

Fuchun Sun  
Dewen Hu  
Huaping Liu *Editors*

# Foundations and Practical Applications of Cognitive Systems and Information Processing

Proceedings of the First International  
Conference on Cognitive Systems  
and Information Processing, Beijing,  
China, Dec 2012 (CSIP2012)

# **Advances in Intelligent Systems and Computing**

Volume 215

*Series Editor*

J. Kacprzyk, Warsaw, Poland

For further volumes:  
<http://www.springer.com/series/11156>

Fuchun Sun · Dewen Hu · Huaping Liu  
Editors

# Foundations and Practical Applications of Cognitive Systems and Information Processing

Proceedings of the First International  
Conference on Cognitive Systems  
and Information Processing, Beijing,  
China, Dec 2012 (CSIP2012)

 Springer

*Editors*

Fuchun Sun  
Huaping Liu  
Department of Computer Science and  
Technology  
Tsinghua University  
Beijing  
People's Republic of China

Dewen Hu  
College of Mechatronics and Automation  
National University of Defense Technology  
Changsha  
People's Republic of China

ISSN 2194-5357

ISBN 978-3-642-37834-8

DOI 10.1007/978-3-642-37835-5

Springer Heidelberg New York Dordrecht London

ISSN 2194-5365 (electronic)

ISBN 978-3-642-37835-5 (eBook)

Library of Congress Control Number: 2013947347

© Springer-Verlag Berlin Heidelberg 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law. The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

# Preface

This book is part of the Proceedings of the Seventh International Conference on Intelligent Systems and Knowledge Engineering (ISKE2012) and the First International Conference on Cognitive Systems and Information Processing (CSIP2012) held in Beijing, China, during December 15–17, 2012. ISKE is a prestigious annual conference on Intelligent Systems and Knowledge Engineering with the past events held in Shanghai (2006, 2011), Chengdu (2007), Xiamen (2008), Hasselt, Belgium (2009), and Hangzhou (2010). Over the past few years, ISKE has matured into a well-established series of International Conferences on Intelligent Systems and Knowledge Engineering and related fields over the world. CSIP2012 is the first conference sponsored by Tsinghua University and Science China Press, and technically sponsored by IEEE Computational Intelligence Society, Chinese Association for Artificial Intelligence. The aim of this conference is to bring together experts from different expertise areas to discuss the state-of-the-art in cognitive systems and advanced information processing, and to present new research results and perspectives on future development. Both ISKE2012 and CSIP2012 provide academic forums for the participants to disseminate their new research findings and discuss emerging areas of research. It also creates a stimulating environment for the participants to interact and exchange information on future challenges and opportunities of intelligent and cognitive science research and applications.

ISKE2012 and CSIP2012 received 406 submissions in total from about 20 countries (United States of America, Singapore, Russian Federation, Saudi Arabia, Spain, Sudan, Sweden, Tunisia, United Kingdom, Portugal, Norway, Korea, Japan, Germany, Finland, France, China, Argentina, Australia, and Belgium). Based on rigorous reviews by the Program Committee members and reviewers, among 186 papers contributed to CSIP2012, high-quality papers were selected for publication in the proceedings with the acceptance rate of 40.9 %. The papers were organized in 25 cohesive sections covering all major topics of intelligent and cognitive science and applications. In addition to the contributed papers, the technical program includes four plenary speeches by Jennie Si (Arizona State University, USA), Wei Li (California State University, USA), Chin-Teng Lin (National Chiao Tung University, Taiwan, China), and Guoqing Chen (Tsinghua University, China).

As organizers of both conferences, we are grateful to Tsinghua University, Science in China Press, Chinese Academy of Sciences for their sponsorship, grateful to IEEE Computational Intelligence Society, Chinese Association for Artificial Intelligence, State Key Laboratory on Complex Electronic System Simulation, Science and Technology on Integrated Information System Laboratory, Southwest Jiaotong University, University of Technology, Sydney, for their technical co-sponsorship.

We would also like to thank the members of the Advisory Committee for their guidance, the members of the International Program Committee and additional reviewers for reviewing the papers, and members of the Publications Committee for checking the accepted papers in a short period of time. Particularly, we are grateful to thank the publisher, Springer, for publishing the proceedings in the prestigious series of Advances in Intelligent Systems and Computing. Meanwhile, we wish to express our heartfelt appreciation to the plenary speakers, special session organizers, session chairs, and student volunteers. In addition, there are still many colleagues, associates, and friends who helped us in immeasurable ways. We are also grateful to them all. Last but not the least, we are thankful to all authors and participants for their great contributions that made ISKE2012 and CSIP2012 successful.

December 2012

Fuchun Sun  
Dewen Hu  
Huaping Liu

# Organizing Committee

|                                     |   |
|-------------------------------------|---|
| <b>General Chairs:</b>              | Bo Zhang, Tsinghua University, China<br>Yen Gray, Oklahoma State University, USA            |
| <b>Honorary Chairs:</b>             | Jianguang Sun, China<br>Deyi Li, China  |
| <b>Advisory Committee Chairs:</b>   | Nanning Zheng, China<br>Wei Li, China<br>Sam Shuzhi Ge, National University<br>of Singapore |
| <b>Steering Committee Chairs:</b>   | Chin-Teng Lin, Taiwan, China<br>Zongben Xu, China   |
| <b>Organizing Committee Chairs:</b> | Zengguang Hou, China<br>Xiaorong Gao, China<br>Huaping Liu, China                           |
| <b>Program Chairs:</b>              | Fuchun Sun, China<br>Jenni Si, USA<br>Dewen Hu, China                                       |
| <b>Plenary Sessions Chair:</b>      | Derong Liu, Chicago, USA  |
| <b>Sessions Chairs:</b>             | Haibo He, USA<br>Fei Song, China  |
| <b>Publications Chairs:</b>         | Dianhui Wang, Australia<br>Guangbin Huang, Singapore  |
| <b>Publicity Chair:</b>             | Jun Zhu, China  |
| <b>Finance Chair:</b>               | Hongbo Li, China  |
| <b>Registration Chairs:</b>         | Jianmin Li, China<br>Caixia Chen, China   |
| <b>Local Arrangements Chairs:</b>   | Huaping Liu, China<br>Linmi Tao, China  |
| <b>Electronic Review Chair:</b>     | Xiaolin Hu, China   |

## Advisory Committee

### Advisory Committee Chairs

|               |  |
|---------------|--|
| Nanning Zheng | President of Xi'an Jiaotong University, Member of Chinese Academy of Engineering, China                                |
| Wei Li        | Previous president of Beijing University of Aeronautics and Astronautics, Member of Chinese Academy of Sciences, China |
| Sam Shuzhi Ge | Vice President of IEEE Control System Society, National University of Singapore, IEEE Fellow                           |

### Members

|                  |   |
|------------------|---|
| Tianyou Chai     | Member of the Chinese Academy of Engineering, North-eastern University, China |
| Aike Guo         | Member of the Chinese Academy of Science, Chinese Academy of Sciences, China  |
| Okay Kaynak      | Bogazici University, Turkey   |
| Yanda Li         | Member of Chinese Academy of Science, Tsinghua University, China              |
| Shoujue Wang     | Member of Chinese Academy of Sciences, Chinese Academy of Sciences, China     |
| Donald Wunsch II | Missouri University of Science and Technology, USA                            |
| Xin Yao          | IEEE Fellow, University of Birmingham, UK                                     |
| Lei Xu           | The Chinese University of Hong Kong, Hong Kong                                |
| Jacek M. Zurada  | IEEE Fellow, University of Louisville, USA                                    |
| Shangkai Gao     | IEEE Fellow, Tsinghua University, China                                       |
| Lin Chen         | Member of Chinese Academy of Engineering, Chinese Academy of Science, China   |

## Steering Committee

### Steering Committee Chairs

|               |  |
|---------------|--|
| Chin-Teng Lin | Deputy President of National Chiao Tung University, Taiwan, China                        |
| Zongben Xu    | Vice President of Xi'an Jiaotong University, Member of Chinese Academy of Science, China |

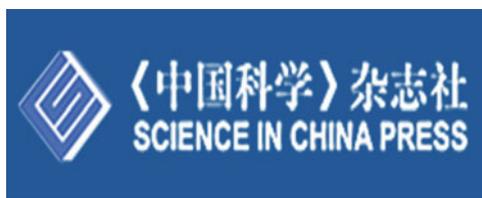
## **Members**

|               |  |
|---------------|--|
| Guangren Duan | Harbin Institute of Technology, China                |
| Jinde Cao     | Southeast University, China                          |
| Wen Yu        | National Polytechnic Institute, Mexico               |
| Aiguo Song    | Southeast University, China                          |
| Cong Wang     | South China University of Technology, China          |
| Jianwei Zhang | Hamburg University, Germany                          |
| Zhigang Zeng  | Huazhong University of Science and Technology, China |

## Sponsors



Tsinghua University



Science in China Press



Chinese Academy of Sciences



Institute of Electrical and Electronics Engineers

# Contents

|   |           |
|---|-----------|
| <b>Effects of Stimulus Views on Mental Rotation of Hands:<br/>An Event-Related Potential Study . . . . .</b>                        | <b>1</b>  |
| Xiaogang Chen, Guangyu Bin and Xiaorong Gao   |           |
| <b>Predictive Coding with Context as a Model<br/>of Image Saliency Map . . . . .</b>  | <b>15</b> |
| Duzhen Zhang and Chuancai Liu   |           |
| <b>Multiclass Pattern Analysis of Whole-Brain Functional Connectivity<br/>of Schizophrenia and Their Healthy Siblings . . . . .</b> | <b>25</b> |
| Yang Yu, Hui Shen, Ling-Li Zeng and Dewen Hu  |           |
| <b>Network Organization of Information Process<br/>in Young Adults' Brain . . . . .</b>   | <b>35</b> |
| Shao-Wei Xue, Yi-Yuan Tang and Lan-Hua Zhang  |           |
| <b>Reconfigurable Control Allocation of Multi-Surfaces<br/>Aircraft Based on Improved Fixed Point Iteration . . . . .</b>           | <b>45</b> |
| Kejun Bi, Weiguo Zhang, Chengzhi Chi and Jingkai Zhang  |           |
| <b>Particle Filter-Based Object Tracking and Handover<br/>in Disjoint View Multi-Cameras. . . . .</b>                               | <b>57</b> |
| Xiaoyan Sun, Faliang Chang and Wenhui Dong  |           |
| <b>Analyzing Effects of Pressing Radio Button on Driver's<br/>Visual Cognition . . . . .</b>  | <b>69</b> |
| Huacai Xian, Lisheng Jin, Haijing Hou, Qingning Niu and Huanhuan Lv   |           |
| <b>MAC Protocol Used Progressive Reservation Mechanism<br/>in Ad hoc Networks. . . . .</b>  | <b>79</b> |
| Hai-dong Yang, Bo Jing, Jian-hai Li and Xin Xiang   |           |

**An Automatic SSVEP Component Selection Measure for High-Performance Brain-Computer Interface . . . . .** 93  
 Zimu Zhang and Zhidong Deng

**Human Gender Differences in Cognitive Preferences Toward Attractive Faces in a Visual Oddball Paradigm: An ERP Study . . . . .** 109  
 Zimu Zhang and Zhidong Deng

**An Optimized Particle Filter Based on Improved MCMC Sampling Method . . . . .** 121  
 Aili Sang and Zhiquan Feng

**A Novel Spectrum Sensing Algorithm in Cognitive Radio System Based on OFDM . . . . .** 131  
 Liu Yun, Qicong Peng, Fuchun Sun, Huaizong Shao, Xingfeng Chen and Ling Wang

**Target Tracking Algorithm Based on Multi-Subblock Feature Matching . . . . .** 139  
 Faliang Chang, Wenhui Dong and Li Ma

**Intuitive Systemic Models and Intrinsic Features for Radar-Specific Emitter Identification . . . . .** 153  
 Tao Han and Yiyu Zhou

**RGBD SLAM for Indoor Environment . . . . .** 161  
 Rongyi Lin, Yangzhu Wang and Songpu Yang

**Neighborhood-Based Variable-Scale Model in Mobile Maps . . . . .** 177  
 Chaode Yan, Wang Guo, Jianjun Bai and Tian He

**ODE-LM: A Hybrid Training Algorithm for Feedforward Neural Networks . . . . .** 187  
 Li Zhang, Hong Li and Dazheng Feng

**Weather Condition Recognition Based on Feature Extraction and K-NN . . . . .** 199  
 Hongjun Song, Yangzhou Chen and Yuanyuan Gao

**Improvements of UMHExagonS Algorithm for Fast Motion Estimation in H.264. . . . .** 211  
 Hong-jian Cao and Gang Song

**Research on Middle-Semantic Manifold Object Annotation . . . . .** 221  
Wengang Feng and Shaozhong Wu

**Research on HVS-Inspired, Parallel, and Hierarchical Scene Classification Framework. . . . .** 233  
Wengang Feng and Xiping Zhou

**Adaptive Sub-Channel Allocation Based on Hopfield Neural Network for Multiuser OFDM . . . . .** 243  
Sufang Li, Mingyan Jiang, Anming Dong and Dongfeng Yuan

**CUDA on Hadoop: A Mixed Computing Framework for Massive Data Processing . . . . .** 253  
Zhanghu Wang, Pin Lv and Changwen Zheng

**Human Appearance Matching Based on Major Color and Spatio-Texture Features Across Disjoint Camera Views. . . . .** 261  
Biao Yang and Guoyu Lin

**Fast Image Dehazing Using Fuzzy System and Hybrid Evolutionary Algorithm. . . . .** 275  
Hongjun Song, Yuanyuan Gao and Yangzhou Chen

**Anti-Interference Performance Improvement Using Probability Control in Cognitive CDMA Communication System . . . . .** 285  
Sheng Hong, Bo Zhang and Hongqi Yang

**A Stereo Matching Algorithm Based on Outline-Assisted Dynamic Programming . . . . .** 297  
Pei Wang, Chen Chen and FangFang Wei

**Dynamic Visual Time Context Descriptors for Automatic Human Expression Classification . . . . .** 307  
Yi Ji, Shengrong Gong and Chunping Liu

**Moving Objects Detecting and Tracking for Unmanned Aerial Vehicle . . . . .** 317  
Binpin Su, Honglun Wang, Xiao Liang and Hongxia Ji

**Object Recognition and Pose Estimation Based on Principle of Homology-Continuity . . . . .** 335  
Zhonghua Hao and Shiwei Ma

**Vision-Based Traffic Sign Recognition System for Intelligent Vehicles. . . . .** 347  
 Jing Yang, Bin Kong and Bin Wang

**An Iterative Method for Classifying Stroke Subjects’ Motor Imagery EEG Data in the BCI-FES Rehabilitation Training System . . . . .** 363  
 Hao Zhang, Jianyi Liang, Ye Liu, Hang Wang and Liqing Zhang

