DLC and the metallic substrate [11]. The TEC of SMA and DLC are typically  $\sim 10 \times 10^{-6}/\text{K}$  and  $2.3 \times 10^{-6}/\text{K}$ , respectively. The difference in the TECs can induce extremely high residual stress. This in turn can result in fragmentation of the coating, because the order of the residual stress is  $\sim \text{GPa}$  even if the DLCs are deposited onto Si substrate ( $\text{TEC} = \sim 3.2 \times 10^{-6}/\text{K}$ ).

To address this issue, two major approaches are often made. One is to fabricate the interlayer between the DLC and metallic substrate. Since the high residual stress is derived from the difference between the TECs of the DLC and the metallic substrates, the adhesive strength can be improved by adding an interlayer that has a TEC between the DLC and the metallic substrate. As such, Kobayashi et al. made a SiC layer to increase the adhesive strength [10]. Likewise, Wei et al. demonstrated an interlayer effect on the adhesive strength of the DLC coatings and concluded that a proper interlayer material should be selected for good adhesion [11]. A second approach is to incorporate an additional element. Several researchers have demonstrated a reduction of the internal stress by the incorporation of Si [12] and  $\text{SiO}_x$  [13]. Choi et al. also reported that the internal stress of Ag-doped DLCs can be reduced by more than half with an increase in the amount of an additional element [14].

Our aim followed the latter approach, because we expect a by-product of the doped-DLCs to act as a sensor function [15][16]. If the DLC coating with a sensor function is fabricated on the SMA actuator, "self-sensing actuators" can be realized in the DLC/SMA structure. Not only does strong adhesive strength result from the reduction of the internal stress, but electrical functionality is also expected in the coating. Here, we focused on the tungsten element for the inclusion and investigated tungsten-containing DLC (W-DLC) coatings onto stainless steel



**Fig. 1.** Plan-view image of the W-DLC coating. The image shows the typical structure. The black and gray part corresponds to the tungsten nanocluster and amorphous carbon host matrix.

substrates (JIS SUS440C,  $\text{TEC} = 10\text{-}12 \times 10^{-6}/\text{K}$ ). The structure of our coating is that tungsten nanoclusters are well dispersed in amorphous carbon host matrix as shown in fig. 1. This kind of structures is classified into granular one [17]. Within our previous results, the internal stress can be reduced from 3-5 GPa for the non-doped DLC coating to the minimum of 0.8 GPa for the W-DLC coating when the coatings are deposited onto Si wafer. We have reported that the improvement of the adhesive strength of the coating deposited onto the stainless steel substrates (JIS SUS440C,  $\text{TEC} = 10\text{-}12 \times 10^{-6}/\text{K}$ ). The structure of our coating consisted of tungsten nanoclusters, which were well dispersed in an amorphous carbon host matrix as shown in Fig. 1. This type of structure is classified as a granular structure [18]. As seen in our previous results, the internal stress can be reduced from 3-5 GPa for the non-doped DLC coating to the minimum of 0.8 GPa for the W-DLC coating when the coatings are deposited onto Si wafers. We have reported improvement of the adhesive strength of the coating deposited onto the stainless steel substrates (JIS SUS440C,  $\text{TEC} = 10\text{-}12 \times 10^{-6}/\text{K}$ ). The degree of the reduction depended on the concentration. We have also demonstrated that the W-DLC coating on Ni-based SMAs showed great adhesive strength over the application of 10 % uniaxial tensile loading when the coating was deposited under certain conditions that achieved the lowest internal stress [19]. These findings allowed us to undertake further steps for the development of a SMA-DLC system for biomedical actuator applications.

In this study, we present SMA actuators with highly adhesive W-DLC coatings. It is expected that the coated-SMA system can be a new type of mechanical

**Table 1.** Deposition parameters used in this study

Coating ID	WD40	WD60	WD80
Ar/CH <sub>4</sub>	10.0/3.0		
Gas pressure [Pa]		1.6	
DC power [kW]	0.04	0.06	0.08
Concentration [vol.%]	11	20	30
Thickness [μm]		~1.0	

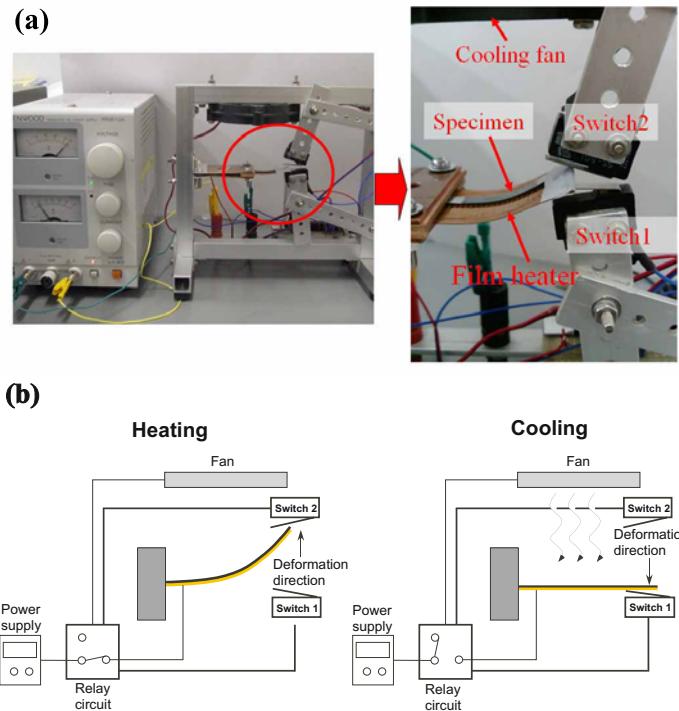
component for biomedical applications, because the system is expected to resolve the issue of the SMA, namely, the elution of toxic Ni ions. This paper is divided into two parts: the first addressing the appropriate conditions to achieve strong adhesion of W-DLC coatings to the SMAs. The best condition for the deposition was determined from the results of the tensile tests. The second part of this study addressed the fabrication of the coated-SMA system and examined the adhesive strength of the coatings. Taking into account the practical usage of the SMAs, the tests were performed by SME of the SMAs. We constructed a simple device to perform the self-bending tests of the SMAs and investigated the usability of this system for biomedical applications.

## 2 Experiment

### 2.1 Sample Preparation

A NiTi SMA plate with a composition of Ni<sub>51</sub>Ti<sub>49</sub> was used in this study. SMAs were cut into dog-bone shape with the rectangular shape of 5.0 mm × 40.0 mm × 0.5 mm, using an electric discharged machine. First, the solution heat treatment was performed at 1123 K for 20 minutes. The SMAs were maintained in a round shape with a curvature radius of 24.5 mm, and the aging treatment was performed at 673 K for 100 hours. The process allows the SMAs to have a SME; a curved shape with 1 % of bending strain in the austenitic phase and a flat shape in the martensitic phase. The start and finish temperature of the transformation were as follows [6]:  $A_s = 47^\circ\text{C}$ ,  $A'_f = 55^\circ\text{C}$ , for the transformation from the A-phase to R-phase, and  $M'_s = 49^\circ\text{C}$ ,  $M'_f = 45^\circ\text{C}$  for the reverse transformation. The SMAs were polished with grid paper and a diamond slurry to remove the thin surface oxidation layers, because the layers can weaken the adhesive strength of the coatings.

The carbon-metal composite coatings were deposited using radio-frequency plasma enhanced chemical vapor deposition (RF-PPECVD) and DC magnetron co-sputtering of the metal target. Details on the deposition system used in this study can be found in [20]. In this study, tungsten was used for the doping metal. The SMAs were firstly pumped upto  $4.0 \times 10^{-4}$  Pa for preliminary evacuation. Prior to the deposition, the sputtering target and the substrate were etched to remove the surface contamination by DC plasma and RF plasma of Ar, respectively. The etching processes were done at 1.3 Pa. Then, the methane



**Fig. 2.** Experimental apparatus; (a)photo and (b)schematic illustration

(CH<sub>4</sub>) gas was also introduced into the chamber and started the deposition. The deposition conditions, metal concentrations and thicknesses of the coatings are summarized in Table II. A sputtering power of 0.04 kW was the lower limit in our apparatus when the flow ratio (Ar/CH<sub>4</sub>) was fixed to 10.0/3.0. The metal concentration was estimated by energy dispersive X-ray spectrometer (EDS), and the concentration increased with an increase in DC input power.

## 2.2 Uniaxial Tensile Tests

Uniaxial tensile tests were performed to determine the optimum deposition conditions. The tests were conducted in ambient air condition. The tensile speed was fixed as 0.2 mm/min. The strain applied to the coatings was determined by an expression as follows:

$$\epsilon = \frac{l - l_0}{l_0} \times 100 \quad (1)$$

where,  $l_0$  and  $l$  are the initial and final displacement between the chucks, respectively. In this study, we applied the tensile strain up to 10 %.

### 2.3 Bending Tests by SME

The cyclic bending tests by SME were performed by the hand-made apparatus shown in Fig. 2. The test apparatus was consisted of a film heater, the two switches, a fan, a relay circuit and a power source. The sequence of the deformation and the recovery were as follows:

1. Turn on the heater to induce the austenitic phase for the deformation
2. Approach the switch 2 to turn off the heater and turn on the circulation fan
3. Cool down the SMA to induce the martensitic phase for the recovery
4. Approach the switch 1 to turn off the fan and turn on the heater

The heater increased the temperature up to 90 °C which is enough to induce the austenitic phase completely. It took 2 seconds and 6 seconds to finish the deformation and the recovery, respectively. The difference in duration for forward and backward in our setup was due to the different efficiencies of heating and cooling. The heating rate was fairly high compared to the cooling because the film heater is directly attached to the SMA. The number of the sequence was counted by the operating duration, and we tested it for  $10^5$  cycles. The surface of the SMA after the tests were observed by scanning electron microscope(SEM).