**The Research and Application of Multi-Resource Heterogeneous Data Fusion on Dynamic Traffic Routing System . . . . .** 375  
 Youli Ren, Depin Peng, Jianping Wu and Yuan Zhou

**Data Fusion for the Diagnostics, Prognostics, and Health Management of Aircraft Systems . . . . .** 389  
 Zheng Liu and Nezih Mrad

**Multiscale Image Segmentation via Exact Inference of Hidden Markov Tree. . . . .** 401  
 Yinhui Zhang, Zifen He, Jinhui Peng and Yunsheng Zhang

**Cognitive Emotion Research of Humanoid Expression Robot . . . . .** 413  
 Jizheng Yan, Zhiliang Wang and Siyi Zheng

**Detecting System of Ink Cells in Gravure Cylinder via Neural Network. . . . .** 427  
 Zifen He, Zhaolin Zhan and Yinhui Zhang

**An Underwater Laser Image Segmentation Algorithm Based on Pulse Coupled Neural Network and Morphology . . . . .** 437  
 Bo Wang, Lei Wan and Tie-dong Zhang

**Locality Preserving Discriminant Projection for Total-Variability-Based Language Recognition . . . . .** 451  
 Xianliang Wang, Jinchao Yang, Chunyan Liang, Ruohua Zhou and Yonghong Yan

**Adaptive Spectrum Detecting Algorithm in Cognitive Radio. . . . .** 461  
 Yun Liu, Qicong Peng, Fuchun Sun, Huaizong Shao, Xingfeng Chen and Ling Wang

**The Autonomous Positioning Method for the Mars Probes Based on Cognizing Optical Information . . . . .** 477  
 Yingli Chang, Xiaohua Yuan and Dongmei Huang

**A Cognitive-Heuristic Framework for Optimization of Spaceplane-System Configurations . . . . .** 489  
 Ali Sarosh, Yun-Feng Dong and Shi-Ming Chen

**Artificial Bee Colony Algorithm for Parametric Optimization of Spacecraft Attitude Tracking Controller . . . . .** 501  
 Shan Zhong, Yun-Feng Dong and Ali Sarosh

**Routing for Predictable LEO/MEO Multi-Layered Satellite Networks . . . . .** 511  
 Heyu Liu and Fuchun Sun

**Improved ICP Algorithm with Bounded Rotation Angle for 2D Point Set Registration . . . . .** 523  
 Chunjia Zhang, Shaoyi Du, Jianru Xue and Xiaolin Qi

**Obstacle Detection for On-Road Vehicle Based on Range and Visual Information Fusion . . . . .** 531  
 Lipu Zhou

**Decoupled Parameter Estimation for Coherently Distributed Source . . . . .** 543  
 Yinghua Han, Jinkuan Wang and Qiang Zhao

**A Novel Blind Image Restoration Algorithm Using A SVR-Based Noise Reduction Technique . . . . .** 555  
 You Sheng Xia and Shi Quan Bin

**Effects of Music’s Emotional Styles and Tempo on Driving Behavior and Eye Movement: A Driving Simulation Study . . . . .** 565  
 Meng Yang, Jianqiao Wang, Yuqi Xia, Fan Yang and Xuemin Zhang

**An Improved HOG Based Pedestrian Detector . . . . .** 577  
 Chao Gao, Fengcai Qiao, Xin Zhang and Hui Wang

**A Reconfigurable Array Synthesis Method Using the Correlation Weightings of Smooth Local Trigonometric Base . . . . .** 591  
 Sheng Hong, Bo Zhang and Hongqi Yang

**A Multi-channel SSVEP-Based Brain–Computer Interface Using a Canonical Correlation Analysis in the Frequency Domain . . . . .** 603  
 Guang Chen, Dandan Song and Lejian Liao

**Acquiring Brain Signals of Imagining Humanoid Robot Walking Behavior via Cerebot. . . . .** 617  
 Wei Li, Yunyi Li, Genshe Chen, Qinghao Meng,  
 Ming Zeng and Fuchun Sun

**Multimodal Mixed Conditional Random Field Model for Category-Independent Object Detection . . . . .** 629  
 Jian-Hua Zhang, Jian-Wei Zhang, Sheng-Yong Chen and Ying Hu

**Research on Orbiting Information Procession of Satellites Based on Parallel Management . . . . .** 643  
 Yining Song and Dongdong Yan

**Multi-Sensor Multi-Target Detection Based on Joint Probability Density . . . . .** 655  
 Can Xu, Zhi Li and Lei Shi

**Robot Learning of Everyday Object Manipulation Using Kinect. . . . .** 665  
 Nan Chen, Ying Hu, Jun Zhang and Jianwei Zhang

**Research and Development of Automatic Driving System for Intelligent Vehicles. . . . .** 675  
 Weizhong Zhang, Tao Mei, Huawei Liang, Bichun Li, Jian Huang,  
 Zhaosheng Xu, Yi Ding and Wei Liu

**The Application of V Test Method in Detecting Auditory Steady-State Response . . . . .** 685  
 Jun Ying, Zheng Yan, Guangyu Bin and Xiaorong Gao

**Neural Network-Based Adaptive Dynamic Surface Control for Inverted Pendulum System. . . . .** 695  
 Enping Wei, Tieshan Li, Junfang Li, Yancai Hu and Qiang Li

**Unmanned Aircraft Vehicle Path Planning Based on SVM Algorithm . . . . .** 705  
 Yanhong Chen, Wei Zu, Guoliang Fan and Hongxing Chang

**Modeling Passenger Flow Distribution Based on Disaggregate Model for Urban Rail Transit . . . . .** 715  
 Da-lei Wang, En-jian Yao, Yang Yang and Yong-sheng Zhang

**Information Consensus in Distributed Systems Under Local Communication and Switching Topologies . . . . .** 725  
 Shijie Zhang and Yi Ning

**How Can We Find the Origin of a Path We Visually Traveled? The Effect of the Visual Environment on Path Integration . . . . .** 739  
 Huiting Zhang and Kan Zhang

**Missile Turbofan Engine Fault Diagnosis Technology and Its Application . . . . .** 751  
 Rui Cheng and Jiayuan Dan

**A Novel Acquisition Scheme for a GPS Software Receiver Based on Two Stand-Alone One-Dimensional Search Processes. . . . .** 763  
 Zhiguo Liu, Dacheng Luo, Shicheng Wang, Zhanxin Cheng and Lihua Chen

**In-hand Manipulation Action Gist Extraction from a Data-Glove . . . . .** 773  
 Gang Cheng, Norman Hendrich and Jianwei Zhang

**Development of an Intelligent Omnivision Surveillance System . . . . .** 783  
 Hannes Bistry and Jianwei Zhang

**Active Scene Analysis Based on Multi-Sensor Fusion and Mixed Reality on Mobile Systems. . . . .** 795  
 Denis Klimentjew, Sebastian Rockel and Jianwei Zhang

**Object Learning with Natural Language in a Distributed Intelligent System: A Case Study of Human-Robot Interaction . . . . .** 811  
 Stefan Heinrich, Pascal Follheer, Peer Springstübe, Erik Strahl, Johannes Twiefel, Cornelius Weber and Stefan Wermter

**Verbally Assisting Virtual-Environment Tactile Maps: A Cross-Linguistic and Cross-Cultural Study . . . . .** 821  
 Kris Lohmann, Junlei Yu, Matthias Kerzel, Dangxiao Wang and Christopher Habel

**Structural Similarity-Optimal Total Variation Algorithm for Image Denoising . . . . .** 833  
 Yu Shao, Fuchun Sun, Hongbo Li and Ying Liu

**Dexterous Robotic-Hand Grasp Learning Using Piecewise Linear Dynamic Systems Model. . . . .** 845  
 Wei Xiao, Fuchun Sun, Huaping Liu and Chao He

**Low-Rank Matrix Recovery for Traffic Sign Recognition in Image Sequences . . . . .** 857  
 Deli Pei, Fuchun Sun and Huaping Liu

**Locality-Constrained Linear Coding with Spatial Pyramid Matching for SAR Image Classification . . . . . 867**  
Shanshan Zhang, Fuchun Sun and Huaping Liu

**Quantitative Evaluation of Interictal High Frequency Oscillations in Scalp EEGs for Epileptogenic Region Localization . . . . . 877**  
Yaozhang Pan, Cuntai Guan, How-Lung Eng, Shuzhi Sam Ge, Yen Ling Ng and Derrick Wei Shih Chan

**Erratum to: A Cognitive-Heuristic Framework for Optimization of Spaceplane-System Configurations . . . . . E1**  
Ali Sarosh, Yun-Feng Dong and Shi-Ming Chen

# Effects of Stimulus Views on Mental Rotation of Hands: An Event-Related Potential Study

Xiaogang Chen, Guangyu Bin and Xiaorong Gao

**Abstract** Mental rotation of hands, which is subject to biomechanical constraints, involves participants engaging in motor imagery processing. To contribute to a better understanding of the process of hand mental rotation, reaction times and event-related potential were measured while participants were performing a left-right hand recognition task. Participants apparently solved the task by imagining their own hands rotating to the orientation of the stimulus for comparison. In line with previous studies, the behavioral results showed that slower reaction times were found for the hand views that could not be easily reached with real movement. More importantly, the event-related potential results revealed that the amplitude of rotation-related negativity (RRN) decreased with the difficulty of the hand views increasing. The previous results are complemented by this study; it is stimulus views that modulate reaction times and the amplitude of RRN during mental rotation task of hands.

**Keywords** Biomechanical constrain · Mental rotation · Motor imagery · Event-related potential

## 1 Introduction

Mental rotation is a well-known paradigm used to study the cognitive process of mentally rotating objects. It was first observed by Shepard and Metzler in 1971 when participants were shown pairs of three-dimensional block drawings at

---

X. Chen · G. Bin · X. Gao (✉)

Department of Biomedical Engineering, School of Medicine, Tsinghua University,  
Beijing 100084, China  
e-mail: gxr-dea@tsinghua.edu.cn

G. Bin

The College of Life Science and Bio-Engineering, Beijing University of Technology,  
Beijing 100124, China

different orientations and required to distinguish whether both drawings were presented as the same or the mirrored image version [1]. Results showed that the reaction times (RTs) increased with increasing angles of rotation. At the psychological level, this effect saliently explained that the image is mentally rotated to align it with its upright position. Since their discovery, mental rotation has been widely investigated [2–6]. The mental rotation paradigm has been extensively applied not only for rotation of 3-D cubes but also for letters, numbers, hands, and feet pictures [7, 8].

The investigation of hand mental rotation is classically performed by judgment of hand laterality. In such mental rotation tasks, participants tend to use a set of mental transformations of their own hands in order to closely match the stimulus. The reaction times for judging hand laterality also increased with increasing angle of rotated stimulus from its vertical position. Furthermore, hand mental rotation is sensitive to proprioceptive constraints, leading to longer reaction times for stimulus depicting anatomically difficult postures. These reaction times profiles were interpreted as evidence of kinesthetic limitations on movement imposed by the biomechanical constraints of the muscles and joints [9]. These constraints led to the postulation that participants utilized an embodied cognitive processing, which was classically referred to as motor imagery.

With the development of functional brain imaging methods such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), a number of neuroimaging studies have been conducted to illustrate brain regions involved in mental rotation tasks [10–12]. Kosslyn et al. measured the regional cerebral blood flow (rCBF) while participants were performing mental rotation of Shepard–Metzler’s drawings and hands. For the Shepard–Metzler type of stimuli, increased rCBF was observed in the parietal lobe and Area 19. In contrast, for mental rotation task of hands, there was significantly greater rCBF in the precentral gyrus, superior and inferior parietal lobes, primary visual cortex, insula, and frontal Areas 6 and 9. The authors claimed that at least two different mechanisms were involved in mental rotation, one for objects like hands related to motor preparation processes and another for objects like Shepard–Metzler’s drawings that did not [13]. A recent fMRI study compared activation regions while participants performed mental rotation task of hands and tools. In addition to replicating classic activation regions in mental rotation of hands and tools (bilateral superior parietal lobes and visual extrastriate cortex), there was an important difference in premotor area activation. Pairs of hands engendered bilateral premotor activation while tools only activated the left premotor cortex. The authors concluded that participants might simulate moving objects with their hands [14]. Results of PET and fMRI indicated that mental rotation was correlated with activity in parietal, occipital, frontal regions, as well as primary motor cortex, and the supplementary motor region of the precentral sulcus [15–17]. The involvement of the motor cortex in mental rotation task of hands is also supported by evidence from transcranial magnetic stimulation (TMS) studies. Pelgrims et al. found Brodmann area 4 (BA4) virtual lesions selectively increased reaction times in the left–right hand recognition task while leaving the mental rotation task of letters

unaffected. The authors suggested that BA4 contribution to mental rotation task of hands [18]. Activation in motor-related cortical areas implied that motor imagery was being applied during hand mental rotation task, indicating that motor structures might play an important role in supporting the mental rotation task of hands.

In contrast to neuroimaging techniques such as PET and fMRI, electroencephalography (EEG) has exquisite temporal resolution that make it possible to measure the timing and order of activity of brain regions during the unfolding of mental rotation task. Electrophysiological studies by event-related potential (ERP) showed that the mental rotation tasks elicited a modulated positive component approximately between 300 and 700 ms after stimulus presentation located at parietal region. This component was known as rotation-related negativity. The amplitude of this component was inversely related to the rotation angle and becomes more negative with increasing angles of rotation [19, 20]. Wijers et al. suggested that the gradual decrease of the positivity was caused by a modulation of a slow negativity that would be understood as a direct electrophysiological correlate of the mental rotation process [21]. This notion was extensively validated in various studies, suggesting that the ERP effect observed during mental rotation was highly specific to the process of mental rotation itself.

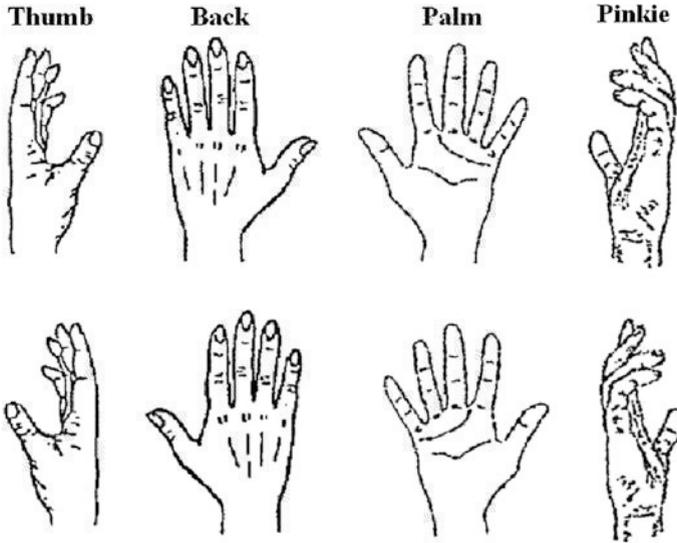
In this study we employed a hand mental rotation task similar to that used in the upper limb amputees' study by Nico et al. on the basis that there was a close similarity in brain activity between mental rotation of a movement and its real counterpart. Participants were presented with left–right hands at four different views and were asked to perform left–right hand recognition task. Previous studies have focused on the effect of rotational angle in general. To our knowledge, this is the first study investigating view effect on mental rotation of hands. If the mental rotation of hands relied on a first person perspective, we expect that it will be also more sensitive to the stimulus view. Accordingly, for more awkward stimulus view, we expect longer reaction times and more negative ERP component in the mental rotation task of hands.