## 3 Results and Discussions

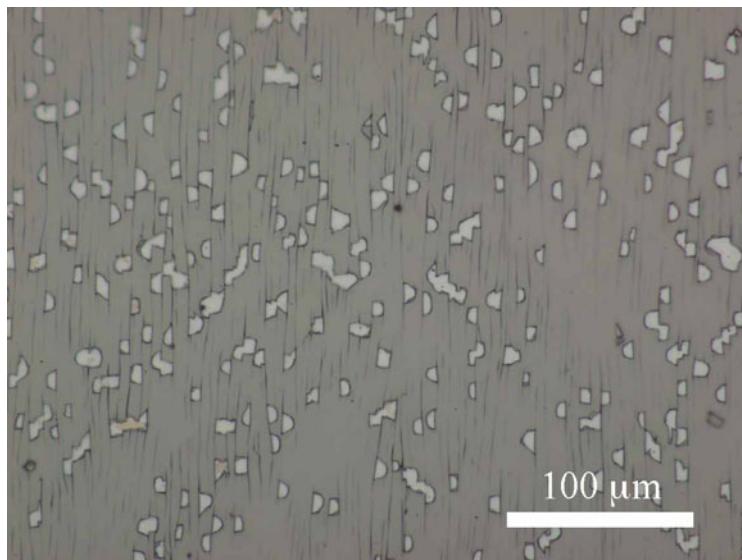
### 3.1 Determination of the Deposition Parameters by Uniaxial Tensile Tests

The typical optical image of a WD80 specimen after 10 % tensile strain is shown in Fig. 3. The tensile loading was applied in a horizontal direction to the image. Two typical fracture types of the coating can be seen in Fig. 3. One includes cracks initiating perpendicular to the tensile loading direction and the cracks located parallel to the other cracks. This type of fracture appears when the tensile or bending load is applied to the coating/substrate system [21]. The other fracture included a coating failure. One can see the white-colored area, like an "island" is apparent. According to the results of the elemental analyses performed using the energy dispersive X-ray spectrometer, it was confirmed that the coating was completely detached and the substrate surface was exposed.

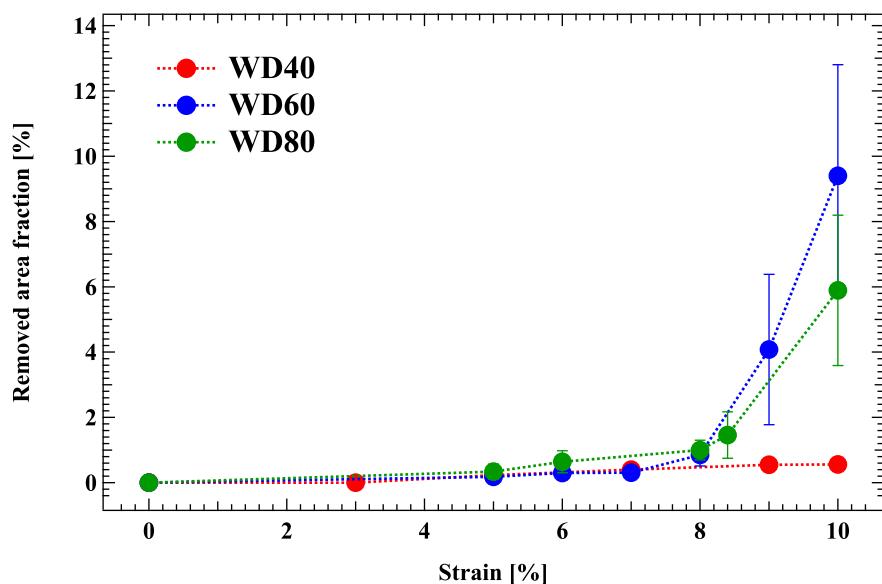
To evaluate how much the coating was detached from the substrate quantitatively, we estimated the removed area from the optical images. The optical images were firstly binarized and then the removed part and the coating were colored as white and black, respectively. The white-colored pixels were then counted, and the removed area fraction,  $RAF$ , was estimated using the expression as follows:

$$RAF = \frac{D}{D_0} \times 100 \quad (2)$$

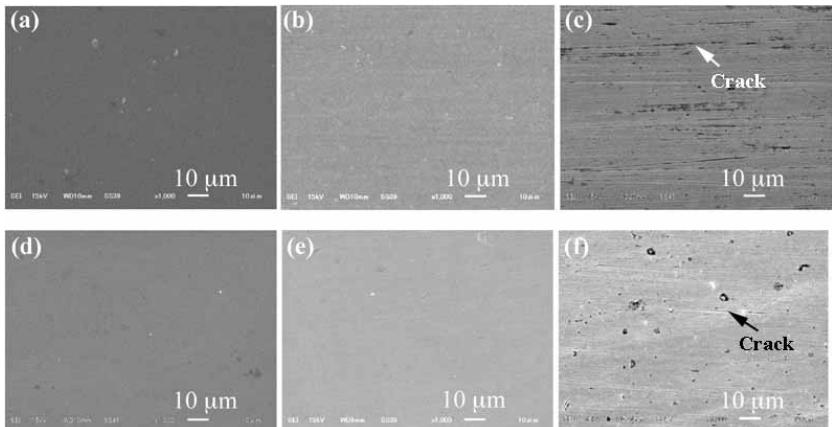
where,  $D$  and  $D_0$  represent the number of the pixels in the removed area of 300  $\mu\text{m}^2$ , respectively.



**Fig. 3.** Surface optical images after 10 % tensile strain in WD80. The tensile direction is longer direction to the optical image.



**Fig. 4.** Removed area fraction,  $RAF$ , plotted against the strain



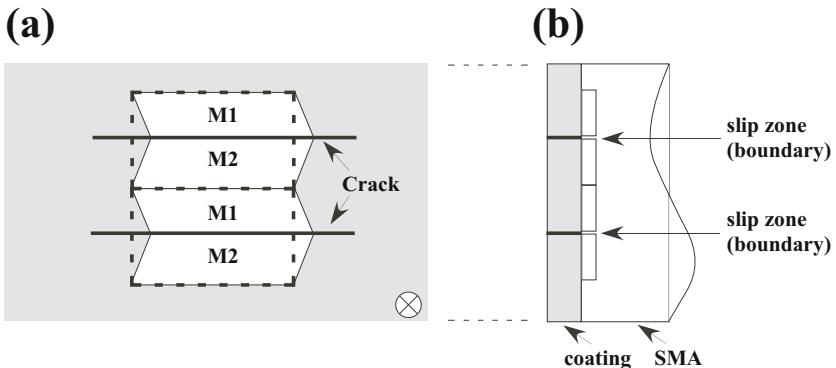
**Fig. 5.** SEM surface images after the bending tests. Images of (a)-(c) and (d)-(f) correspond to the surface of the tension side and the compression side; (a) and (d) shows the initial surface, (b) and (e) after  $5 \times 10^4$ , (c) and (f) after  $10^5$  cycles. The SMA deforms on the back side of the images for (a)-(c) and on the near side for (d)-(f).

Fig. 4 shows the plot of *RAF* against the tensile strain. The *RAF* started to increase at 7 % strain for the WD60 and 5 % strain for the WD80, then, it gradually increased with the strain. At 10 % strain, the *RAFs* of the WD60 and the WD80 were estimated as 10 % and 6 % for WD 80, respectively. However, the *RAF* of the WD40 was almost a constant value compared to the other specimens. Although the cracks in the coating appeared, no clear detachment of the coating could be seen from the optical images. Additionally, even if the *RAF* of the WD40 was estimated, its value was  $\sim 1\%$ . It is not clear why such a strong adhesive strength was achieved in WD40 is not clear up to now. However, as far as our investigations of the internal stress of the W-DLCs on Si substrates, the stress was minimized when the coating was deposited under the same parameters as the WD40. Since an ideal adhesive strength of the coating was achieved for the WD40, the deposition parameter for the WD40 was thereafter adopted.

### 3.2 Self-bending Tests by SME

The self-bending test by SME was conducted to evaluate the actuation of SMAs for a long-term operation. The test took almost 9 days to complete the  $10^5$  bending cycles. During the operation, we did not see any malfunctions such as the depression of the bending force that originated the SME with the fatigue or visible failure of the coating.

The surface SEM images after the cyclic bending tests are shown in Fig. 5. As seen in Figs.(a) and (d), the initial surface were very smooth. The maximum height of the coatings was measured as 23 nm, and the surface height without a coating was 21 nm. Since the dispersion of these values are  $\sim 5$  nm, it was understood that the surface did not become rough by the coating. The surface



**Fig. 6.** Schematic illustration of the coating with cracks after the self-bending tests by SME; (a) top view and (b) cross-sectional view. The SMA deforms on the near side of the sheet. The M1 and M2 represent the different martensitic phase. The cracks initiate along the boundary of the phases.

was kept smooth under the bending motion, and the cracks started to initiate when the number of cycles reached  $5 \times 10^4$  as shown in Figs. 5(b) and (d). By further increasing the cycles, the number of the cracks increased, and the distance between the cracking became narrower.