## 2 Materials and Methods

### 2.1 Participants and Stimuli

Fourteen healthy participants (seven males) aged between 19 and 26 years [mean  $\pm$  standard deviation (SD): 24.3  $\pm$  1.93 years] took part in this experiment. All participants were right-handed, had normal or corrected to normal vision and had no history of psychiatric or neurological disorders. All participants signed an informed consent prior to the experiment and were paid for their participation.

Participants sat approximately 60 cm away from a 19-inch CRT monitor. Stimuli were black and white line drawings of both left and right hands presented from different views (thumb, back, palm, and pinkie side). Each of the eight hands



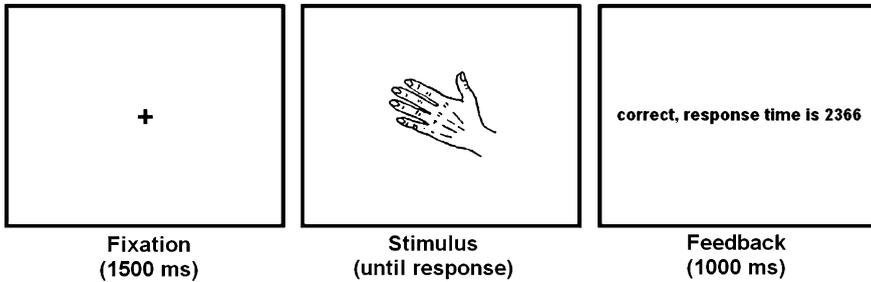
**Fig. 1** The stimuli used in the hand rotation task

was rotated clockwise in a step of  $30^\circ$  (12 rotations), which produced a total of 96 pictures (Fig. 1). The pictures were displayed randomly at the center of the screen on a background, which created an approximate  $8^\circ$  viewing angle. The stimulus presentation was managed by a computer program (VC++) written by the authors.

## 2.2 Procedure

Participants sat comfortably in a sound attenuated and electrically shielded EEG room. They were instructed to rest their right index and ring fingers on the left and right arrow of the keyboard, representing the left and right hand, respectively. Participants were asked to respond as quickly and accurately as possible, with a motor imagery strategy on how to solve the task, and not to be too cautious in their response. The experiment began with a practice phase to familiarize participants with the procedure. When the accuracy rate reached 75 %, the practice phase was ended. During the formal test phase, participants were instructed to respond by the fingers alone, without any hand, arm, or wrist movement. Furthermore, they were told to keep their eyes fixated on the central area of the screen while keeping their hands on their laps and out of sight, avoiding eye blinking and any sort of movement. After finishing one block, each participant was asked whether the motor imagery strategy was used or not.

The experiment was separated into three blocks of 96 trials. Each trial began with a green fixation cross lasting for 1,500 ms at the central area of the screen.



**Fig. 2** Sequence of events within one trial in the mental rotation paradigm

Subsequently, the stimulus was displayed and did not disappear until the participant made a response. Then feedback on their performance of the participants were shown and lasted for 1,000 ms in the center of the screen, including the reaction time and whether the stimulus was correctly sorted (Fig. 2). At the end of each block, the accuracy rate and reaction time were presented to give participants insight into their progress.

### **2.3 Data Acquisition**

For behavioral data acquisition, reaction times were defined as the period during which the stimulus appeared on the screen and the onset of the participant's response. Corresponding information regarding the participant's response was automatically recorded by appropriate software.

EEG data were recorded continuously using SynAmps2 amplifiers and Scan 4.3 software (Compumedics Neuroscan, EI Paso, TX, USA) with a 64-channel ElectroCap according to the international 10–20 system. EEG data were referenced online to a point midway between Cz and CPz and grounded midway between Fz and FPz. Electrode impedances were kept below 10 kohm. The horizontal electrooculogram (EOG) was recorded from electrodes placed next to the outer canthi of each eye. For the vertical EOG, electrodes were placed at the supraorbital and infraorbital regions of the left eye. Data were digitally amplified and collected at a sampling rate of 1,000 Hz.

### **2.4 Data Analysis**

Reaction times and accuracy rates were analyzed to ensure that the participants used the motor imagery strategy in the procedure of the task. If they did, it would be expected that more time was taken for difficult hand views [22]. Reaction times and accuracy rates were calculated in separate analyses of variance (ANOVA)

with factor view (thumb, back, palm, and pinkie side). Errors and trials with RTs shorter than 200 ms or longer than 3 standard deviations from the mean were removed before analysis of reaction times and left out from the EEG analyses; 7.9 % data were eliminated for all participants in this task. The Greenhouse–Geisser correction was applied if the data did not conform to the sphericity assumption by Mauchly’s Test of Sphericity. Whenever a main effect reached significance, pairwise comparisons were conducted by using Bonferroni adjustments for multiple comparisons. The alpha level was set at 0.05.

Raw EEG data were bandpass filtered between the frequency ranges from 0.5 to 40 Hz (24 dB/octave) and digitally re-referenced to the average of the left and right mastoids. Ocular artifacts were corrected using a linear regression procedure [23]. The continuous EEG was segmented subsequently into epochs of  $-200$  to  $+1000$  ms relative to the stimulus onset. Then, the prestimulus baseline of 200 ms was performed for baseline correction. Epochs with amplitude above  $\pm 100$   $\mu$ V were rejected (0.9 % of the whole trials). Then epochs were averaged separately for each of the four conditions.

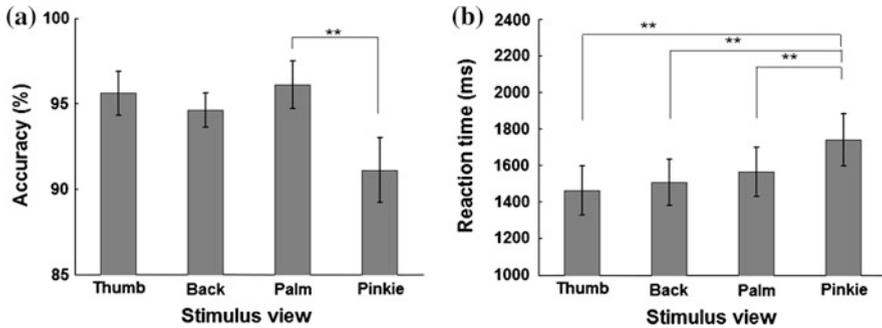
In the present study, there was no statistically significant difference between left- and right-hand stimuli, thus these data were pooled for further study. The view effect was studied by analyzing mean amplitude measures in the 400–550 ms time window, spanning the RRN component, selected according to visual inspection of grand average ERP waveforms. A  $4 \times 6 \times 3$  repeated-measures ANOVA was performed on the mean amplitudes at 18 electrodes (F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4, P3, Pz, P4, PO3, POz, and PO4), tasking as factors view (thumb, back, palm, and pinkie), region (frontal, fronto-central, central, central-parietal, parietal, and occipito-parietal), and laterality (left, middle, and right). Statistical analyses were carried out by the same way described for behavioral data.

To investigate the relationship between reaction times and RRN amplitudes, we conducted a linear regression analysis. Two-tailed Pearson correlation was implemented to compare the resulting regression slope coefficients across participants.

## 3 Results

### 3.1 Behavioral Data

All participants engaged in the mental rotation task of hands, with a mean accuracy of 94.37 %. A one-way ANOVA on accuracy was conducted and a marginal main effect of view was observed [mean  $\pm$  standard error of the mean (SEM);  $F(3,39) = 3.23$ ,  $P = 0.07$ ]. Pairwise comparisons revealed significant mean differences between the palm and pinkie condition ( $P < 0.01$ ). In contrast to accuracy, reaction times revealed a significant effect of view [mean  $\pm$  SEM;  $F(3,39) = 14.27$ ,  $P < 0.01$ ; thumb,  $1462 \pm 134$  ms; back,  $1508 \pm 126$  ms; palm,



**Fig. 3** **a** Accuracy and **b** reaction times as a function of stimulus viewpoint. The sign “\*\*” indicates significance  $P < 0.01$ . Error bars represent standard error of the mean

1563  $\pm$  135 ms; pinkie, 1740  $\pm$  143 ms]. Pairwise planned comparisons confirmed that hands presented in pinkie-side views took significantly longer RTs than those presented in more usual perspectives, such as thumb, back, or palm views (all  $P < 0.01$ ), but there were no differences among thumb, back, and palm views, as shown in Fig. 3.

### 3.2 ERP Data

Figure 4 shows the grand average ERP for each view at Fz, FCz, Cz, CPz, and Pz sites. As can be seen in Fig. 4, the rotation-related negativity component becomes more negative for stimuli depicting hands in unnatural views. That is, views difficult to reach with a real movement. Figure 5 shows the mean amplitude in the 400–550 ms time window as a function of stimulus view at Fz, FCz, Cz, CPz, Pz, and POz sites. The plots show again that response to pinkie-side views are more negative than the others views. Voltage maps in Fig. 6 show the spatial distribution of the view effect over all electrodes at the scalp surface in the 400–550 ms time window. These topographic maps show that the greatest mean amplitude was found at the CPz site.

The statistical analysis performed on the average ERP amplitude from 400 to 550 ms supports these observations. Three-way repeated-measures ANOVA showed that a significant main effect of view [mean  $\pm$  SEM;  $F(3,39) = 8.79$ ,  $P < 0.01$ ; thumb, 1.01  $\pm$  0.17  $\mu$ V; back, 1.10  $\pm$  0.13  $\mu$ V; palm, 1.10  $\pm$  0.14  $\mu$ V; pinkie, 0.83  $\pm$  0.15  $\mu$ V]. Pairwise planned comparisons revealed that pinkie-side views were distinctly different from back and palm views (all  $P < 0.01$ ) and marginally different between pinkie-side views and thumb views ( $P = 0.09$ ). Specially, the amplitude decreased with increasingly awkward/uncomfortable views. There were also significant effects of region [mean  $\pm$  SEM;  $F(5,65) = 5.24$ ,  $P < 0.05$ ], laterality [mean  $\pm$  SEM;  $F(2,26) = 4.53$ ,  $P < 0.05$ ], and region  $\times$  laterality [mean  $\pm$  SEM;  $F(10,130) = 3.55$ ,  $P < 0.01$ ], indicating

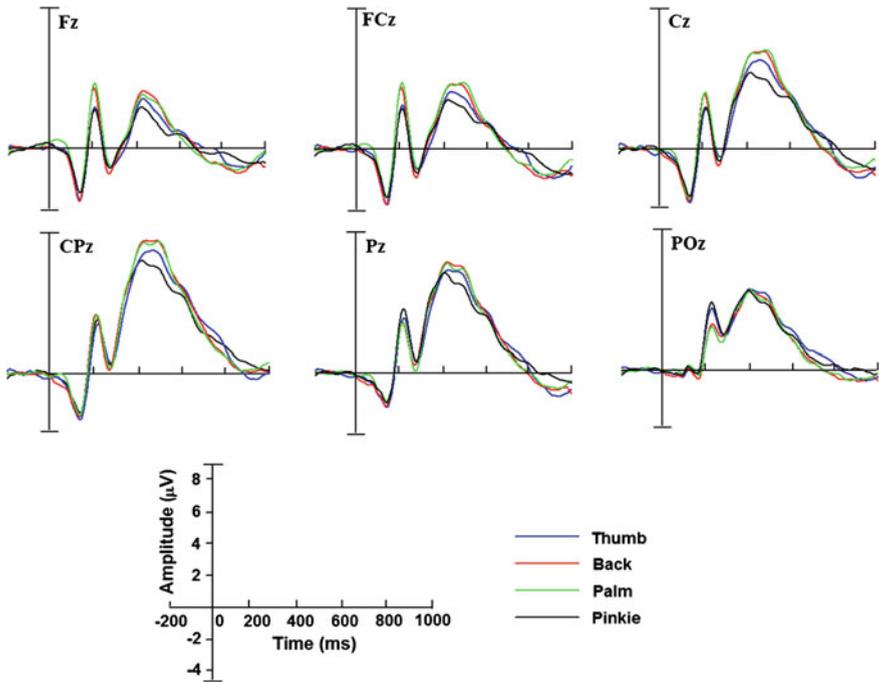


Fig. 4 Grand average ERP ( $n = 14$ ) in each view at the Fz, FCz, Cz, CPz, Pz, and POz electrodes

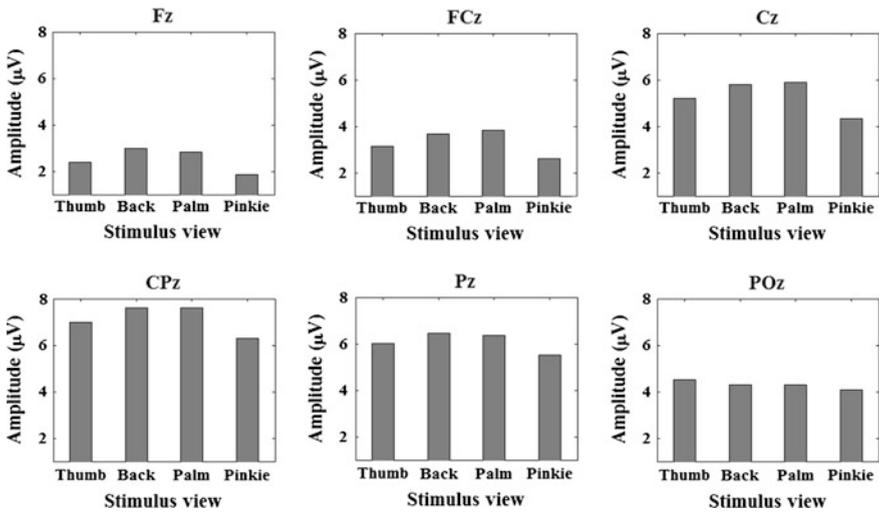
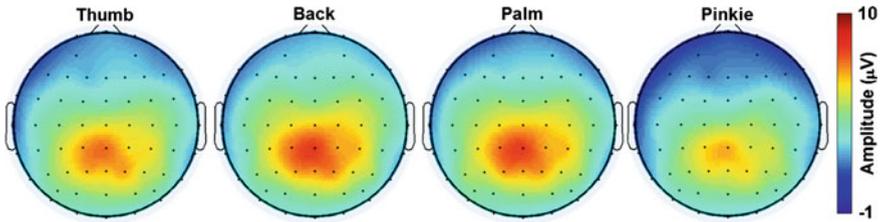


Fig. 5 Mean amplitude in the 400–550 ms time window as a function of stimulus view at the Fz, FCz, Cz, CPz, Pz, and POz sites



**Fig. 6** Topographical maps of the view effect based on the average ERP amplitude from 400 to 550 ms

**Table 1** The view effects reflected by the mean amplitudes of RRN in each of the six regions

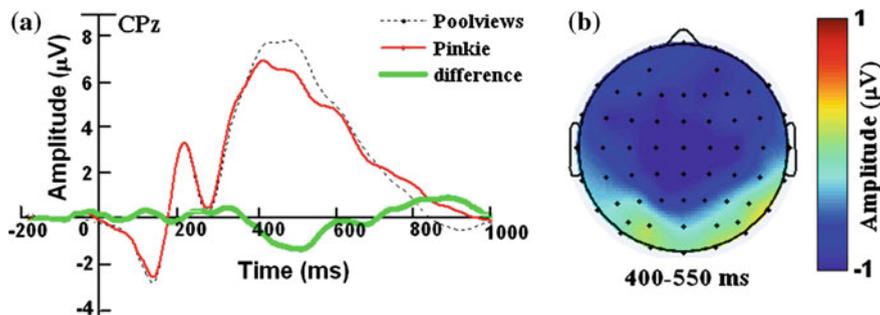
| Region            | RRN      |          |
|-------------------|----------|----------|
|                   | <i>F</i> | <i>P</i> |
| Frontal           | 9.72     | 0.002    |
| Frontal-central   | 13.02    | 0.001    |
| Central           | 23.06    | <0.001   |
| Central-parietal  | 13.16    | 0.001    |
| Parietal          | 3.12     | 0.070    |
| Occipito-parietal | 1.19     | 0.358    |

the topographic specificity of the RRN component. The interaction of region  $\times$  view [mean  $\pm$  SEM;  $F(15,195) = 3.51$ ,  $P < 0.05$ ] was also significant. Pairwise planned comparisons showed that the largest effect size was in the central region, suggesting the most significant view effect in this region (Table 1). However, no other significant interactions were obtained.