Notably, the cracks were initiated in the longer direction. In general, if the brittle coating was forced to be expanded, cracks were initiated perpendicular to the longer direction [21]. When the tensile strain reaches a certain value, or in other words when the stress reaches the critical fracture stress, the first crack in the coating appears, and the stress of the coating is relieved by initiating the cracks [22]. If further strain is applied to the coating and the stress reaches a critical fracture stress again, other cracks are initiated. In the tensile loading, the number of the cracks increases in this process and is finally saturated. However, in this study, no perpendicular cracks were observed as shown in Figs. 5(c) and (f). A possible explanation is that the stress of the coating under the maximum bending strain did not exceed the critical fracture stress.

The occurrence of the cracks parallel to the longer direction may be due to a boundary slip phenomena between the small phase domains in SMA during the SME [23]. A schematic illustration of the coating after the self-bending tests is shown in Fig. 6. The SME is the recovery effect with the martensitic transformation from the martensitic phase to the austenitic phase and vice versa. As a principal, the transformation without any dislocations is derived from the existence of the variant phases in SMAs. When there is deformation, a different phase as shown schematically in Fig. 6 is introduced in order to compensate the strain. In case of the inverse-transformation, the variant phase completely disappears and the shape is recovered. Despite the complete recovery of its shape, the boundary of the phases moves with an increase in the number of phase transformations [24][25]. The movement of the boundary introduces the shear stress in the coatings. Since the degree of the movement of the boundary becomes larger

as there is an increase in the number of phase transformation, the stress reaches the fracture strength of the coating and cracks are initiated. In the case of the SMAs used in this study, the boundary was parallel to the longer direction of the substrate, and the cracks were initiated along the boundary.

The RAF after the self-bending test was estimated as  $RAF < 1\%$ . Since the island-type delamination as shown in Fig. 3 was not observed in the coating, the value of  $RAF$  corresponded only to the cracks. The small amount of  $RAF$  led to the strong adhesive strength of the coating. Overall, the WD40 exhibited a strong adhesive strength to the substrate, and may have potential usage for a DLC/SMA actuator system.

## 4 Summary

In this study, SMAs with a highly adhesive W-DLC coating were presented with the aim of preventing the toxic Ni ions from eluting. The coatings were deposited by a PECVD technique and DC magnetron co-sputtering of a tungsten metal target. Three types of the W-DLC coatings were investigated to determine the appropriate deposition conditions by uniaxial tensile tests: the WD40 showed the best performance of the adhesive strength to the SMAs. The self-bending tests with SME were then performed in up to  $10^5$  cycles. The optical images indicated that the RAF of the coating was less than 1 %, and no clear detachment of the coating was observed. The cracks after the self-bending tests initiated along the longer direction to the substrate, which can be distinguished from the conventional crack direction when tensile loading is applied. The occurrence of the longer cracking may be due to the slip of the boundary between the parent phases in the SMA during the cyclic bending accompanied by the SME.

## Acknowledgements

This work was partially supported by a Grant-in-Aid for Scientific Research (B) (20360380) from the Japan Society for the Promotion of Science (JSPS), the Global COE program, "World Center of Education and Research for Trans-disciplinary Flow Dynamics" by the Ministry of Education, Culture, Sports, Science and Technology, Japan, and by the State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, China. We are thankful to Mr. Takeshi Sato in the Institute of Fluid Science, Tohoku University, Japan, for his technical assistance.

## References

1. Kahn, H., Huff, M.A., Heuer, A.H.: The TiNi shape-memory alloy and its applications for MEMS. *Journal of Micromechanics and Microengineering* 8, 213–221 (1998)
2. Li, D.Y.: Exploration of TiNi shape memory alloy for potential application in a new area: Tribological engineering. *Smart Materials and Structures* 9(5), 717–726 (2000)

3. Van Humbeeck, J.: Non-medical applications of shape memory alloys. *Materials Science and Engineering A* 273-275, 134–148 (1999)
4. Fu, Y., Du, H., Huang, W., Zhang, S., Hu, M.: TiNi-based thin films in MEMS applications: A review. *Sensors and Actuators, A: Physical* 112(2-3), 395–408 (2004)
5. Filip, P., Lausmaa, J., Musialek, J., Mazanec, K.: Structure and surface of TiNi human implants. *Biomaterials* 22(15), 2131–2138 (2001)
6. Luo, Y., Okuyama, T., Takagi, T., Kamiyama, T., Nishi, K., Yambe, T.: Thermal control of shape memory alloy artificial anal sphincters for complete implantation. *Smart Materials and Structures* 14(15), 29–35 (2005)
7. Robertson, J.: Diamond-like amorphous carbon. *Material Science and Engineering R* 37, 129–281 (2002)
8. Shirakura, A., Nakaya, M., Koga, Y., Kodama, H., Hasebe, T., Suzuki, T.: Diamond-like carbon films for pet bottles and medical applications. *Thin Solid Films* 494(1-2), 84–91 (2006)
9. Abbas, G.A., Roy, S.S., Papakonstantinou, P., McLaughlin, J.A.: Structural investigation and gas barrier performance of diamond-like carbon based films on polymer substrates. *Carbon* 43(2), 303–309 (2005)
10. Kobayashi, S., Ohgoe, Y., Ozeki, K., Sato, K., Sumiya, T., Hirakuri, K., Aoki, H.: Diamond-like carbon coatings on orthodontic archwires. *Diamond and Related Materials* 14, 1094–1097 (2004)
11. Wei, C., Yen, J.Y.: Effect of film thickness and interlayer on the adhesion strength of diamond like carbon films on different substrates. *Diamond and Related Materials* 16(4-7), 1325–1330 (2007)
12. Ban, M., Hasegawa, T.: Internal stress reduction by incorporation of silicon in diamond-like carbon films. *Surface and Coatings Technology* 162(1), 1–5 (2003)
13. Randeniya, L., Bendavid, A., Martin, P., Amin, M., Preston, E., Ismail, F.M., Coe, S.: Incorporation of Si and SiO<sub>x</sub> into diamond-like carbon films: Impact on surface properties and osteoblast adhesion. *Acta Biomaterialia* 5(5), 1791–1797 (2009)
14. Choi, H., Choi, J.H., Lee, K.R., Ahn, J.P., Oh, K.: Structure and mechanical properties of Ag-incorporated DLC films prepared by a hybrid ion beam deposition system. *Thin Solid Films* 516(2-4), 248–251 (2007)
15. Takagi, T., Takeno, T., Miki, H.: Metal-containing diamond-like carbon coating as a smart sensor. *Materials Science Forum* 638–642, 2103–2108 (2010)
16. Takeno, T., Takagi, T., Bozhko, A., Shupegin, M., Sato, T.: Metal-containing diamond-like nanocomposite thin film for advanced temperature sensors. *Materials Science Forum* 475–479(III), 2079–2082 (2005)
17. Takeno, T., Miki, H., Takagi, T., Onodera, H.: Electrically conductive properties of tungsten-containing diamond-like carbon films. *Diamond and Related Materials* 15(11-12), 1902–1905 (2006)
18. Takeno, T., Sugawara, T., Miki, H., Takagi, T.: Deposition of DLC film with adhesive W-DLC layer on stainless steel and its tribological properties. *Diamond and Related Materials* 18(5-8), 1023–1027 (2009)
19. Takeno, T., Shiota, H., Sugawara, T., Miki, H., Takagi, T.: Highly adherent tungsten-containing diamond-like carbon (W-DLC) coating on a NiTi shape memory alloy under 10% tensile strain. *Diamond and Related Materials* 18(2-3), 403–406 (2009)
20. Takeno, T., Miki, H., Sugawara, T., Hoshi, Y., Takagi, T.: A DLC/W-DLC multilayered structure for strain sensing applications. *Diamond and Related Materials* 17(4-5), 713–716 (2008)
21. Agrawal, D.C., Raj, R.: Measurement of the ultimate shear strength of a metal-ceramic interface. *Acta Metallurgica* 37(4), 1265–1270 (1989)

22. Jeong, J.H., Kwon, D.: Evaluation of the adhesion strength in DLC film-coated systems using the film-cracking technique. *Journal of Adhesion Science and Technology* 12, 29–46 (1998)
23. Liu, Y., Liu, Y., Van Humbeeck, J.: Two-way shape memory effect developed by martensite deformation in NiTi. *Acta Materialia* 47(1), 199–209 (1998)
24. Liu, Y., Xie, Z., Van Humbeeck, J., Delaey, L.: Asymmetry of stress-strain curves under tension and compression for NiTi shape memory alloys. *Acta Materialia* 46(12), 4325–4338 (1998)
25. Xie, Z., Liu, Y., Van Humbeeck, J.: Microstructure of NiTi shape memory alloy due to tension-compression cyclic deformation. *Acta Materialia* 46(6), 1989–2000 (1998)

# Closed-Loop Stiffness Modeling and Stiffness Performance Analysis for Multi-axis Process System

Yan Rong<sup>1</sup>, Peng Fangyu<sup>2,\*</sup>, Chen Wei<sup>1</sup>, Yang Jianzhong<sup>1</sup>, and Linsen<sup>1</sup>

<sup>1</sup> National NC System Engineering Research Center, China

cgxiong@sina.com, owenchenwei@sina.com, afxyang@yahoo.com.cn,  
eversun1@163.com, zwm8917@263.net

<sup>2</sup> State Key Laboratory of Digital Manufacturing Equipment and Technology  
Huazhong University of Science and Technology, Wuhan 430074, P. R. C.  
zwm8917@263.net

**Abstract.** General stiffness performance of manufacturing process system directly affects machining accuracy and stability. When machining complex surface, general stiffness performance changes with tool posture that is determined by machining position and feed direction. In order to analyze general stiffness performance including machine tool, cutter and workpiece, a closed loop stiffness model is established by Jacobi matrix and finite element method. It is a general stiffness model for multi-axis process system, which can be used to evaluate general stiffness performance in workspace. The distribution law of stiffness performance can be evaluated by isolines map of stiffness index plotted in workpiece coordinate system. The result helps engineers plan tool path based on both geometric and physics constraints.