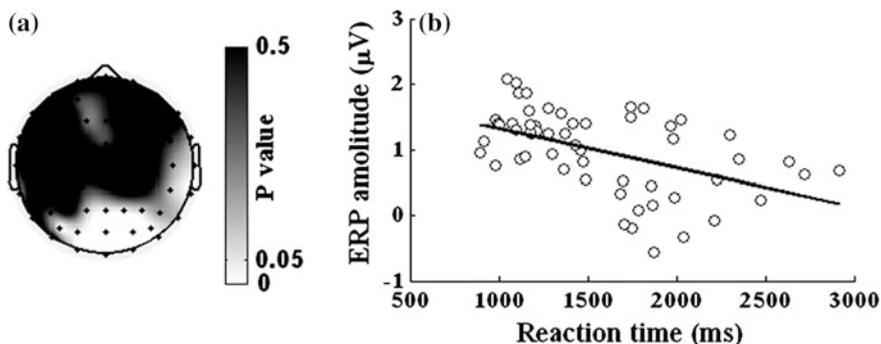
The three-way repeated-measures ANOVA was performed for these analyses. The significant level was set at  $P < 0.05$ .

Since pairwise planned comparisons revealed no significant difference between the mean amplitude with the thumb, back, and palm views. These results were pooled (Poolviews) for comparison with pinkie views. A significant reduction in mean amplitude was observed in the pinkie views (Fig. 7a) and topographic maps of difference wave Poolviews versus pinkie (Fig. 7b) showed the following spatio-temporal distribution for the ERP data. The amplitude was more negative in pinkie than in Poolviews in the 400–550 ms time window over the central region.

The mean values of strongest correlations between RRN amplitudes and reaction times for each electrode are depicted in Fig. 8a. The strongest and statistically significant correlations were obtained only for POz site. In Fig. 8b, a scatter plot illustrating the relationship between RRN amplitudes change and reaction times for participants during hand mental rotation task. Change in RRN amplitudes showed a significant negative correlation with reaction times [Pearson correlation coefficient ( $r$ ) =  $-0.49$ ;  $P < 0.01$ ], which indicated that more negative RRN amplitude was associated with slower responses to hand stimuli.



**Fig. 7** **a** Grand average ERP to Poolviews (average of thumb, back, and palm) and pinkie conditions and the difference wave (pinkie-Poolviews) at Cz site. **b** Topographical maps of the voltage amplitudes for the Poolviews versus pinkie condition difference wave in the 400–550 ms time window



**Fig. 8** **a** The *P* value at each electrode, as assessed the statistical significance by a linear regression analysis between reaction times and ERP amplitudes from 400 to 550 ms. **b** Association of individual mean ERP amplitudes from 400 to 550 ms and reaction times for stimulus view (thumb, back, palm, and pinkie) at the POz site

## 4 Discussion

The primary purpose of the present study was to further understand the process of hand mental rotation task. We investigated reaction times and event-related potential for the neural dynamics of mental hand rotation task by using EEG.

### 4.1 Behavioral Response

For all participants with low error rates, it was found that significantly longer reaction times were taken for pinkie-side views compared to the other three

conditions, which was in line with previous studies [22] (see Fig. 1). It can be concluded that the pinkie-side views were the most biomechanically difficult mental spatial transformations. This modulation of reaction times has been linked to the increase of biomechanical constraints when imaging the hand rotation to match the difficult stimulus view. Therefore, the biomechanical constraints might be primarily responsible for the greater mental simulation times and play an important role in stimuli-related cognitive processing. In addition, it has been reported that this cognitive processing can be affected by both current and desired states [24]. In a hand mental rotation task, the desired state is the presented stimulus view. Because motor imagery is subject to biomechanical constraints, reaction times are increased to judge the stimuli oriented in anatomically difficult positions if participant uses the motor imagery strategy to solve this task. In an informal debriefing following the experimental session, all participants reported to have solved the task by using a set of mental transformations of their own hands in order to closely match the view presented in the line drawing (motor imagery strategy). These behavioral results and participants' debriefings indicated the engagement in motor imagery.

#### ***4.2 ERP Amplitude Modulation Association with Mental Rotation***

Our ERP data showed that this mental rotation task had its main impact on neural activity during the RRN segment. In contrast to natural views such as thumb, back, and palm, the pinkie views elicited more negative component. This slow ERP component was functionally related to the P3b component of the P300 complex, which was generally interpreted as reflecting quality of information processing and the amount of effort allocated to a task. Therefore, we conclude that it takes more effort for the mental rotation in a difficult view than that in an easier one.

The RRN showed a maximal correlation with reaction times at parieto-occipital sites. Increased reaction time was associated with reduced ERP amplitude over parieto-occipital regions. Therefore, we can conclude that this effect may be an index of the increase in posterior parietal/occipital cortex activity. This notion is in accord with numerous previous studies that have reported that a large region of activation centered in the intraparietal sulcus and extending into the transverse occipital sulcus during mental rotation task [13, 25, 26]. Participants were slower and less accurate when making a 'same' judgment for a difficult view in comparison to an easy one. Therefore, the tendency of RRN amplitude in the parieto-occipital region to be negatively correlated with the reaction time might also indicate that attention to the mental rotation task is more demanding in a difficult view than in an easy one.

### ***4.3 Role of Central Motor Regions in Mental Rotation***

There is growing evidence that the parietal lobe play an important role in mental rotation. A variety of studies using PET, fMRI, and EEG have shown parietal lobe activation during mental rotation [13, 14, 19]. Furthermore, as task demands are increased, parietal lobe activation increases. In addition to activating the parietal lobe, there is also considerable evidence reporting activity in the precentral sulcus during mental rotation tasks [18].

In the current study we observed view-sensitive amplitude modulations and the most significant view effect in the central region. Thus, we conclude that mental rotation of hands is based on biomechanical constraints related to motor simulation and the central region is indeed involved in mental rotation of hands. Regions in the central cortex are associated with motor planning and motor execution. Motor imagery shares many cognitive aspects of movement with actual motor execution. These results suggest that the central lobes activity may reflect the use of action simulation to solve this task. In other words, participants use a set of mental transformations of their own hands in order to closely match the view presented in the line drawing. Evidence from PET and fMRI has provided support for the notion that when participants use the motor imagery strategy to solve the mental rotation tasks, medial/superior motor regions are activated [13, 14].

## **5 Conclusion**

We employed a mental rotation task of hands which is similar to that previously adopted to involve in motor imagery [22]. The behavioral data showed that reaction times were subject to biomechanical constraints because longer reaction times were found for those views when hands are not able to reach easily with real movement. ERP results showed that the RRN amplitude was more negative for difficult to maintain or unfamiliar views with the most significant view effect in the central region. Behavior was closely linked with the slope of the parieto-occipital ERP amplitude modulation. These results suggest that the neural activity underlying mental rotation of hands are primarily conducted in parietal, occipital, and central lobes.

**Acknowledgments** This work is supported by National Basic Research Program (973) of China (No. 2011CB933204), National Natural Science Foundation of China under Grant 90820304, 91120007, 61105123, Chinese 863 Project: 2012AA011601.

## References

1. Shepard RN, Metzler J (1971) Mental rotation of three-dimensional objects. *Science* 171:701–703
2. Prime DJ, Jolicoeur P (2010) Mental rotation requires visual short-term memory: evidence from human electric cortical activity. *J Cogn Neurosci* 22:2437–2446
3. Milivojevic B, Hamm JP, Corballis MC (2009) Functional neuroanatomy of mental rotation. *J Cogn Neurosci* 21:945–959
4. Quinn PC, Liben LS (2008) A sex difference in mental rotation in young infants. *Psychol Sci* 19:1067–1070
5. Peters M, Battista C (2008) Application of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. *Brain Cogn* 66:260–264
6. Chu M, Kita S (2008) Spontaneous gestures during mental rotation tasks: insights into the microdevelopment of the motor strategy. *J Exp Psychol Gen* 137:706–723
7. Parsons LM (1987) Imagined spatial transformations of one's hands and feet. *Cognitive Psychol* 19:178–241
8. Ionta S, Blanke O (2009) Differential influence of hands posture on mental rotation of hands and feet in left and right handers. *Exp Brain Res* 195:207–217
9. Petit LS, Harris IM (2005) Anatomical limitations in mental transformations of body parts. *Vis Cogn* 12:737–758
10. Mourao-Miranda J, Ecker C, Sato JR, Brammer M (2009) Dynamic changes in the mental rotation network revealed by pattern recognition analysis of fMRI data. *J Cognitive Neurosci* 21:890–904
11. Vingerhoets G, Santens P, Laere KV, Lahorte P, Dierckx RA, Reuck JD (2001) Regional brain activity during different paradigms of mental rotation in healthy volunteers: a positron emission tomography study. *Neuroimage* 13:383–391
12. Shimoda N, Takeda K, Imai I, Kaneko J, Kato H (2008) Cerebral laterality differences in handedness: a mental rotation study with NIRS. *Neurosci Lett* 430:43–47
13. Kosslyn SM, Digirolamo GJ, Alpert NM, Thompson WL (1998) Mental rotation of objects versus hands: neural mechanisms revealed by positron emission topography. *Psychophysiology* 35:151–161
14. Vingerhoets G, de Lange FP, Vandemaële P, Deblaere K, Achten E (2002) Motor imagery in mental rotation: an fMRI study. *Neuroimage* 17:1623–1633
15. de Lange FP, Helmich RC, Toni I (2006) Posture influences motor imagery: an fMRI study. *Neuroimage* 33:609–617
16. Zacks JM (2008) Neuroimaging studies of mental rotation: a meta-analysis and review. *J Cognitive Neurosci* 20:1–19
17. Beste C, Heil M, Konrad C (2010) Individual differences in ERPs during mental rotation of characters: lateralization, and performance level. *Brain Cogn* 72:238–243
18. Pelgrims B, Michaux N, Olivier E, Andres M (2011) Contribution of the primary motor cortex to motor imagery: a subthreshold TMS study. *Hum Brain Mapp* 32:1471–1482
19. Núñez-Peña MI, Aznar-Casanova JA (2009) Mental rotation of mirrored letters: evidence from event-related brain potentials. *Brain Cogn* 69:180–187
20. Makeig S, Delorme A, Westerfield M, Jung TP, Townsend J, Courchesne E, Sejnowski TJ (2004) Electroencephalographic brain dynamics following manually responded visual targets. *PLoS Biol* 2:e176
21. Wijers AA, Otten LJ, Feenstra S, Mulder G, Mulder LJM (1989) Brain potentials during selective attention, Memory search, and mental rotation. *Psychophysiology* 26:452–467
22. Nico D, Daprati E, Rigal F, Parsons L, Sirigu A (2004) Left and right recognition in upper limb amputees. *Brain* 127:120–132
23. Semlitsch HV, Anderer P, Schuster P, Presslich O (1986) A solution for reliable and valid reduction of ocular artifacts, Applied to the P300 ERP. *Psychophysiology* 23:695–703

24. de Lange FP, Roelofs K, Toni I (2008) Motor imagery: a window into the mechanisms and alterations of the motor system. *Cortex* 44:494–506
25. Gogos A, Gavrilescu M, Davison S, Searle K, Adams J, Rossell SL, Bell R, Davis SR, Egan GF (2010) Greater superior than inferior parietal lobe activation with increasing rotation angle during mental rotation: an fMRI study. *Neuropsychologia* 48:529–535
26. Koscik T, Leary DO, Moser DJ, Andreasen NC, Nopoulos P (2009) Sex differences in parietal lobe morphology: relationship to mental rotation performance. *Brain Cogn* 69:451–459

# Predictive Coding with Context as a Model of Image Saliency Map

Duzhen Zhang and Chuancai Liu

**Abstract** Predictive coding/biased competition (PC/BC) is a computational model of primary visual cortex (V1). Recent literature demonstrates that PC/BC model provides an implementation of the V1 bottom-up saliency map hypothesis. In this paper, we propose a novel approach toward natural color images saliency detection via the PC/BC model with top-down cortical feedback as context. We compare our method with the five state-of-the-art models of saliency detectors. Experimental results show that our method performs competitively for visual saliency detection task.

**Keywords** Saliency map · PC/BC model · Primary visual cortex (V1) · Top-down · Bottom-up · Context

## 1 Introduction

The visual system pays attention to the salient object. A number of psychophysical experiments suggest that primary visual cortex (V1) may be involved in the computation of visual salience. Spratling introduced the nonlinear predictive coding/biased competition (PC/BC) model [1], a reformulation of predictive

---

D. Zhang (✉) · C. Liu  
School of Computer Science and Engineering, Nanjing University of Science and Technology, NJUST, Nanjing, China  
e-mail: zhduzhen@yahoo.cn

C. Liu  
e-mail: liu.ccnj@yahoo.com.cn

D. Zhang  
School of Computer Science and Technology, Jiangsu Normal University, JSNU, Xuzhou, China

coding consistent with the biased competition theory of attention, that can simulate a very wide range of V1 response properties including tuning and suppression [2, 3]. The paper [4] extends his previous work by showing that the PC/BC model of V1 can also simulate a wide range of psychophysical experiments on visual saliency, and demonstrates that PC/BC provides a possible implementation of the V1 bottom-up saliency map hypothesis. It proposes that the perceptual saliency of the image is consistent with the relative strength of the prediction error calculated by PC/BC. Saliency can therefore be interpreted as a mechanism by which prediction errors attract attention in an attempt to improve the accuracy of the brain's internal representation of the world [4].

Visual saliency plays important roles in natural vision in that saliency can direct eye movements, deploy attention, and facilitate tasks like object detection and scene understanding. Many models have been built to compute saliency map. There are two major categories of factors that drive attention: bottom-up factors and top-down factors [5]. Bottom-up factors are derived solely from the visual scene. Regions of interest that attract our attention are in a bottom-up way and the responsible feature for this reaction must be sufficiently discriminative with respect to surrounding features. Most computational models focused on bottom-up attention, where the subjects are free-viewing a scene and salient objects attract attention. Inspired by the feature-integration theory [6], Itti et al. [7] proposed one of the earliest bottom-up selective attention models by utilizing color, intensity, and orientation of images. Bruce et al. [8] introduced an idea of using Shannon's self-information to measure the perceptual saliency. Saliency using natural image statistics (SUN) is a bottom-up bayesian framework [9]. Recently, Hou et al. [10] proposed a dynamic visual attention approach to calculate the saliency map based on Incremental Coding Length (ICL). Bottom-up attention can be biased toward targets of interest by top-down cues such as object features, scene context and task-demands. Bottom-up and top-down factors should be combined to direct attentional behavior. A recent review of attention models from computational perspective can be found in [11].

Reference [4] uses synthetic stimuli to test the saliency of the PC/BC model. In this paper, inspired by the work of Spratling, we propose an approach toward natural color images saliency detection via the PC/BC model with top-down cortical feedback as context. We compare our method with the five state-of-the-art models of saliency detectors. Experimental results show that our method performs competitively for visual saliency detection task. The rest of this paper is organized as follows. [Section 2](#) introduces and analyzes Spratling's PC/BC model, and based on his work, a novel method combining top-down cortical feedback for measuring image saliency is proposed. Experimental results and comparisons with state-of-the-art models are presented in [Sect. 3](#), and discussions are given in [Sect. 4](#).

## 2 The Model Description

Figure 1 illustrates the retina/LGN model and the PC/BC model of V1, from left to right, capital characters I, X, E, Y, and A represent input image, image preprocessing stage by the retina/LGN, the error-detecting neurons, the prediction neurons, feedback from higher cortical regions, respectively.

### 2.1 The Retina/LGN Model

To simulate the effects of circular-symmetric center-surround receptive fields (RFs) in lateral geniculate nucleus (LGN) and retina, input image (I) preprocessed by convolution with a Laplacian-of-Gaussian (LoG) filter ( $l$ ) and a saturating nonlinearity:

$$X = \tanh\{2\pi(I * l)\}. \quad (1)$$

The positive and rectified negative responses were separated into two images  $X_{\text{ON}}$  and  $X_{\text{OFF}}$  simulating the outputs of cells in retina and LGN with on-center/off-surround and off-center/on-surround RFs, respectively. These ON- and OFF-channels provided the input to the PC/BC model of V1.

### 2.2 The V1 Model

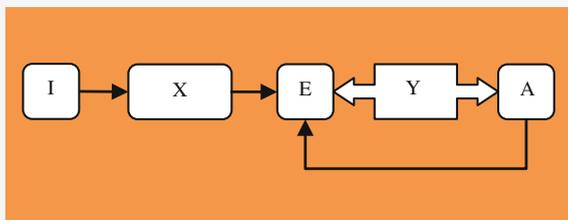
The PC/BC model of V1 is described by the following equations:

$$E_o = X_o \oslash \left( \varepsilon_2 + \sum_{k=1}^p (\hat{\omega}_{ok} * Y_k) \right). \quad (2)$$

$$Y_k \leftarrow (\varepsilon_1 + Y_k) \otimes \sum_o (\omega_{ok} \circ E_o). \quad (3)$$

$$Y_k \leftarrow Y_k \otimes (1 + \eta A_k). \quad (4)$$

**Fig. 1** The retina/LGN model and the PC/BC model of V1



where  $o \in [\text{ON}, \text{OFF}]$ ;  $X_o$  represents the input to the model of V1,  $E_o$  represents the error-detecting neuron responses,  $Y_k$  represents the prediction neuron responses,  $A_k$  represents the weighted sum of top-down predictions, all of them are two-dimensional array, equal in size to the input image;  $\omega_{ok}$  is a two-dimensional kernel representing the synaptic weights for a particular class ( $k$ ) of neuron normalized so that the sum of all the weights is equal to  $\psi$ ,  $\hat{\omega}_{ok}$  is a two-dimensional kernel representing the same synaptic weights as  $\omega_{ok}$  but normalized so that the maximum value is equal to  $\psi$ , the Gabor function is used to define the weights of each kernel  $\omega_{ok}$  and  $\hat{\omega}_{ok}$  (a family of 32 Gabor functions with eight orientation ( $0^\circ$ – $157.5^\circ$  in steps of  $22.5^\circ$ ) and four phases ( $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ , and  $270^\circ$ ) were used);  $p$  is the total number of kernels;  $\varepsilon_1$ ,  $\varepsilon_2$ ,  $\eta$  and  $\psi$  are parameters;  $\oslash$  and  $\otimes$  indicate element-wise division and multiplication, respectively;  $\circ$  represents cross-correlation (which is equivalent to convolution without the kernel being rotated  $180^\circ$ ); and  $*$  represents convolution (which is equivalent to cross-correlation with a kernel rotated  $180^\circ$ ). Parameter values  $\psi = 5000$ ,  $\varepsilon_1 = 0.0001$ ,  $\varepsilon_2 = 250$ , and  $\eta = 1$  were used in the experiments.

Equation (2) describes the calculation of the neural activity for each population of error-detecting neurons. The activation of the error-detecting neurons can be interpreted as representing the residual error between the input and the reconstruction of the input generated by the prediction neurons. The values of  $E$  are related to the image saliency, with high error values corresponding to high saliency.

Equation (3) describes the updating of the prediction neuron activations. The values of  $Y_k$  represent predictions of the causes underlying the inputs to the model of V1. If the input remains constant, the values of  $Y_k$  will converge to steady-state values that reconstruct the input with minimum error.

### 2.3 Modeling the Top-Down Effects

Equation (4) describes the effects on the V1 prediction neuron activations of top-down inputs from prediction neurons at later processing stages (i.e., in extra-striate cortical regions). In Eq. (4), the effects of cortical feedback are modeled by using an array of inputs ( $A$ ) to the V1 model which represents the weighted sum of top-down predictions. In the simulations of Ref. [4], feedback was either simple orientation preferences, the elements of  $A$  were set to values of 0.25 and zero, or assumed to be negligible, the elements of  $A$  were given a value of zero, in which cases Eq. (4) had no effect. We add the following equation between Eq. (3) and Eq. (4) to model the top-down effects:

$$A_k \leftarrow \sum_{k=1}^p (\hat{\omega}_{ok} * Y_k). \quad (5)$$

This top-down feedback will have two effects on the PC/BC model of V1. (1) Increasing the response of the prediction neurons that represent information consistent with the top-down expectation [see Eq. (4)]. This will result in these prediction neurons sending stronger feed-forward activation, and hence, make this information more conspicuous for cortical regions at subsequent stages along the processing hierarchy. (2) The enhanced activity in the prediction neurons consistent with top-down expectations will in turn decrease the response of the error-detecting neurons from which these prediction neurons receive their input [see Eq. (2)] [4]. Since the strength of the responses of the error-detecting neurons is assumed to be related to saliency, in this way, top-down feedback modulates bottom-up saliency.

### 3 Experimental Comparisons

#### 3.1 Saliency Results Comparison

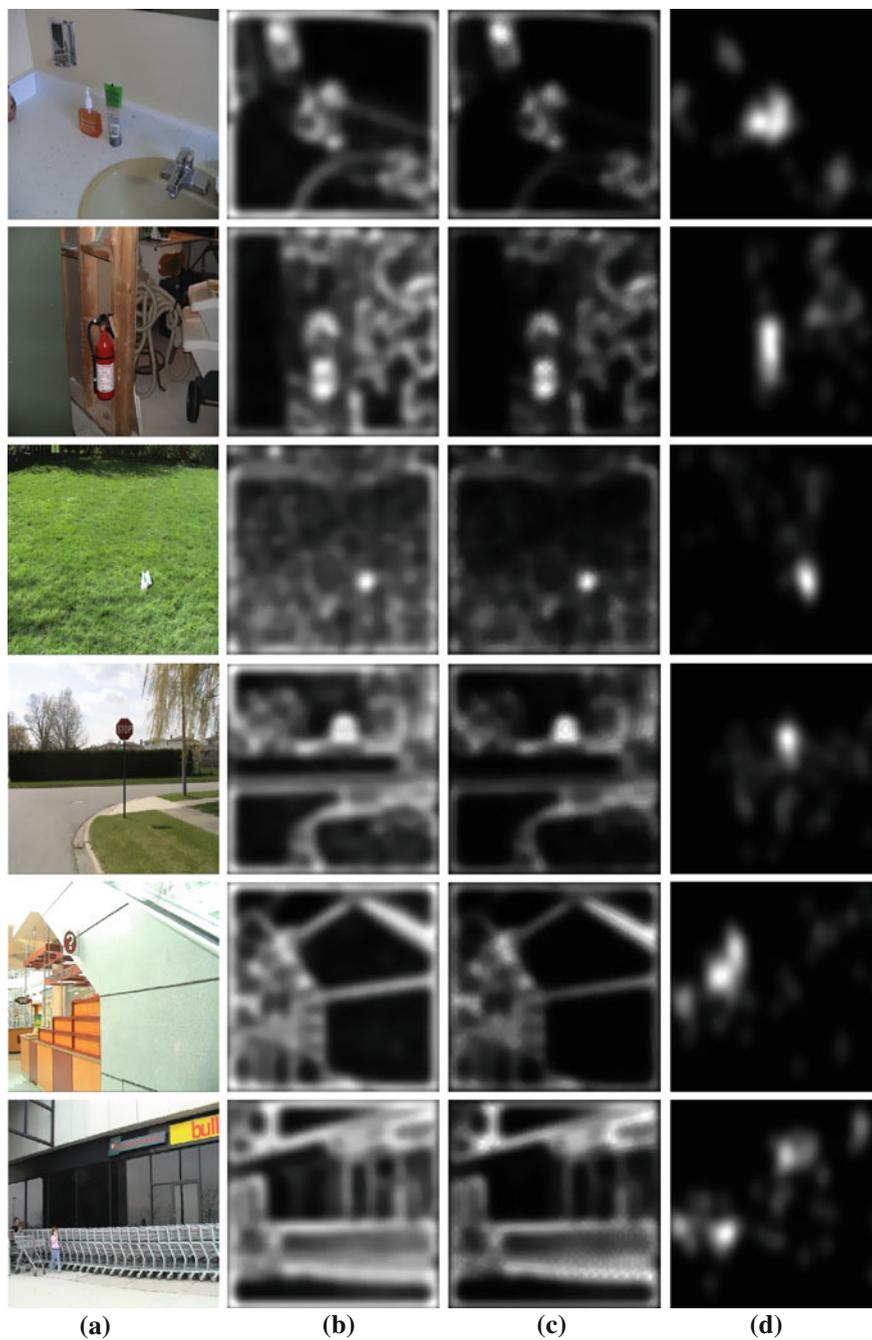
We evaluated our method on human visual fixation data from natural images. The dataset we used was collected by Bruce and Tsotsos [8] as the benchmark dataset for comparing human eye predictions between methods. The dataset contains eye fixation data from 20 subjects for a total of 120 natural images.

Figure 2 affords a qualitative comparison of the output of the proposed models (without/with context) for a variety of images. Visually, top-down effects increase the performance of salient object detection, i.e., top-down signals modulate bottom-up saliency. This is in line with preceding analysis. Figure 2d is fixation density map based on experimental human eye tracking data as the “ground truth” saliency map of each image.

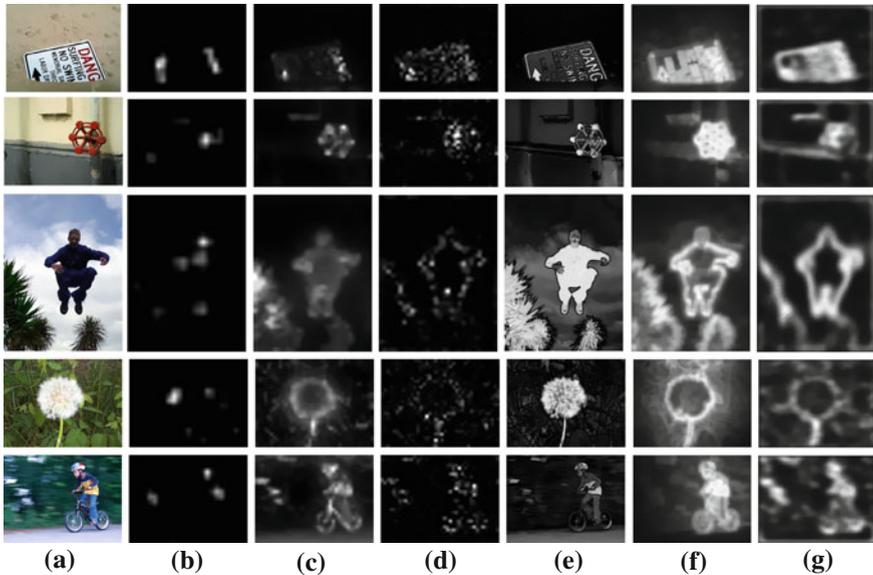
#### 3.2 Comparing Our Saliency Results with Other Methods

We compare our saliency method with context against other five state-of-the-art methods using the database from the publicly available database used by Achanta et al. [12]. Each of the 1,000 images in the database contains a salient object or a distinctive foreground object, so we can compare the performance of different algorithms.

The five saliency detectors are Itti et al. [7], Harel et al. [13], Hou and Zhang [14], Achanta [12], and Goferman et al. [15], hereby referred to as IT, GB, SR, IG, and CA. We refer to our proposed method as PC. The choice of these algorithms is motivated by the following reasons: citation in literature (the classic approach of IT is widely cited), recency (IG, and CA are recent), and variety (IT is biologically



**Fig. 2** Results for qualitative comparison: **a** Original image; **b** Saliency map without context; **c** Saliency map with context; **d** Fixation density map based on experimental human eye tracking data



**Fig. 3** Visual comparison of saliency map. **a** Original, **b** IT [7], **c** GB [13], **d** SR [14], **e** IG[12], **f** CA [15], **g** PC

motivated, CA is purely computational, GB is a hybrid approach, SR and IG estimates saliency in the frequency domain).