**Keywords:** Multi-axis, general stiffness, complex surface.

## 1 Introduction

With the development of national defense, transportation and energy etc., more requirement is brought forward for key parts of fundamental equipments such as ship propeller, turbine blade and air compressor impeller. Since it's difficult to machine those parts with complex surface by 3 axes machine tools, 4 or 5 axes machine tools have to be applied in machining industry[1].

Multi-axis machine tools have more movement axes and driving part. Increasing driving parts directly results in weakening stiffness performance of the whole process system. Besides the poor rigidity of movable joint surfaces, cutter and workpiece also influence general stiffness performance if typical thin-wall structure like impellers is processed by a flexible cutter. To make the rigidity suitable for cutting process, it's essential to analyze the general stiffness performance which consists of machine tool, cutter and workpiece and to help engineers plan tool path. But most researches plan tool path only based on geometry constraints[2, 3]. During machining process, even if feedrate, vertical depth of cut and crossfeed is uniform, general stiffness performance

---

\* Corresponding author.

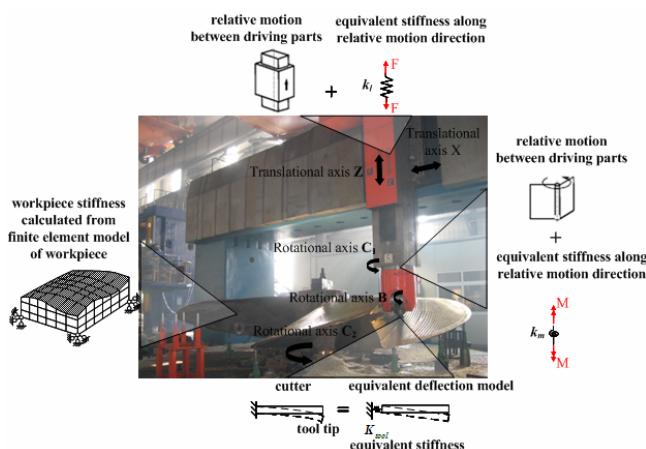
changes along with tool posture. Machining efficiency needs to be improved and machining chatter needs to be restrained by controlling tool posture. Hence, engineers can plan tool path based on both geometric and physics constraints.

General stiffness performance is characterized by the anti-deformation capacity of cutter relative to workpiece. Budak [4], Lacalle[5, 6], Philippe[7] and G.M. Kim[8] proposed a model by analytical method where only the influence of poor rigidity of cutter is concerned. Ratchev[9, 10] applied finite element method to predict general stiffness performance, where only poor rigidity of workpiece is taken into account. Wan[11] used analytical method and finite element method synthetically concerning only cutter and workpiece. In these various stiffness models, no one can reflect the distribution law of general stiffness performance consisting of machine tool, cutter and workpiece[12]. Moreover, no one is applicable for arbitrary process system.

Based on Jacobi matrix and finite element method in this paper, a general closed-loop stiffness model for multi-axis process system is established. The stiffness model involves the majority factors, including joint surface between movement axes, cutter and workpiece. Taking large scale propeller machined by standard mill head of 5 axis machine tool for an example, the model is established and the distribution law is evaluated by isoclines map of stiffness index plotted in 2D space of workpiece coordinate system, which can be used by engineers to plan tool path.

## 2 Closed-Loop Stiffness Model of Multi-axis Process System

When machining complex surface, cutting force is applied on the cutter, therefore various deformation happens in machine tool, cutter and workpiece. In the general stiffness model, the deformation of machine tool, cutter and workpiece would be considered, all of which deform elastically with the assumption of small deformations. As shown in figure 1, the model reflects the anti-deformation capacity of cutter relative to workpiece in the closed-loop which is shaped from machine to cutter tip and workpiece, called as “close-loop general stiffness model”.



**Fig. 1.** Stiffness model of multi-axis machine tool

The relationship of the displacement  $\delta v$  of cutter tip relative to workpiece and cutting force  $F$  can be expressed in the following formula:

$$F = K \delta v . \quad (1)$$

herein,  $K$  is 6×6 order general stiffness matrix.

General stiffness matrix  $K$  of the multi-axis series machine tool is:

$$K = (K_j^{-1} + K_t^{-1} + K_p^{-1})^{-1} . \quad (2)$$

herein,  $K_j$ ,  $K_t$  and  $K_p$  is stiffness of machine tool (only poor stiffness of the driving parts of movement axis is considered), cutter and workpiece respectively.

Stiffness matrix of  $n$  driving parts can be obtained by Jacobi matrix method[13]:

$$K_j = (J_j K_{joint}^{-1} (J_j)^T)^{-1} . \quad (3)$$

herein,  $J_j$  is  $6 \times n$  order Jacobi matrix which reflects the differential displacement mapping from joint space to workpiece coordinate system(CSW), which can be derived by multi-body kinematics model.  $K_{joint}$  is  $n \times n$  order diagonal matrix consisting of

transmission stiffness of  $n$  driving parts, that is:  $K_{joint} = \begin{bmatrix} k_1 & & \cdots & 0 \\ & k_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & k_n \end{bmatrix}.$

As the same way, stiffness matrix of cutter can be expressed as:

$$K_t = (J_t (K_{tool})^{-1} (J_t)^T)^{-1} . \quad (4)$$

herein,  $J_t$  is  $6 \times 6$  order Jacobi matrix, which reflects the differential displacement mapping from cutter coordinate system to CSW.  $J_t$  can be expressed as

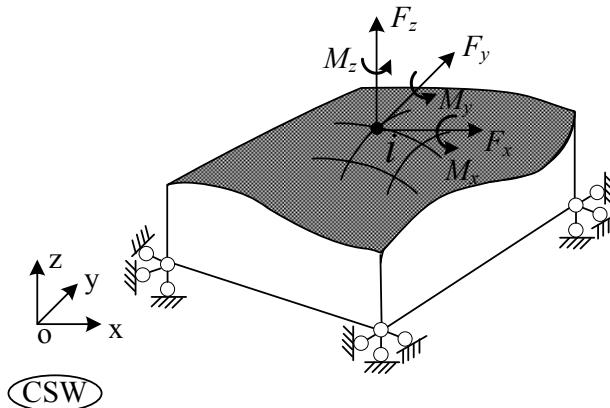
$$J_t = \begin{bmatrix} R_t & | & \mathbf{0} \\ \mathbf{0} & | & R_t \end{bmatrix}_{6 \times 6} \text{ where } R_t \text{ is rotation transformation matrix from CST to CSW.}$$

$K_{tool}$  is  $6 \times 6$  order stiffness matrix of cutter relative to cutter coordinate system CST, which reflects the rigidity of cutter.  $K_{tool}$  can be expressed as:

$$\mathbf{K}_{tool} = \begin{bmatrix} k_{xx} & 0 & 0 & 0 & k_{x\theta_y} & 0 \\ 0 & k_{yy} & 0 & -k_{y\theta_x} & 0 & 0 \\ 0 & 0 & k_{zz} & 0 & 0 & 0 \\ 0 & -k_{y\theta_x} & 0 & k_{\theta_{xx}} & 0 & 0 \\ k_{x\theta_y} & 0 & 0 & 0 & k_{\theta_{yy}} & 0 \\ 0 & 0 & 0 & 0 & 0 & k_{\theta_{zz}} \end{bmatrix}.$$

herein,  $k_{ij}$  of matrix  $\mathbf{K}_{tool}$  is the stiffness along direction  $i$  by the unit displacement along direction  $j$  on tool tip.

Because of complicated structure and complex surface of workpiece, stiffness performance varies at different positions. Finite element method is used for stiffness matrix  $\mathbf{K}_p$  of workpiece. Stiffness coefficient is related to the position and direction of applied external force. So before setting up  $\mathbf{K}_p$ , special positions need to be chosen as control points, of which flexibility coefficient would be calculated. Then, these flexibility coefficients are assembled into flexibility matrix  $\mathbf{K}_p^{-1}$ . The calculated model of the flexibility coefficient of control point  $i$  is shown as figure 2. In the model, boundary condition is set according to actual fixture condition.



**Fig. 2.** Model of the flexibility coefficient of control point  $i$

During calculating process, number unit force  $F_x$ 、 $F_y$ 、 $F_z$ 、 $M_x$ 、 $M_y$ 、 $M_z$  in the order of 1 to 6, which will be placed on control point  $i$  for flexibility coefficients. Combining these coefficients into flexibility matrix  $\mathbf{S}_p$  of control point  $i$ , which can be expressed as:

$$\mathbf{S}_p = \begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{21} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} & s_{36} \\ s_{41} & s_{42} & s_{43} & s_{44} & s_{45} & s_{46} \\ s_{51} & s_{52} & s_{53} & s_{54} & s_{55} & s_{56} \\ s_{61} & s_{62} & s_{63} & s_{64} & s_{65} & s_{66} \end{bmatrix}. \quad (5)$$

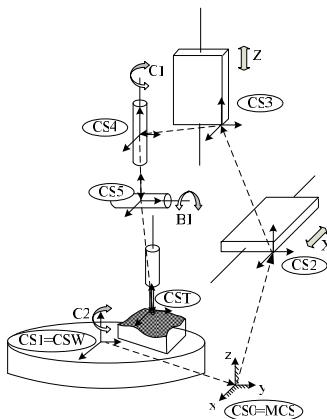
herein  $\mathbf{S}_p$  is symmetrical, where 21 coefficients needs to be determined.  $s_{jk}$  means displacement along direction  $j$  by the unit force along direction  $k$ .

According to  $\mathbf{K}_p = (\mathbf{S}_p)^{-1}$ ,  $\mathbf{K}_p$  could be directly evaluated.

Based on formula (2)~(5), general stiffness model could be obtained.

### 3 Modeling Closed-Loop Stiffness Model of 5 Axis Turn-Milling Machine Tool

5 axis turn-milling machine tool provides function of vertical lathe, drill, mill, bore and tap etc. The machine tool is double column structured, including column, beam, table base, tool holder and standart mill head. Movement axes are translational axes X and Z, rotational axes C1, B1 and C2. The kinemic chain is: C2 → X → Z → C1 → B → cutter, as shown in Figure 3.



**Fig. 3.** Kinematics model of 5-axis turn-milling machine tool

When machining different positions on the surface of the propeller, general stiffness index is different. Not only stiffness cahracteristics of machine and cutter but also that of workpiece varies. Based on the geometry of propeller, general stiffness characteristic is analyzed. Because working medium flows along the helix of propeller, feed direction

has to be coincident with the direction of helix. The lead angle of cutter along the feed direction is  $\alpha$ , which is set to get rid of overcut.

Close-loop general stiffness matrix of 5-axis turn-milling machine tool derived from the formulation of general stiffness model, namely formula (2).

$$\mathbf{K} = \left( \left( \mathbf{J}_j (\mathbf{K}_{joint})^{-1} (\mathbf{J}_j)^T \right)^{-1} + \left( \mathbf{J}_t (\mathbf{K}_{tool})^{-1} (\mathbf{J}_t)^T \right)^{-1} + \mathbf{K}_p^{-1} \right)^{-1}. \quad (6)$$

The cutter tip coincides with the origin of workpiece coordinate system (CSW) originally. Kinematics model is got by D-H method and Jacobi matrix  $\mathbf{J}_j$  is:

$$\mathbf{J}_j = \begin{bmatrix} \cos C_2 & 0 & l_t \sin(C_1 + C_2) \sin B & -l_t \cos(C_1 + C_2) \cos B & l_t \sin(C_1 + C_2) \sin B - X \sin C_2 \\ \sin C_2 & 0 & -l_t \cos(C_1 + C_2) \sin B & -l_t \sin(C_1 + C_2) \cos B & -l_t \cos(C_1 + C_2) \sin B + X \cos C_2 \\ 0 & 1 & 0 & l_t \sin B & 0 \\ 0 & 0 & 0 & -\sin(C_1 + C_2) & 0 \\ 0 & 0 & 0 & \cos(C_1 + C_2) & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}. \quad (7)$$

herein,  $\mathbf{J}_j$  is  $6 \times 5$  order Jacobi matrix.  $C_2$ ,  $X$ ,  $Z$ ,  $C_1$  and  $B$  are motion parameters of machine axes. All of them can be inversely calculated by postprocess algorithm.  $l_t$  is the distance between gyration centre of axis  $B$  and the tool tip  $t$ .

$\mathbf{J}_t$  is also derived from kinematics model as follows:

$$\mathbf{J}_t = \begin{bmatrix} \cos(C_1 + C_2) \cos B & -\sin(C_1 + C_2) & \cos(C_1 + C_2) \sin B & 0 & 0 & 0 \\ \sin(C_1 + C_2) \cos B & \cos(C_1 + C_2) & \sin(C_1 + C_2) \sin B & 0 & 0 & 0 \\ -\sin B & 0 & \cos B & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos(C_1 + C_2) \cos B & -\sin(C_1 + C_2) & \cos(C_1 + C_2) \sin B \\ 0 & 0 & 0 & \sin(C_1 + C_2) \cos B & \cos(C_1 + C_2) & \sin(C_1 + C_2) \sin B \\ 0 & 0 & 0 & -\sin B & 0 & \cos B \end{bmatrix} \quad (8)$$

Referring to design data of this machine tool, transmission stiffness of each machine axis is calculated to form diagonal stiffness matrix  $\mathbf{K}_{joint}$  of machine axes as follows:

$$\mathbf{K}_{joint} = \begin{bmatrix} 1.1100 & 0 & 0 & 0 & 0 \\ 0 & 1.0186 & 0 & 0 & 0 \\ 0 & 0 & 0.1560 & 0 & 0 \\ 0 & 0 & 0 & 0.0785 & 0 \\ 0 & 0 & 0 & 0 & 0.1560 \end{bmatrix} \times 10^9. \quad (9)$$

herein the diagonal elements are respectively axial stiffness (N/m) of translational axis  $X$  and  $Z$ , and rotation stiffness (N • m/rad) of rotational axis  $C_1$ ,  $B$  and  $C_2$ .

In this paper, disc-type cutter is adopted to machine the propeller. Stiffness matrix of tool  $\mathbf{K}_{tool}$  is derived from cantilever beam model (length of cutter disc  $l_t = 50$  mm,

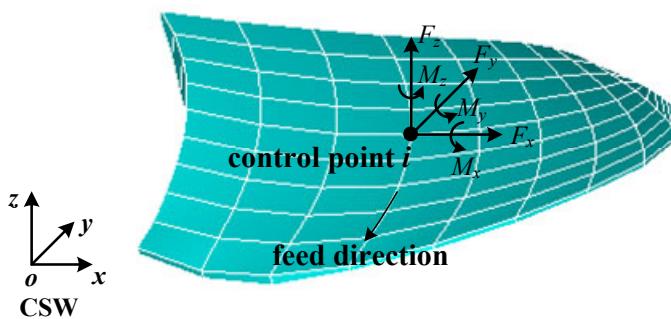
diameter of cutter disc D = 80mm, length of cutter bar  $l_2$  = 58 mm, diameter of cutter bar d=60mm; elastic modulus E = 450Gpa, Poisson ratio  $\nu$  = 0.28 ):

$$\mathbf{K}_{tool} = \begin{bmatrix} 3.3550 & 0 & 0 & 0 & 0.0992 & 0 \\ 0 & 3.3550 & 0 & -0.099 & 0 & 0 \\ 0 & 0 & 2.6045 & 0 & 0 & 0 \\ 0 & -0.099 & 0 & 0.0131 & 0 & 0 \\ 0.0992 & 0 & 0 & 0 & 0.0131 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0009 \end{bmatrix} \times 10^9 . \quad (10)$$

where dimension are the same as that are prescribed in matrix  $\mathbf{K}_{joint}$ .

Finite element model of the propeller blade is established to calculate stiffness matrix  $\mathbf{K}_p$  of workpiece at control points. Boundary condition is set to the blade root, and only vertical constraint is applied near the edge of blade tip. The finite element model and its control points are illustrated in figure 4 ( Elastic modulus 105Gpa and Poisson ratio 0.25). Control point  $i$  on the surface is selected. Stiffness matrix  $\mathbf{K}_p$  at this control point is obtained by finite element calculation as follows:

$$\mathbf{K}_p = \begin{bmatrix} 5.6698 & -1.5310 & 0.0587 & 0.1221 & 0.3783 & -0.2561 \\ -1.5310 & 4.3226 & 1.7703 & -0.3238 & -0.0590 & 0.1110 \\ 0.0587 & 1.7703 & 2.4165 & -0.2789 & 0.0125 & 0.0575 \\ 0.1221 & -0.3238 & -0.2789 & 0.0555 & 0.0055 & -0.0162 \\ 0.3783 & -0.0590 & 0.0125 & 0.0055 & 0.0412 & -0.0261 \\ -0.2561 & 0.1110 & 0.0575 & -0.0162 & -0.0261 & 0.0195 \end{bmatrix} \times 10^9 . \quad (11)$$



**Fig. 4.** Propeller blade and control point  $i$  on the surface

Substituting Eq. (7), (8), (9), (10), and (11) into Eq. (6), close-loop general stiffness matrix at this control point is obtained as follows:

$$\mathbf{K} = \begin{bmatrix} 0.5349 & 0.0127 & 0.0394 & -0.0077 & 0.0035 & -0.0003 \\ 0.0127 & 0.6196 & 0.1040 & -0.0192 & 0.0051 & 0.0014 \\ 0.0394 & 0.1040 & 0.418 & -0.0249 & 0.0061 & 0.0021 \\ -0.0077 & -0.0192 & -0.0249 & 0.0074 & -0.0018 & -0.0006 \\ 0.0035 & 0.0051 & 0.0061 & -0.0018 & 0.0012 & -0.0003 \\ -0.0003 & 0.0014 & 0.0021 & -0.0006 & -0.0003 & 0.0003 \end{bmatrix} \times 10^9 . \quad (12)$$

Close-loop general stiffness matrix of multi-axis process system at other control points can be obtained as well by the same method as above.

#### 4 Stiffness Performance Analysis and Improvement for 5 Axis Turn-Milling Machine Tool

Based on general closed-loop stiffness model, stiffness performance can be expressed as genreal force ellipsoid[15]. In order to evaluate stiffness performance with respecte to force, 3D force ellipsoid is obtained by decoupling method[16]. Firstly, dividing general force  $\mathbf{F}$  into force  $\mathbf{g}$  and moment  $\mathbf{m}$ :

$$\mathbf{F} = \begin{Bmatrix} \mathbf{g} \\ \mathbf{m} \end{Bmatrix} = \mathbf{K} \delta \mathbf{v} = \begin{bmatrix} \mathbf{K}_g \\ \mathbf{K}_m \end{bmatrix} \delta \mathbf{v} . \quad (13)$$

where  $\mathbf{g}$  is cutting force at the tool tip formulated as  $\mathbf{g} = \mathbf{K}_g \delta \mathbf{v}$  and  $\mathbf{m}$  is cutting moment at the tool tip formulated as  $\mathbf{m} = \mathbf{K}_m \delta \mathbf{v}$ . Matrix  $\mathbf{K}_g$  and  $\mathbf{K}_m$  are force-displacement Jacobi matrix and moment-displacement Jacobi matrix.