We randomly choose some images from the database. Figure 3 is the output of the five state-of-the-art methods and our method for comparison. Our method is a competitive, promising algorithm.

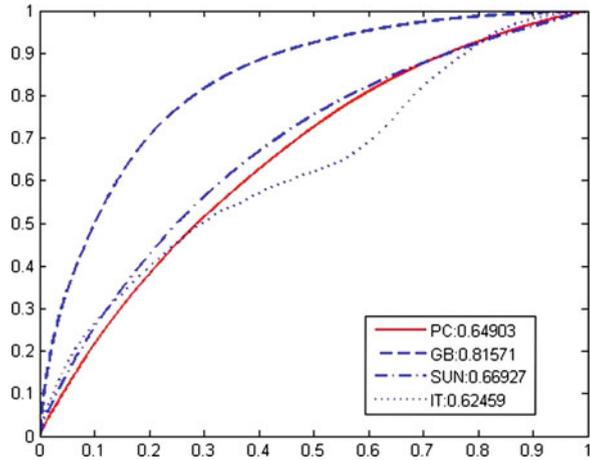
### 3.3 Quantitative Evaluation

To obtain a quantitative evaluation we compare ROC curves and Area Under Curve (AUC) on the database presented in [8]. Figure 4 is the result of our method and other three methods.

## 4 Discussions

PC/BC is a computational model of primary visual cortex (V1) which provides an implementation of the V1 bottom-up saliency map. In this paper, we propose a novel approach to natural color image saliency detection method with top-down cortical feedback as context. Our experimental result is consistent with recent literature conclusion: top-down signals modulate (override) bottom-up saliency (in

**Fig. 4** ROC curves for the database of [8]



a feature-specific way) [16]. We compare our method with the five state-of-the-art models of saliency detectors. Experimental results show that our method performs competitively for visual saliency detection task.

When the organism is not actively searching for a particular target (the free-viewing condition), the organism’s attention should be directed to the most salient points which potential targets in the visual field. Bottom-up attention mechanisms have been more thoroughly investigated than top-down mechanisms. One reason is that data-driven stimuli are easier to control than cognitive factors such as task-demands, knowledge, and expectations. Even less is known on the interaction between the two processes [17].

In future work, we will incorporate color feature and other task-demands features as context to detect saliency, “Combining such features-specific top-down signals with (learnt) contextual priors on target location therefore may provide a promising approach to searching for real-world objects in their natural context [16]”, and develop applications of our model.

**Acknowledgments** This work is supported by the National Natural Science Fund of China (Grant Nos. 60632050, 9082004) and by the basic key technology project of Ministry of Industry and Information Technology of China (Grant No. E0310/1112/JC01).

## References

1. Spratling MW (2008) Predictive coding as a model of biased competition in visual attention. *Vis Res* 48:1391–1408
2. Spratling MW (2010) Predictive coding as a model of response properties in cortical area V1. *J Neurosci* 30:3531–3543
3. Spratling MW (2011) A single functional model accounts for the distinct properties of suppression in cortical area V1. *Vis Res* 51:563–576

4. Spratling MW (2011) Predictive coding as a model of the V1 saliency map hypothesis. *Neural Networks* 20:1–22
5. Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Ann Rev Neurosci* 18:193–222
6. Treisman A, Gelade G (1980) A feature-integration theory of attention. *Cogn Psychol* 12:97–136
7. Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans Pattern Anal Mach Intell* 20:1254–1259
8. Bruce N, Tsotsos J (2005) Saliency based on information maximization. In: *Proceedings in advances in neural information processing systems*, vol 18, pp 155–162
9. Zhang L, Tong MH, Marks TK, Shan H, Cottrell GW (2008) SUN: a Bayesian framework for saliency using natural statistics. *J Vis* 8:1–20
10. Hou X, Zhang L (2008) Dynamic visual attention: searching for coding length increments. In: *Proceedings of advances in neural information processing systems*, pp 681–688
11. Borji A, Itti L (2012) State-of-the-art in visual attention modeling. <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2012.89>
12. Achanta R, Hemami S, Estrada F, Süsstrunk S (2009) Frequency-tuned salient region detection. In: *Proceedings of IEEE international conference on computer vision and pattern recognition*, pp 1597–1604
13. Harel J, Koch C, Perona P (2007) Graph-based visual saliency. *Adv Neural Inf Proc Syst* 19:545–552
14. Hou X, Zhang L (2007) Saliency detection: a spectral residual approach. In: *Proceedings of IEEE conference on computer vision and pattern recognition*, pp 18–23
15. Goferman S, Zelnik-Manor L, Tal A (2010) Context-aware saliency detection. In: *Proceedings of IEEE international conference on computer vision and pattern recognition*, pp 2376–2383
16. Einhäuser W, Rutishauser U, Koch C (2008) Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *J Vis* 8:1–19
17. Frintrop S, Rome E, Christensen HI (2010) Computational visual attention systems and their cognitive foundations: a survey. *ACM Trans Appl Percept* 7:1–46

# Multiclass Pattern Analysis of Whole-Brain Functional Connectivity of Schizophrenia and Their Healthy Siblings

Yang Yu, Hui Shen, Ling-Li Zeng and Dewen Hu

**Abstract** Recently, a growing number of neuroimaging studies have begun to pay attention to exploring the brains of schizophrenic patients to identify heritable biomarkers for this disorder involving their healthy siblings. Based on whole-brain resting-state functional connectivity of schizophrenic patients, their healthy siblings and healthy controls, the objective of the present study aimed to use multiclass pattern analysis to reveal three types of neural signature: (i) state connectivity patterns, reflecting the state of having schizophrenia; (ii) trait connectivity patterns, reflecting the genetic vulnerability to develop schizophrenia; and (iii) compensatory connectivity patterns, underlying special brain connections by which healthy siblings compensate for an increased genetic risk for developing schizophrenia. The current study may provide additional insights into the pathophysiological mechanisms underlying schizophrenia and be helpful in further highlighting genetic contribution to the etiology of schizophrenia.

**Keywords** Schizophrenia · Healthy siblings · Functional magnetic resonance imaging · Resting-state · Functional connectivity · Multiclass pattern analysis

---

Y. Yu · H. Shen · L.-L. Zeng · D. Hu (✉)  
College of Mechatronics and Automation, National University of Defense Technology,  
Changsha 410073 Hunan, China  
e-mail: dwhu@nudt.edu.cn

Y. Yu  
e-mail: yy87724@sina.com

H. Shen  
e-mail: shenhui\_nudt@126.com

L.-L. Zeng  
e-mail: losernudt@163.com

## 1 Introduction

Schizophrenia is a highly heritable psychiatric disorder, and evidence from family and twin studies demonstrated that heritable factors play an important role in the pathogenesis of schizophrenia [1–4]. It was suggested that the similar genetic backgrounds between schizophrenic patients and their healthy siblings contribute to approximately ninefold higher risk for siblings to develop schizophrenia than the general population [5, 6], whereas the healthy siblings of schizophrenic patients behave normally in their daily lives and do not actually meet the diagnostic criteria for schizophrenic disorder. Thus, comparison among schizophrenic patients, healthy siblings of schizophrenic patients, and healthy controls likely provides additional insights into the pathophysiological mechanisms underlying this disorder and may be helpful in further highlighting genetic contribution to the etiology of schizophrenia.

In recent years, a growing number of neuroimaging studies have begun to pay attention to exploring the brains of schizophrenic patients to identify heritable biomarkers for this disorder involving their healthy siblings [7–10]. Increasing rs-fcMRI studies have demonstrated altered functional connectivity in the brain of schizophrenic patients and their healthy siblings [11–15]. However, most of these previous studies focused mainly on the functional connectivities associated with one or a few preselected seed regions of interest (ROIs), ignoring other potentially informative connectivity patterns. Or they focused on group-level statistical methods which are less helpful to clinical diagnosis on individual-level. In this study, we investigated the whole-brain functional connectivity patterns of schizophrenic patients, healthy siblings, and healthy controls, which could make full use of information contained in the neuroimaging data.

Recent years, multivariate pattern analyses based on machine learning have aroused great interest for its capacity of extracting functional patterns from neuroimaging data and its potential capability of finding valuable neuroimaging-based biomarkers [16]. Our recent study [17] based on multivariate pattern analysis has applied nonlinear support vector machine (SVM) to whole-brain rs-fMRI data to train a one-against-one classifier between schizophrenic patients and their healthy siblings, resulting in significant separation. However, one limitation of this work is that reconstruction is unavailable for the use of nonlinear SVM classifiers, so that it is difficult to discover informative functional connectivity patterns revealing the pathophysiological mechanisms of schizophrenia.

In the present study, we used a multiclass pattern classifier to explore the whole-brain resting-state functional connectivity patterns of schizophrenic patients, healthy siblings and healthy controls based on rs-fMRI. Instead of using a one-against-one classifier, we established our one-against-rest classifier for addressing this multiclass pattern classification issue. There were three procedures involved in this study: (i) dimensional reduction of the original feature space using principal component analysis (PCA), (ii) one-against-rest classification based on

linear SVM was trained to solve the multiclassification problems, and (iii) reconstruction to identify features with high discriminative power.

## 2 Materials and Methods

### 2.1 Participants

Subjects consisted of 24 schizophrenic patients, 25 healthy siblings, and 22 healthy controls. Schizophrenic patients were recruited from outpatient departments and inpatient units at the Department of Psychiatry, Second Xiangya Hospital of Central South University, and all the patients were evaluated by the Structured Clinical Interview for DSM-IV, and fulfilled the criteria for schizophrenia according to DSM-IV. Symptom severity for patients was assessed using the positive and negative syndrome scale [18]. No patients had a history of neurological disorders, severe medical disorder, substance abuse, or electroconvulsive therapy. Six of the schizophrenic patients were medication-free, while the others accepted atypical psychotropic drugs during time of scanning (risperidone [ $n = 10$ , 2–6 mg/day], clozapine [ $n = 4$ , 200–350 mg/day], quetiapine [ $n = 4$ , 400–600 mg/day], and sulpiride [ $n = 1$ , 200 mg/day]). Twenty-five healthy siblings of schizophrenia were recruited so that each schizophrenic patient had a corresponding sibling, and they do not fulfill the DSM-IV criteria for any Axis-I psychiatric disorders. Twenty-five healthy controls who had no first-degree relatives with a history of psychiatric disorders were recruited from Changsha City, China. The schizophrenic patients, healthy siblings, and the healthy controls were well demographically similar on age, gender, and education levels.

All participants gave their written informed consents to participate in the study and they were studied under protocols approved by the Second Xiangya Hospital of Central South University.

### 2.2 Resting Experiment and Data Acquisition

MRI scans were performed with a 1.5 T GE Signa System (GE Signa, Milwaukee, Wisconsin, USA) via using a gradient-echo EPI sequence. The imaging parameters are as follows: TR = 2000 ms, TE = 40 ms, FOV = 24 cm, FA = 90°, matrix = 64 × 64, slice thickness = 5 mm, gap = 1 mm, slices = 20. In the experiment, subjects were instructed simply to keep their eyes closed, to relax, remain awake, and perform no specific cognitive exercise. Foam pads and earplugs were used to minimize head motion and scanner noise, respectively. Each functional resting-state session lasted 6 min, resulting in 180 volumes.

### ***2.3 Data Preprocessing***

Image preprocessing is performed for all rs-fMRI images using the statistical parametric mapping software package (SPM8, Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK, <http://www.fil.ion.ucl.ac.uk/spm>). For each subject, the first 5 volumes of the scanned data were discarded for magnetic saturation effects. The remaining volumes were corrected by registering and reslicing for head movement. All subjects in this study had less than 1.5 mm translation and/or 2° of rotation in each axis. After the head motion effect was removed, the volumes were normalized to the standard EPI template in the Montreal Neurological Institute (MNI) space. The resulting images were then spatially smoothed with a Gaussian filter of 8 mm full-width half-maximum kernel to increase signal-to-noise ratio and temporally filtered with a Chebyshev band-pass filter (0.01–0.08 Hz). All the registered fMRI volumes with the MNI template were further divided into 116 regions according to the anatomically labeled template previously validated and reported by Tzourio-Mazoyer et al. [19].

Regional mean time series were acquired for each individual by averaging the fMRI time series over all voxels in each of the 116 regions. For each regional mean time series, we further regressed out the global mean signals and the effects of translations and rotations of the head estimated in the course of initial movement correction by image realignment. The residuals of the above regressions constituted the set of regional mean time series used for functional connectivity analysis [20]. We then calculated the Pearson’s correlation coefficients between the time series of all ROIs, resulting in a 6,670 dimensional feature vector.

### ***2.4 Multiclass Pattern Analysis***

Before training the multiclass classifier, PCA was applied to reduce the dimensionality of original feature space [21]. We used one-against-rest strategy in designing our classifiers. A leave-one-out cross-validation (Loocv) strategy was used to estimate the generalization ability of our classifier [13]. Statistical significance of the classification accuracy was determined by permutation test [22], involving repeating the classification procedure 1,000 times with a different random permutation of the classification labels of the training data.

### ***2.5 Identification of Features with High Discriminative Power***

In this study, we determined the functional connectivity features with the highest discriminative power by reconstruction based on the performance of each one-

against-rest classifiers. Because each feature influenced the classification via its weight, the larger the absolute magnitude of a feature’s weight is, the stronger it affected the classification result. For every one-to-rest classifier, we obtained a weight vector in each Loocv experiment. The weight vector for the one-against-rest classifier was finally acquired by averaging these above weight vectors. We therefore obtained three weight vectors respectively representing the features’ discriminative power for each one-against-rest classifiers. Because we performed the classification in the dimension-reduced subspace, in order to determine the original functional connectivities which have significantly contribution to the classification, we used the method detailed in previous study [23] to map back each weight vector to the original high-dimensional space. Thus, for all the 6,670 resting-state functional connectivities, we obtained the order of their contribution to the classification for each one-against-rest classifier.

Similar to the definition in the previous study [24], we defined three types of neural signature patterns: (i) state connectivity pattern, functional connectivity patterns responding to the state of having schizophrenia which potentially underlying the nature of abnormal in brain network of schizophrenic patient; (ii) trait connectivity pattern, functional connectivity patterns of dysfunction shared by schizophrenic patients and their healthy siblings, providing a possible neuroendophenotype to help to bridge genomic complexity and disorder heterogeneity and thereby accelerate the investigate of pathophysiological mechanism; and (iii) compensatory connectivity pattern, unique to healthy siblings, underlying special brain connectivities by which healthy siblings might compensate for an increased genetic risk for developing schizophrenia.