Let  $\mathbf{S}_g = (\mathbf{K}_g^T \mathbf{K}_g)^{-1} \mathbf{K}_g^T$ , the relationship of force  $\mathbf{g}$  and deflection  $\delta \mathbf{v}$  is:

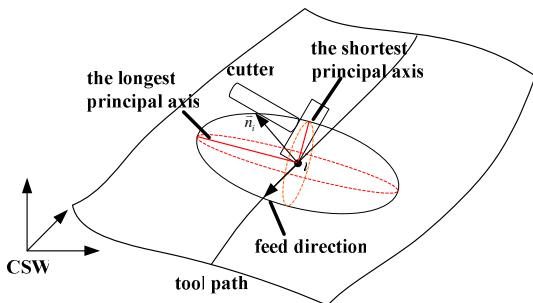
$$\mathbf{S}_g \mathbf{g} = \delta \mathbf{v} . \quad (14)$$

where  $\mathbf{S}_g$  is displacement-force Jacobi matrix.

Let  $\delta \mathbf{v}^T \delta \mathbf{v} = 1$ , then a formula about force-ellipsoid in 3-D space is:

$$\mathbf{g}^T \mathbf{S}_g^T \mathbf{S}_g \mathbf{g} = 1 . \quad (15)$$

The intersection of any vector from the center of force-ellipsoid reflects the stiffness along the vector. In order to improve production efficiency, the machine stiffness is required high enough to improve feed velocity. The length of the shortest principal axis of force ellipsoid reflects the poorest stiffness of whole machine in figure 5. Longer axis means that the stiffness of machine tool is higher, vice versa.



**Fig. 5.** 3D force ellipsoid

By extracting stiffness index (that is, the shortest principal axis of force ellipsoid) at each control point, isoclines map of stiffness index can be plotted in 2D space of CSW and be analyzed the distribution law of stiffness performance.

According to this process, stiffness performance of 5 axis process system including turn-milling machine tool, disc-type cutter and a large propeller is analyzed. Isoclines map of logarithmic scale of stiffness index are plotted. Improvement is proposed by changing cutter and adding the supporting points.

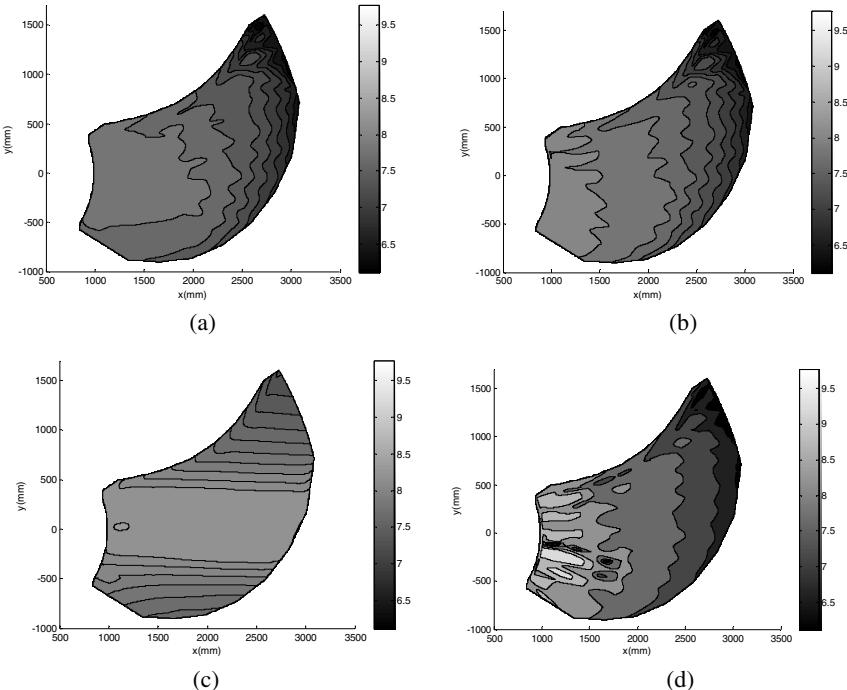
By analyzing the distribution law of stiffness performance, engineers can grasp the distribution law of anti-deformation capability of process system in CSW and judge the disadvantage area where chatter is prone to appear, where stiffness performance is much poorer. Possibly chatter or forced vibration can be reduced by adjusting machining strategy such as reducing back cutting depth, feedrate etc.

#### 4.1 Stiffness Performance Analysis for 5 Axis Turn-Milling Machine Tool

The general stiffness matrix can be expressed as generral force ellipsoid in geometric space. In order to evaluate general stiffness performance, 3D force ellipsoid is obtained by decoupling, and the shortest axis of the force ellipsoid is chosen as stiffness index  $\lambda_{min}$ . The stiffness index at the control point  $i$  is  $6.75 \times 10^7$  N/m. Other stiffness indices at other control points can also be obtained.

Stiffness index of all control points is derived from general stiffness matrix by the proposed method. And isoclines map of logarithmic scale of stiffness index are plotted for further analysis. As shown in figure 6(a),  $\lambda_{min}$  at the blade root is longer than that at the blade tip. It means that anti-deformation capability of manufacturing process system is the best at blade root and the worst at blade tip.

Figure 6(a)~(d) present isoclines map of logarithmic scale of stiffness index. The result in figure 6(b) is closed to figure 6(a). It means that the effect of workpiece and cutter on stiffness index is larger than machine tool. Stiffness performance of work-piece and cutter can be improved by altering conditions of the workpiece fixture and the cutter size. Then general stiffness performance can be further improved.

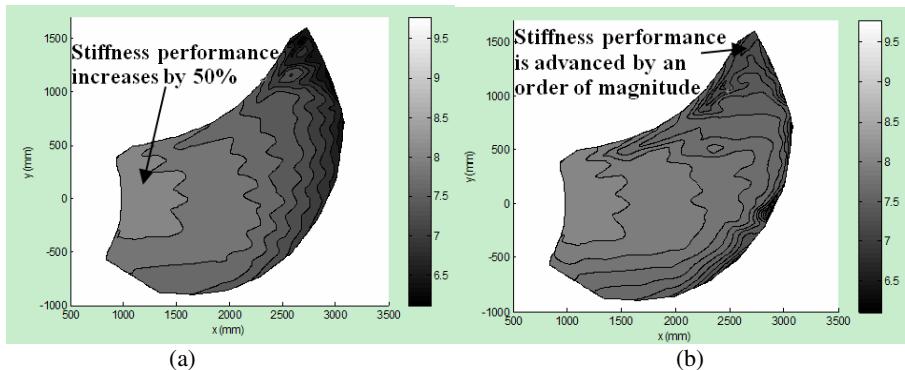


**Fig. 6.** Isoclines map of stiffness index (Firstly chosen cutter), (a)Stiffness model proposed in this paper, (b)Stiffness model in which only poor stiffness of cutter and workpiece is considered,(c) Stiffness model in which only poor stiffness of machine tool is considered, (d) Stiffness model in which only poor stiffness of workpiece is considered.

#### 4.2 Stiffness Performance Improvement for 5 Axis Turn-Milling Machine

Another disc-type cutter( length of disc 41mm, diameter of disc 115mm, length of cutter bar 61.3mm, diameter of cutter bar 75mm; material is the same) is chosen to machine the blade. Isoclines map can be obtained as shown in figure 7(a). Comparing figure 7(a) to 6(a), stiffness index increases by 50% nearly in peak (in figure 6(a), peak value of stiffness index is  $8.13 \times 10^7$  N/m; in figure 7(a), peak value of stiffness index is  $12.30 \times 10^7$  N/m). Stiffness index near the blade tip can be increased by future improving conditions of workpiece fixture.

In order to improve conditions of workpiece fixture, the supporting points on the edge of blade tip is added, where only vertical constraint is applied. Isoclines map is plotted in figure 7(b). Stiffness index near the blade tip is better than figure 7(a). In figure 7(a), stiffness index near the blade tip is low to  $10^6$  N/m. After improvement, stiffness index near the blade tip is up to  $10^7$  N/m in figure 7(b). Stiffness index near the blade tip is enhanced by roughly an order of magnitude. By improving conditions of workpiece fixture, not only general stiffness performance near the blade tip is enhanced, but also in other areas, for instance the right area on the surface of the blade in figure 7(b).



**Fig. 7.** Isoclines map of stiffness index: (a) Stiffness model proposed in this paper in which another cutter is adopted, (b) Stiffness model proposed in this paper in which fixture conditions is improved besides the replacement of cutter

## 5 Conclusion

Closed-loop general stiffness model of multi-axis process system is established by Jacobi matrix and finite element method, where main factors (including poor stiffness of machine tool, cutter and workpiece ) affecting general stiffness performance is considered. In this model, both of differential displacement mapping relationships from joint space to workpiece coordinate system(CSW) and that from cutter coordinate system to CSW are established by Jacobi matrix method to calculate stiffness matrix of machine tool and cutter. Stiffness matrix of workpiece is calculated by finite element method through extracting flexibility coefficient at any control point on the surface of workpiece. All of the stiffness matrix of machine tool, cutter and workpiece are finally used to form close-loop general stiffness matrix of multi-axis process system. Once general stiffness matrix of multi-axis process system is established, the distribution law of general stiffness performance can be evaluated in workspace and help plan tool path. Future research will focus on tool path planning and processing parameters optimization.