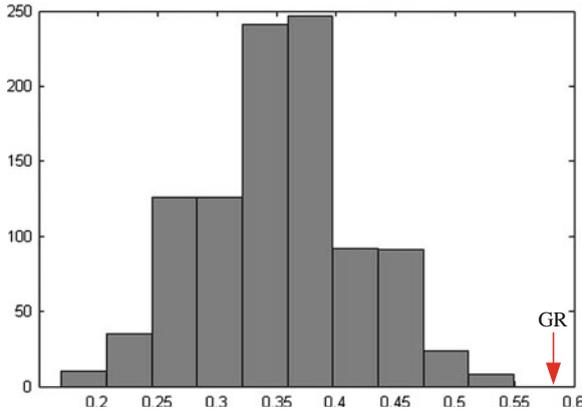
## 3 Results

### 3.1 Classification Results

The classification results indicated that the correct classifications of schizophrenic patients, healthy siblings, and healthy controls were respectively 62.5, 48.0, and 63.64 % (details were shown in Table 1) and the generalization rate (GR) achieved an accuracy of 57.75 %. Permutation test results ( $p < 0.001$ ) indicated that the overall classification accuracy was statistically significance. The permutation distribution of the estimate is shown in Fig. 1.

**Table 1** Confusion matrix for results in leave-one-out cross-validation

| Classes          | Schizophrenia (%) | Healthy siblings (%) | Healthy controls (%) |
|------------------|-------------------|----------------------|----------------------|
| Schizophrenia    | 62.5              | 12.5                 | 25.0                 |
| Healthy siblings | 16.0              | 48.0                 | 36.0                 |
| Healthy controls | 13.7              | 22.7                 | 63.6                 |



**Fig. 1** Histogram of the permutation distribution of the estimate (repetition times: 1,000). x-label and y-label represent the generalization rate and occurrence number, respectively. GR is the generation rate obtained by the classifier trained on the real class labels. With the generalization rate as the statistic, this figure reveals that the classifier learned the relationship between the data and the labels with a probability of being wrong of  $<0.001$

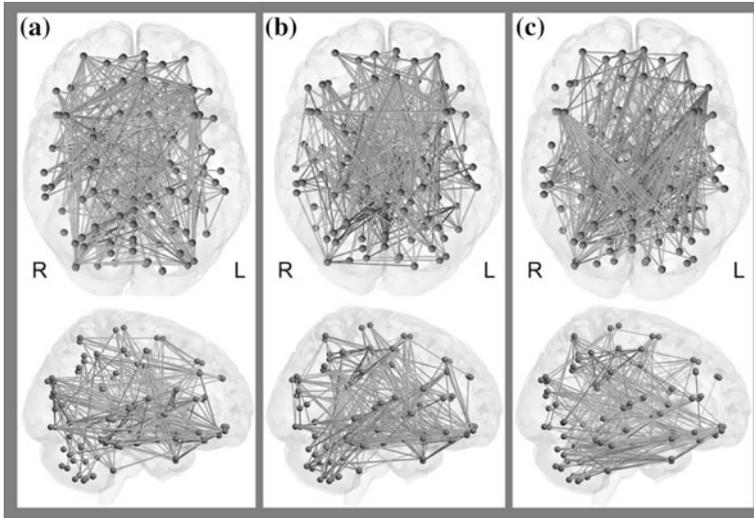
### 3.2 Functional Connectivity with High Discriminative Power

To identifying the three kinds of brain neural signature patterns: state, trait, compensatory, we selected 5 % (330) of the total functional connectivities with the highest discriminative power for each classifier. The regions related to the three types of connectivity patterns were widely distributed throughout the entire brain rather than restricted to a few specific brain regions (See Fig. 2 for details).

## 4 Discussion

To the best of our knowledge, this was the first study that used a multiclass pattern analysis method based on resting-state functional magnetic resonance imaging to investigate the difference of whole-brain functional connectivity among schizophrenic patients, healthy siblings, and healthy controls. PCA was employed to reduce the dimensionality of the original feature space for its ability to perform a direct connectivity mapping of the whole brain within a reasonable time frame. And the use of PCA leads to a minimum loss of information, which did not affect the functional connectivity analysis [25]. We chose the generalization rate as the statistic to estimate the statistical significance of the classification accuracy by permutation test. The results demonstrate that the multiclass classifier learned the relationship between the data and the labels with a probability of being wrong of  $<0.001$ .

Our classification accuracies for these three groups were respectively 62.5, 48.0, and 63.64 %, resulting in a total accuracy of 57.75 %, which is significantly above



**Fig. 2** Region weights and the distribution of the 330 discriminative functional connectivities responding to the state patterns **a** trait pattern **b** and compensatory pattern **c** in bottom and right view, respectively. Regions are color-coded by category. The line colors represent the discriminative power of the connectivities

the chance level of  $1/3$ . These results suggested our methods reliably captured discriminative resting-state functional connectivity patterns among schizophrenic patients, their healthy siblings, and healthy controls at the individual subject level with a high degree of accuracy. From the confusion matrix for classification results, we found that the classification accuracy of the healthy siblings was comparatively lower, and the healthy siblings were easier to be misclassified as schizophrenic patients than healthy controls. In addition, the healthy siblings were much easier to be misclassified as the healthy controls than as the schizophrenic patients. These results might help to explain the normal daily behaviors exhibition of the healthy siblings and demonstrated that healthy siblings of schizophrenic patients have an potential higher risk for developing schizophrenia compared with the general population, which was consistent with previous finds [17].

The identified state connectivity patterns represented the brain functional connectivities which are observably different in schizophrenic patients and non-psychotic individuals (including healthy siblings and healthy controls), and they might reveal the nature of abnormality in brain connectivity pattern of schizophrenic patients. The trait patterns, obtained from the healthy controls against rests (including schizophrenia patients and their healthy siblings) classifier, likely reflect the genetic vulnerability to develop schizophrenia. The compensatory connectivity patterns, composed of connectivities which contributed most in separating the healthy siblings from schizophrenia patients and healthy controls, potentially reveal unique brain functional connectivity pattern owned by the healthy siblings. Interestingly, we found few shared connectivities among these

three types of connectivity patterns, this raises the question that whether the functional connectivity patterns of the brains of schizophrenia patients, their healthy siblings, and healthy controls are different in a whole-brain level rather than restricted to a few specific brain regions This is an interesting finding need to be considered with caution and confirmed by further investigations.

## 5 Limitations

There are two possible limitations in the current study that should be considered. First, due to the limited size of samples and various noises resulting from acquisition hardware, our findings needed to be confirmed with a larger sample size in the future. Second, we focused on the whole-brain functional connectivity pattern, and identified the discriminative connectivities using multiclass pattern analysis, while how do these connectivities interact remained to be further investigated, i.e., whether these connectivities were decrease or increase what might reflect the dysfunctional mechanism in the pathophysiology of schizophrenia was ignored in this study. Finally, some of the schizophrenic patients in this study were medicated. Previous studies suggest antipsychotic treatments that tend to change aberrant connectivity [26]. We are yet unable to exclude the possibility effects of antipsychotic treatment.

## 6 Conclusion

We used multiclass pattern analysis to investigate the functional connectivities throughout the entire brain in the schizophrenia patients and their healthy siblings with resting-state fMRI. We identified brain connectivity pattern related to the state of having schizophrenia, provided a possible neuroendophenotype reflecting the genetic vulnerability to develop schizophrenia, and revealed unique brain functional connectivity pattern owned by the healthy siblings. Our investigation suggested that the schizophrenia patients, their healthy siblings, and healthy controls might have different brain connection patterns, which needs to be considered with caution and confirmed by further investigations.

## References

1. Laursen TM, Labouriau R, Licht RW, Bertelsen A, Munk-Olsen T, Mortensen PB (2005) Family history of psychiatric illness as a risk factor for schizoaffective disorder: a Danish register-based cohort study. *Arch Gen Psychiatry* 62:841–848

2. Cardno A, Marshall E, Coid B, Gottesmann I, Farmer A, McGuffin P (1999) Heritability estimates for psychotic disorders. *Arch Gen Psychiatry* 56:162–168
3. Tuulio-Henriksson A, Arajarvi R, Partonen T, Haukka J, Varilo T, Schreck M, Cannon T, Lonnqvist J (2003) Familial loading associates with impairment in visual span among healthy siblings of schizophrenia patients. *Soc Biol Psychiatry* 54:623–628
4. Li X, Sundquist J, Hemminki K, Sundquist K (2009) Familial risks of psychotic disorders and schizophrenia among siblings based on hospitalizations in Sweden. *Psychiatry Res* 166:1–6
5. Gottesman II, Shield J (1973) Genetic theorizing and schizophrenia. *Br J Psychiatry* 122:15–30
6. Sadock B (2007) Kaplan and Sadock's synopsis of psychiatry: behavioral sciences/clinical psychiatry (10th ed). *Indian J Psychiatry* 51:331
7. Brahmabhatt SB, Haut K, Csernansky JG, Barch DM (2006) Neural correlates of verbal and nonverbal working memory deficits in individuals with schizophrenia and their high-risk siblings. *Schizophr Res* 87:191–204
8. Calabrese DR, Wang L, Harms MP, Ratnanather JT, Barch DM, Cloninger CR, Thompson PA, Miller MI, Csernansky JG (2008) Cingulate gyrus neuroanatomy in schizophrenia subjects and their non-psychotic siblings. *Schizophr Res* 104:61–67
9. Fan Y, Gur RE, Gur RC, Wu X, Shen D, Calkins ME, Davatzikos C (2008) Unaffected family members and schizophrenia patients share brain structure patterns: a high-dimensional pattern classification study. *Biol Psychiatry* 63:118–124
10. Li X, Branch CA, Nierenberg J, DeLisi LE (2010) Disturbed functional connectivity of cortical activation during semantic discrimination in patients with schizophrenia and subjects at genetic high-risk. *Brain Imag and Behav* 4:109–120
11. Huang XQ et al (2010) Localization of cerebral functional deficits in treatment-naive, first-episode schizophrenia using resting-state fMRI. *NeuroImage* 49:2901–2906
12. Liang M, Zhou Y, Jiang TZ, Liu ZN, Tian LX, Liu HH, Hao YH (2006) Widespread functional disconnectivity in schizophrenia with resting-state functional magnetic resonance imaging. *Neuroreport* 17(2):209–213
13. Shen H, Wang L, Liu Y, Hu D (2010) Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI. *NeuroImage* 49:3110–3121
14. Zhou Y, Shu N, Liu Y, Song M, Hao YH, Liu HH, Yu CS, Liu ZN, Jiang TZ (2008) Altered resting-state functional connectivity and anatomical connectivity of hippocampus in schizophrenia. *Schizophr Res* 100:120–132
15. Repovs G, Csernansky JG, Barch DM (2011) Brain network connectivity in individuals with schizophrenia and their siblings. *Biol Psychiatry* 69:967–973
16. Pereira F, Mitchell T, Botvinick M (2009) Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage* 45:S199–S209
17. Liu M, Zeng L-L, Shen H, Liu Z, Hu D (2012) Potential risk for healthy siblings to develop schizophrenia: evidence from pattern classification with whole-brain connectivity. *Neuroreport* 23:265–268
18. Kay SR, Fiszbein A, Opler LA (1987) The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 13:261–276
19. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15:273–289
20. Achard S, Salvador R, Whitcher B, Suckling J, Bullmore E (2006) A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J Neurosci* 26:63–72
21. Maćkiewicz A, Ratajczak W (1993) Principal components analysis. *Comput Geosci* 19(3):303–342
22. Golland P, Fischl B (2003) Permutation tests for classification: towards statistical significance in image-based studies. *Inf Process Med Imaging* 2732:330–341

23. Wang L, Shen H, Tang F, Zang Y, Hu D (2012) Combined structural and resting-state functional MRI analysis of sexual dimorphism in the young adult human brain: an MVPA approach. *NeuroImage* 61:931–940
24. Kaiser MD et al (2010) Neural signatures of autism. *Proc Natl Acad Sci USA* 107:21223–21228
25. Zhou Z, Chen Y, Ding M, Wright P, Lu Z, Liu Y (2009) Analyzing brain networks with PCA and conditional Granger causality. *Hum Brain Mapp* 30:2197–2206
26. Schlosser R, Gesierich T, Kaufmann B, Vucurevic G, Hunsche S, Gawehn J, Stoeter AP (2003) Altered effective connectivity during working memory performance in schizophrenia: a study with fMRI and structural equation modeling. *NeuroImage* 19:751–763

# Network Organization of Information Process in Young Adults' Brain

Shao-Wei Xue, Yi-Yuan Tang and Lan-Hua Zhang

**Abstract** In order to characterize non-random organization patterns of information process in the brain, we combine complex network analysis and resting-state functional magnetic resonance imaging to investigate brain activity derived from young adults, and then extract the tree layout and module structure of whole-brain network. These network organizations may be associated with the emergence of complex dynamics that supports the brain's moment-to-moment responses to the external world and widen understanding potentially biological mechanisms of brain function.

**Keywords** Brain networks · Maximum spanning tree · Modularity · Hubs

## 1 Introduction

Complex networks have attracted great attention as a compelling framework with which complex systems are being studied in many fields, such as computer science, physics, social sciences, and neuroscience [1]. Under this framework, many natural and man-made networks generated from different datasets have been exhibited common principles that govern network behavior and can be quantitatively characterized by the same parameters [2].

In neuroscience, studies on relationship between brain and neural networks have deep historical roots. However, the application of complex networks theory to the brain is under development during recent years. Newly developed recording

---

S.-W. Xue · L.-H. Zhang

Institute of Neuroinformatics, Dalian University of Technology, Dalian 116024, China

Y.-Y. Tang (✉)

Department of Psychology, Texas Tech University, TX 79409, USA

e-mail: yy2100@126.com

and imaging techniques help us to acquire large brain activity-based datasets dynamically and noninvasively. Especially, resting-state functional magnetic resonance imaging offers a novel approach to assess the spontaneous brain activity as low-frequency (usually  $<0.1$  Hz) fluctuations in blood oxygen level-dependent signals. These low-frequency fluctuations potentially provide insight into the brain's functional organization because their extremely disproportionate allocation of energy resources is devoted to functionally significant intrinsic activity. Moreover, such fluctuations are not random noise, but demonstrate temporal synchrony within widely distributed brain regions. Communication between neurons and neuronal ensembles is probably established to a large extent through the synchronization of their activity [3]. Network analysis of these datasets is helpful for uncovering the organizational principles that underlies information process in the brain.

However, what can we extract about non-random organization patterns of information process in the brain? Graph theory-based network analysis techniques are employed to characterize brain activity data derived from young adults. Functional brain networks are regarded as graph representations of brain activity, where the nodes represent anatomically defined regions and the edges describe their functional connectivity between each pair.

## 2 Datasets and Network Construction

The experimental data for network construction were acquired from 33 healthy volunteers (17 males, mean age:  $21.6 \pm 2.23$  (SD), right-handed) on a 3.0 Tesla Philips Achieva MR imaging system. Images of the whole brain using an echo-planar imaging (EPI) sequence were collected in 36 slices (TR = 2000 ms, TE = 30 ms, flip angle =  $80^\circ$ , FOV = 23 cm, matrix =  $64 \times 64$ , 4 mm thickness and 0 mm gap). Images preprocessing and further analysis are performed including slice timing, head motion correction, spatial normalization and smoothing with an 8 mm Gaussian kernel, full width at half maximum. Finally, temporal band-pass filtering (between 0.01 and 0.08 Hz) is performed in order to reduce the effects of low-frequency drift and high-frequency noise. Thus, the dataset available for network analysis comprise thirty-three multivariate time series.