**Acknowledgments.** The authors would like to thank the anonymous referees for their valuable comments and helpful suggestions. This work was supported by Natural Science Foundation of China (50805057), Natural Science Foundation of China (50835004), National High-Tech Research and Development Program of China (2009AA04Z149) during all the period of this research.

## References

1. Zelinski, P.: Four types of five-axis machining centers. *Modern Machine Shop*. 71(10), 94–103 (1999)
2. Kruth, J.P., Klewais, P.: Optimization and dynamic adaptation of the cutter inclination during five axis milling of sculptured surfaces. *CIRP Annals - Manufacturing Technology* 43(1), 443–448 (1994)

3. Lee, Y.S.: Admissible Tool orientation control of gouging avoidance for 5-axis complex surface machining. *Computer Aided Design* 29(7), 507–521 (1997)
4. Budak, E.: Analytical models for high performance milling Part I: Cutting forces structural deflections and tolerance integrity. *International Journal of Machine Tools and Manufacture* 46, 1478–1488 (2006)
5. de López Lacalle, L.N., Lamikiz, A., Sánchez, J.A., Salgado, M.A.: Effects of tool deflection in the high-speed milling of inclined surfaces. *International Journal of Advanced Manufacture Technology* 24, 621–631 (2004)
6. Uriarte, L., Herreroa, A., Zataraina, M., Santisoa, G., Lopéz de Lacalle, L.N., Lamikiz, A., Albizurib, J.: Error budget and stiffness chain assessment in a micromilling machine equipped with tools less than 0.3 mm in diameter. *Precision Engineering* 31(1), 1–12 (2007)
7. DePincé, P., Hascoet, J.-Y.: Active integration of tool deflection effects in end milling Part 1. Prediction of milled surfaces. *International Journal of Machine Tools and manufacture* 46, 937–944 (2006)
8. Kim, G.M., Kim, B.H., Chu, C.N.: Estimation of cutter deflection and form error in ball-end milling processes. *International Journal of Machine Tools and Manufacture* 43, 917–924 (2003)
9. Ratchev, S., Liu, S., Huang, W., Becker, A.A.: An advanced FEA based force induced error compensation strategy in milling. *International Journal of Machine Tools and Manufacture* 46, 542–551 (2006)
10. Ratchev, S., Liu, S., Huang, W., Becker, A.A.: Milling error prediction and compensation in machining of low-rigidity parts. *International Journal of Machine Tools and Manufacture* 44, 1629–1641 (2004)
11. Wan, M., Zhang, W.H., Qin, G.H., Wang, Z.P.: Strategies for error prediction and error control in peripheral milling of thin-walled workpiece. *International Journal of Machine Tools and Manufacture* 48(12-13), 1366–1374 (2008)
12. de López Lacalle, L.N., Lamikiz, A.: Machine Tool Performance and Precision. *Machine Tools for High Performance Machining*, 219–260 (2009)
13. Youlun, X.: Robotic technology Foundation. Huazhong University of Science and Technology Press, wuhan (1996)
14. Rong, Y., Fangyu, P., Bin, L.: A Method of General Stiffness Modeling for Multi-axis Machine Tool. *Intelligent Robotics and Applications* 5315, 1013–1021 (2008)
15. Yoshikawa, T.: Foundation of robotics analysis and control. MIT Press, Cambridge (1990)
16. Rong, Y., Fangyu, P., Bin, L.: General Stiffness Analysis for Multi-Axis Machine Tool. In: *Advanced Design and Manufacture to Gain a Competitive Edge*. In: ICADAM (2008)

# Author Index

- Abdelgawad, Amr I-288  
Abdul-Rahman, Hussein S. II-717  
Adachi, Masaru I-226  
Alaqtash, Murad I-288  
Al-Bluwi, Ibrahim II-116  
Alboul, Lyuba II-717  
Aljanaideh, Omar I-104  
Alpen, Mirco II-215  
Antoun, Shérine M. II-130
- Bai, Yucheng II-227  
Bi, Qingzhen II-553  
BoYang, I-530  
Brower, Richard I-288  
Bui, Quoc Phuong I-442
- Cabibihan, John-John II-498, II-683  
Cao, Heng I-590  
Cao, Jiangtao II-486  
Cao, Qixin I-226  
Cao, Zhiqiang I-538  
Chang, Yau-Zen II-154  
Che, Demeng I-47  
Chen, Bing II-22  
Chen, Daguo II-475  
Chen, Genliang II-252  
Chen, Hui II-1  
Chen, Huiying I-749  
Chen, Peihua I-226  
Chen, Qiang II-56, II-695  
Chen, S.Y. I-300  
Chen, Wei II-754  
Chen, Xiang I-692  
Chen, Xiao I-590  
Chen, Xinpu I-81  
Chen, Xubing II-345  
Chen, Yixiong II-417  
Chen, Zhaopeng I-1  
Chen, Zheng I-159  
Cheng, Chen I-499  
Cheng, Shili I-409  
Chin, Chungwei II-683  
Choi, Jong-suk I-582
- Chou, Yu-Shin II-154  
Chuang, Chin-Wen I-266
- Dai, Hongjun II-695  
Dai, Jian S. I-430  
Dai, Lei I-171  
Dario, Paolo I-477  
Deng, Hua II-329, II-339  
Deng, Lei I-548, II-191  
Deng, Mingcong I-135  
Dijk, Judith I-665  
Dillmann, Ruediger I-26  
Ding, Han II-553, II-576  
Ding, Weili II-368  
Ding, Ye II-541  
Dong, Wei I-600  
Dorst, Leo I-665  
Du, Jian II-285  
Du, Juan I-92  
Du, Yingkui I-609  
Du, Zhengchun II-276
- Elnagar, Ashraf II-116  
Esteban, Isaac I-665  
Ewert, Daniel II-44
- Fan, Baojie I-609  
Fan, Hongchao II-390  
Fan, Huijie I-600  
Fan, Shaowei I-1, I-38  
Fang, Yinfeng I-300  
Fang, Zhou II-191  
Feng, Ying I-92  
Fu, Ruiqing I-548  
Fu, Xin I-477, II-296  
Fu, Xingwei II-227  
Fu, Zhuang I-400, I-519
- Ganapathy, Velappa II-624, II-636  
Gao, Zhen I-418  
Gasteratos, Antonios II-705  
Ge, Shuzhi Sam II-498, II-683  
Gong, Guofang I-329  
Gong, Hu II-67

- Goto, Takafumi II-97  
 Gu, Dongbing I-724  
 Gu, Guoying I-114  
 Guan, Enguang I-400, I-519  
 Guo, Qiang II-565  
 Guo, Qiuquan II-649  
 Guo, Wei I-257  
 Guo, Xinglin II-305  
 Guo, Yongxin I-147
- Hai, Dan I-713  
 Han, Dapeng II-75  
 Han, Feng I-570  
 Hashimoto, Koichi II-378  
 Hashimoto, Tskuya II-505  
 Haynes, Paul S. II-717  
 Hei, Mo I-509  
 Herrera, Jose Luis Alarcon I-692  
 Horn, Joachim II-215  
 Hou, Jiaoyi II-296  
 Hou, Jung-Fu II-154  
 Hou, Lei II-378  
 Hou, Zengguang II-417  
 Hu, Huosheng II-659  
 Hu, Pinghua I-81  
 Hu, Xiangtao I-383  
 Huang, Dan I-171  
 Huang, Jian II-440  
 Huang, Xinhuan I-558  
 Huang, Yong'an I-383  
 Hwang, Do-hyeong I-582
- Janaideh, Mohammad Al I-104  
 Jauw, Veronica Lestari II-612  
 Jeschke, Sabina II-44  
 Jia, Xinghua II-285  
 Jia, Zhenyuan II-285, II-565  
 Jiang, Dongsheng I-400  
 Jiang, Jiaxin II-390  
 Jiang, Li I-38  
 Jiang, Xianzhi II-463  
 Jin, Ming-He I-1, I-762  
 Ju, Zhaojie I-71
- Kagami, Shingo II-378  
 Kang, Zhan II-317  
 Kato, Naoki II-505  
 Kawano, Yuki II-357  
 Kawasue, Kikuhiro II-357  
 Khoury, Mehdi I-680
- Ko, Kwang-Eun I-702  
 Kobayashi, Hiroshi II-505  
 Kouskouridas, Rigas II-705  
 Kubota, Naoyuki II-486, II-517, II-529  
 Kunze, Ralph II-44
- Lai, Leijie I-114  
 Lee, Hyun I-237  
 Lee, Jang Myung I-237  
 Lee, Jun seok I-237  
 Leong, Benny I-276  
 Li, Bolan I-372  
 Li, Fei I-477  
 Li, Gang II-305, II-317  
 Li, Jau-Rong I-266  
 Li, Jiting II-398  
 Li, Miao II-10, II-240, II-264  
 Li, Nan I-38, I-642, I-654  
 Li, Pengfeng II-417  
 Li, Pengzhi I-114  
 Li, Ping II-486  
 Li, Qingling II-417  
 Li, Shan II-22  
 Li, Shunchong I-81  
 Li, Xiuli I-300  
 Li, Yanan II-683  
 Li, Yangmin I-214, II-451  
 Li, Yanming I-642, I-654  
 Li, Yingsong II-227  
 Li, Youfu I-749  
 Li, Zheng I-246  
 Li, Zhijun I-171
- Lii, Neal Y. I-1  
 Lin, Zhongqin II-252  
 Linsen II-754
- Liu, Chang-Huan II-203  
 Liu, Chengliang I-195, I-642, I-654  
 Liu, Dehao II-305  
 Liu, Deshi II-305  
 Liu, Gequn I-204  
 Liu, Haining I-642, I-654  
 Liu, Hong I-1, I-38, I-570, I-762  
 Liu, Honghai I-71, I-680, II-486  
 Liu, Huaiyin I-329  
 Liu, Jiajun II-695  
 Liu, Jianqin I-257  
 Liu, Jing-Sin II-154  
 Liu, Lei I-204  
 Liu, Mei II-649  
 Liu, Ran I-310

- Liu, Shuangjun II-285  
 Liu, Wei II-285  
 Liu, Weiting I-477  
 Liu, Xinhui II-180  
 Liu, Xin-Jun I-318  
 Liu, Yi-Wei I-762  
 Liu, Yu I-570  
 Liu, Zhenglin I-183  
 Lu, Huimin II-105  
 Lumia, Ron I-630  
 Luo, Dezhi I-356  
 Luo, Jun II-649  
 Luo, Yaqin I-310  
 Luo, Yun II-742  
 Luo, Zi-rong I-499, I-509
- Ma, Hongxu I-736  
 Ma, Wentao I-519  
 Ma, Xiande II-56  
 Maestas, Dominic R. I-630  
 Mahmoud, Abeer II-166  
 Manabe, Takao II-97  
 Mao, Jianqin I-125, I-147  
 Martínez-García, Edgar I-453  
 Matson, Eric T. I-276  
 Mavrinac, Aaron I-692  
 Mayer, Marcel II-44  
 McKerrow, Phillip J. II-130  
 Meng, Tao I-724  
 Miao, Zhonghua I-195  
 Miki, Hiroyuki II-742  
 Min, Huasong I-183  
 Mu, Hua I-736
- Nagatomo, Satoshi II-357  
 Nguyen, Cory Q. I-276  
 Ni, Fenglei I-570  
 Nishiyama, Masahiro II-357  
 Niu, Peifeng II-368
- Okada, Tokuji II-166  
 Ou, Yongsheng I-548, II-191  
 Oyong, Arif Wicaksono II-612
- Pan, Minghua I-392  
 Parasuraman, S. II-612, II-624  
 Park, Gye-do I-237  
 Penders, Jacques II-717  
 Peng, Fangyu II-754  
 Peng, Fuyuan I-530
- Peng, Liling II-390  
 Peng, Yan II-649  
 Poignet, Philippe II-428  
 Ponnambalam, S.G. II-587, II-599, II-636
- Qian, Jijie I-418  
 Qiao, Jinli I-257  
 Qiao, Jin-wei I-499
- Rakheja, Subhash I-104  
 Ren, Haoling II-730  
 Rönnau, Arne I-618  
 Ruan, Xiaogang II-87  
 Ruwadi, II-428
- Sarkodie-Gyan, Thompson I-288  
 Schamm, Thomas I-618  
 Schilberg, Daniel II-44  
 Shang, Jian-zhong I-499, I-509  
 Shang, Yonggang II-305  
 Shen, Xiong I-530  
 Shi, Hu I-329  
 Shi, Kun I-538  
 Shiota, Hiroyuki II-742  
 Sim, Kwee-Bo I-702  
 Smith, Anthony I-276  
 Spier, Eric I-288  
 Starr, Gregory I-630  
 Stefanini, Cesare I-477  
 Stuedemann, Stephan II-215  
 Su, Chun-Yi I-92  
 Sun, Haitao I-59  
 Sun, Jianzhong I-310  
 Sun, Jie I-59  
 Sun, Ronglei II-463  
 Sun, Wei I-310  
 Sun, Yuwen II-565  
 Sun, Zhenguo II-56, II-695
- Takagi, Toshiyuki II-742  
 Takeno, Takanori II-742  
 Tamura, Hiroki II-97  
 Tan, Dawei I-488  
 Tan, Guo yuan I-530  
 Tan, Min II-417  
 Tan, Qing I-348  
 Tang, Yandong I-600, I-609  
 Tanno, Koichi II-97  
 Tech, Ang Wei II-428

- Teh, Joshua II-636  
 Teng, Tao I-356  
 Tharin, Julien II-717  
 Thelen, Sebastian II-44  
 Tian, Bo I-442  
 Tian, Ke I-363  
 Tian, Yao-bin II-203  
 Toh, Chen Koon I-442  
 Torres-Córdoba, Rafael I-453  
 Veera Ragavan, S. II-636  
 Votelho, Wagner II-166  
 Wachs, Juan P. I-276  
 Wang, Aihui I-135  
 Wang, Bangguo II-285  
 Wang, Chaoqun I-409  
 Wang, Dongyun I-135  
 Wang, Hao II-252  
 Wang, Jingguo I-214, II-451  
 Wang, Jinsong I-318  
 Wang, Kai I-348  
 Wang, Licheng II-576  
 Wang, Lintao I-329  
 Wang, Qiyuan II-87  
 Wang, Shaojun II-576  
 Wang, Shuang II-398  
 Wang, Tianmiao I-183  
 Wang, Tongjian II-180  
 Wang, Xiangke II-75, II-105  
 Wang, Xi-ming II-203  
 Wang, Xinqing I-38  
 Wang, Xuezhu I-159  
 Wang, Xuyong I-195  
 Wang, Yongji II-440  
 Wang, Yu I-590  
 Wang, Yuhang II-553  
 Wang, ZhenYan I-125  
 Wang, Zhuo I-509  
 Wang, Zongyao I-724  
 Wei, Guowu I-430  
 Wei, Hongxing I-183  
 Wei, Lai II-659  
 Wood, John I-630  
 Wu, Chao I-318  
 Wu, Gongping II-227  
 Wu, Hongtao I-409  
 Wu, Jun II-440  
 Wu, Meiping I-736  
 Wu, Wenqi I-736  
 Wu, Xinyu I-548, II-191  
 Xia, Chunming I-590  
 Xia, Yi-min I-341, I-348, I-356  
 Xiang, Kangtai II-695  
 Xiao, hua II-227  
 Xiao, Zhihu I-530  
 Xie, Haibo II-730  
 Xie, Jinghua I-363  
 Xie, Ming I-442  
 Xie, Xiaoliang II-417  
 Xing, Kexin II-440  
 Xiong, Caihua II-463, II-541  
 Xiong, Youlun II-463  
 Xiu, CaiJing II-1  
 Xu, Feifei II-565  
 Xu, Guohua I-530  
 Xu, He I-488, II-142  
 Xu, Ke I-548  
 Xu, Wenfu I-570  
 Xu, Xiaolong I-530  
 Xu, Xiaoming I-204  
 Xu, Zhen II-329, II-339  
 Xu, Zijun I-348  
 Xue, Zhixing I-26  
 Xue, Yong I-499  
 Yaguchi, Aiko II-529  
 Yamashita, Yuki II-97  
 Yan, Rong II-754  
 Yan, Weixin I-400, I-519  
 Yan, Yonghua II-475  
 Yan, yu II-227  
 Yang, Dapeng I-38  
 Yang, Dawei I-363  
 Yang, Hejin I-442  
 Yang, Huayong II-730  
 Yang, Jianzhong II-754  
 Yang, Jun II-649  
 Yang, Lixin II-329  
 Yang, Ting II-22  
 Yang, Wenyu II-10, II-240, II-264  
 Yao, Yan-an II-203  
 Yao, Yu I-409  
 Yao, Zhenqiang II-276  
 Yap, Shee Cheng II-428  
 Yasuda, Gen'ichi II-32  
 Yi, Dong kyu I-237  
 Yin, Hang II-475

- Yin, Zhouping I-383  
Yogeswaran, M. II-587  
Yorita, Akihiro II-517  
Yu, Huiying I-288  
Yu, Naigong II-87  
Yuan, Wenzhen II-671  
Yun, Soh Chin II-624
- Zeng, Xiangjin I-558  
Zhang, Dan I-418  
Zhang, Dingguo I-81  
Zhang, Dinghua II-22  
Zhang, Feng II-417  
Zhang, Guoqing I-442  
Zhang, Houxiang I-300  
Zhang, Hui I-713, II-75, II-105  
Zhang, Kai II-317  
Zhang, Kui I-341  
Zhang, Pu II-276  
Zhang, Ruirui I-430  
Zhang, Wenwen I-538  
Zhang, Wenzeng I-13, I-47, I-59,  
I-465, II-56, II-671  
Zhang, XiaoJian II-541  
Zhang, Xiaoping II-10, II-240, II-264  
Zhang, Xu I-372  
Zhang, Xuguang II-368  
Zhang, Yanduo I-558  
Zhang, Yang II-180  
Zhang, Yitong I-257
- Zhang, Yong-de II-405  
Zhang, Yuan-Fei I-762  
Zhang, Zhen I-226, II-390  
Zhang, Zhenyu I-488, II-142  
Zhao, Deyang I-13, I-465  
Zhao, Jinfeng I-488  
Zhao, Kai II-252  
Zhao, Yan-jiang II-405  
Zhao, Yanzheng I-400, I-519  
Zhao, Yanzhi I-724  
Zheng, Guilin I-159  
Zheng, Lee Yi II-599  
Zheng, Ruoyin II-398  
Zheng, Zhiqiang I-713, II-105  
Zhou, Chao I-538  
Zhou, Fan II-345  
Zhou, Hong I-590  
Zhou, Hua II-296  
Zhu, Guoli I-392  
Zhu, Jianying I-409  
Zhu, Jun I-590  
Zhu, Limin I-114, I-372, II-553  
Zhu, Linlin I-609  
Zhu, Xiangyang I-81  
Zhu, Xiaojin I-195  
Zhu, Zongming I-356  
Zhuan, Xiangtao I-159  
Zöllner, J. Marius I-618  
Zou, Jun II-296