In this paper, we investigate the topological characters of brain networks derived from resting-state functional magnetic resonance imaging. Detailed steps include: (1) using a used anatomical automatic labeling atlas [4, 5] to parcellate the whole brain into 90 cortical and sub-cortical regions as network nodes, (2) estimating functional relationship between network nodes by partial correlation analysis of 90 regional time series, (3) generating a correlation matrix ( $90 \times 90$ ) by compiling all coefficients between nodes and applying a network density threshold to each element to produce a binary adjacency matrix or undirected graph, (4) calculating the network parameters of interest in this graphical model of a brain network.

### 3 Experimental Results

Figure 1 shows a maximum spanning tree derived from the mean matrix across 33 participants using Kruskal's algorithm [6], which connects all nodes ( $N = 90$ ) such that the sum of their 89 connectivity coefficients is maximal. Each vertex is first regarded as a separate tree, and then extra edges are added in order of the decreasing correlation coefficients if those edges connect two different trees, until all nodes form a maximum spanning tree. The resulting connectivity skeleton is represented using a procedure known as the Kamada-Kawai algorithm [7] implemented in Network Workbench Tool (<http://nwb.slis.indiana.edu>). The algorithm generates the layout in an esthetically pleasing way, and positions the network nodes in two dimensional spaces so that all the edges are of more or less equal length and there are as few crossing edges as possible. As shown in Fig. 1, most pairs of inter-hemispheric homogenous regions are tightly coupled, and two major sub-networks with intrinsic functional activity (the default brain activity and attention-related regions) seem to be at the center of the network layout, whose nodes are highlighted in black.

A maximum spanning tree can be used as a sparse representation of whole-brain networks. However, tree is by definition acyclic and its edges do not form clusters. Accordingly, we grow the tree by adding extra edges in order of the decreasing coefficients. The corresponding binary adjacency matrices are obtained for further mesoscale analysis. The brain network comprising about 200 edges or 5 % of the possible  $N(N - 1)/2$  edges (4005) is selected for modularity analysis by Newman's spectral optimization algorithm [8]. Modularity is used to evaluate the goodness of a partition of a graph in terms of a modularity matrix and its eigenvector associated with the largest positive eigenvalue. Thus, the module identification problem becomes a modularity optimization problem. The brain network is supposed to consist of  $n$  nodes connected by  $m$  edges. The index vector  $s$  represents any partition in two groups, whose component  $s_i$  is +1 if vertex  $i$  is in one group and -1 if it is in the other group. Modularity function is expressed as follows:

$$Q = \frac{1}{4m} s^T B s. \quad (1)$$

The modularity can be optimized using the eigenvalues and eigenvectors of a special matrix, the modularity matrix, whose elements are

$$B_{ij} = A_{ij} - \frac{k_i k_j}{2m}, \quad (2)$$

where  $k_i$  and  $k_j$  are the degrees of the nodes. Like a Laplacian matrix, B has always an eigenvector (1, 1, 1 ...) with eigenvalue 0, because the elements of each of its rows and columns sum to zero. The algorithm generates a modularity matrix with an associated modularity score. The modularity score ( $Q = 0.52$ ) of the brain network is significantly higher ( $P < 0.05$ ) than the mean value of the equivalent



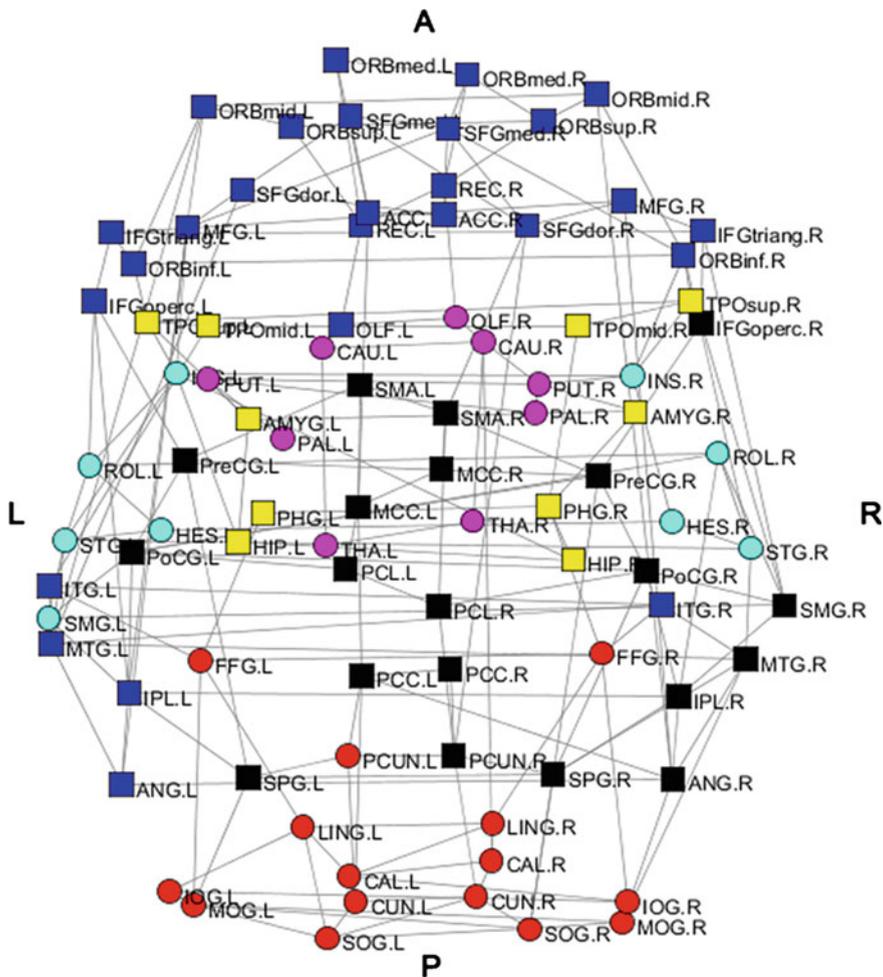


Fig. 2 Module structure of brain networks. *P* posterior, *L* left, *R* right

where  $m_i$  is the module containing node  $i$ ,  $k_i(m_i)$  is the number of links of node  $i$  to other nodes in its module  $m_i$ , and  $\bar{k}(m_i)$  and  $\sigma_{k(m_i)}$  are the respective mean and standard deviation of the within-module  $m_i$  degree distribution. Two nodes with the same  $z$ -score play different roles if one is connected to several nodes in other modules and another is not.

We then classify nodes based on their connections to nodes in other modules than their own. The participation coefficient  $P_i$  of node  $i$  is computed as

$$P_i = 1 - \sum_{m \in M} \frac{k_i(m_i)^2}{k_i} \tag{4}$$

where  $M$  is the set of modules, and  $k_i(m_i)$  is the number of links between  $i$  and all nodes in module  $m_i$ , and  $k_i$  is the total degree of node  $i$ . Thus the participation coefficient of a node tends to one if its links are uniformly distributed among all the modules and zero if all links are within its own module.

We assign four roles to the nodes of the brain network. We first classify nodes with  $z \geq 1$  as module hubs and nodes  $z < 1$  as non-hubs. Both hub and non-hub nodes are then more finely characterized by using the values of  $P_i$ , and the hub nodes are further subdivided into R1 connector hubs ( $P_i > 0.3$ ) and R2 provincial hubs ( $P_i < 0.3$ ), and non-hub nodes are divided into R3 connector non-hubs ( $P_i > 0.3$ ) and R4 peripheral non-hubs ( $P_i < 0.3$ ).

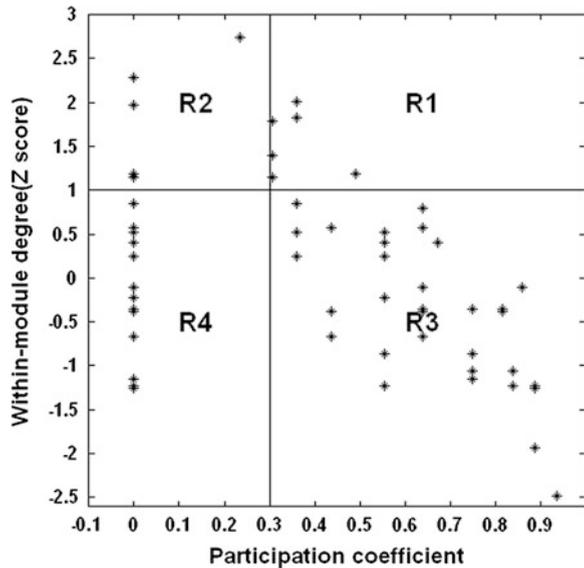
Figure 3 shows the distribution of the roles obtained from the brain network over the  $z$ - $P$  parameter space. Most of the nodes (76/90; 84 %) can be assigned as non-hubs. The anatomical distribution of the node roles is depicted in Fig. 4. Interestingly, these parameter distributions could help us to characterize the complex statistical properties and heterogeneity of edges and nodes.

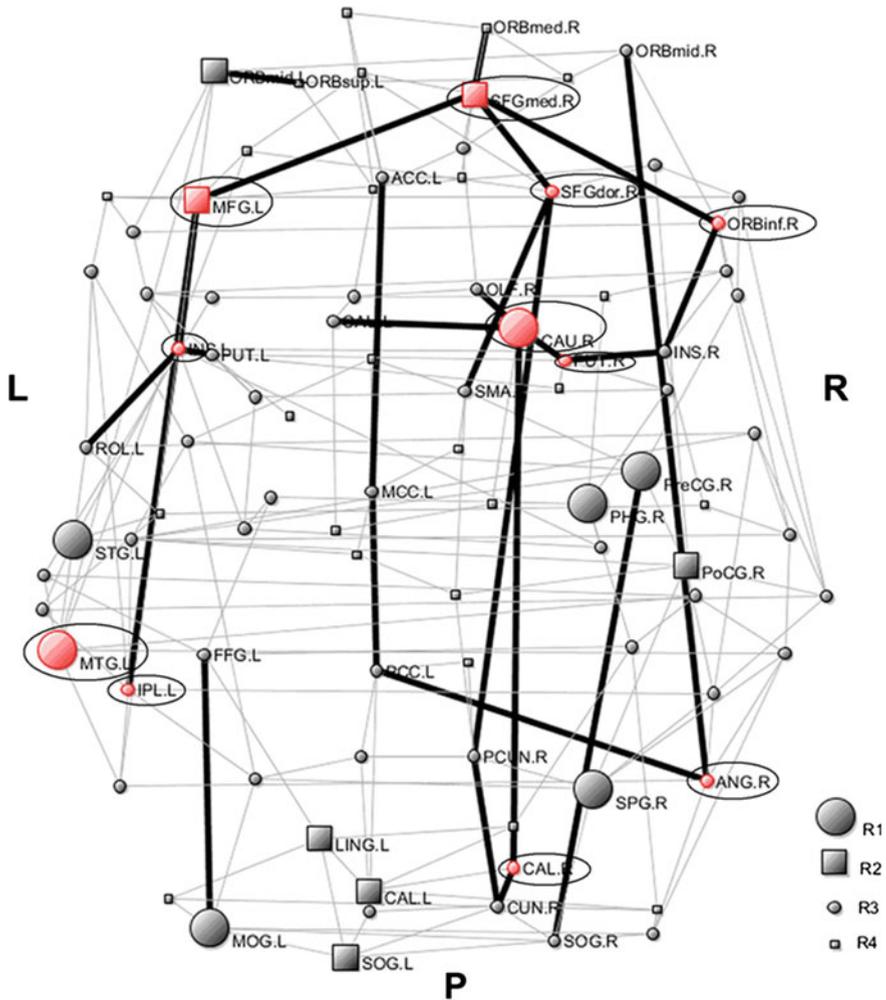
We further characterize global roles of an individual node or edge in the brain network by calculating betweenness centrality [9]. The betweenness centrality of a node can be defined as follows:

$$N_i = \sum_{s \neq i, s \neq t, t \neq i} \frac{g_{st}(i)}{g_{st}}, \quad (5)$$

where  $g_{st}$  is the total number of all shortest paths from node  $s$  to node  $t$ , and  $g_{st}(i)$  is the number of shortest paths from node  $s$  to node  $t$  that pass through the node  $i$ . Betweenness measures the ability of a node over information flow between other

Fig. 3 Regional node roles





**Fig. 4** The spatial distribution of hubs and bridges in the brain network

nodes in the whole-brain network. The brain regions with high betweenness ( $N_i > \text{mean} + \text{SD}$ ) are considered global hubs in the brain network. The identified hub regions include 6 association cortex regions, 2 limbic/paralimbic cortex regions, 1 primary motor cortex region and 2 subcortical regions. Red nodes in Fig. 4 are used to represent the global hubs. The roles of these global hub regions in multiple highly integrated functional systems have been previously reported [10]. Similarly, edges are identified as the bridges in the brain network, if their edge betweenness [9] values are at least one standard deviation (SD) greater than the average edge betweenness of the network. Thick lines in Fig. 4 show the bridge connections between the two brain regions. It should be noted that most of

the hub nodes (10 of 11) are linked with the identified bridge edges, suggesting there is a tendency for these hub nodes and bridge edges to converge on functionally integrated core system.

## 4 Discussion

In this paper, we measure brain activity and employ network analysis approach based on graph theory to reveal functional organization in young adults' brain. This approach, unlike previous conventional functional imaging studies, quantitatively assesses functionally integrated core at various levels. At a global level of tree analysis, most brain areas of the default and attention-related network activity are distributed at the center of the network layout and form important functional hubs of information processing at rest to support the brain's moment-to-moment responses to the external world [11].

At a mesoscale level of modularity analysis, our identified module structure of brain networks is closely associated with some biologically meaningful functional systems of the brain, and the non-random structure with a balanced interplay between local segregation and global integration of distant anatomofunctional brain regions can contribute to efficiency of parallel information transfer at low physical connection cost [12]. Moreover, modularity has the advantage of allowing evolutionary adaptation of one functional subsystem, without risking loss of function in others [13].

At individual node or edge level, we found hub nodes and bridge edges of brain networks had a tendency to converge on functionally integrated core system or rich-club organization [14], indicating playing key roles in integration of information process in the brain [15].

## 5 Conclusion

In summary, we have applied graph theory-based network analysis techniques to study the organization of intrinsic spontaneous information process in the young adults' brain at rest. Several specific non-random network organization patterns including whole network layout and modular structure have been found. This non-randomness may be associated with the brain's moment-to-moment responses to the external world. Many applications of this work to brain dysfunction will be performed in our future work.

**Acknowledgments** The authors thank the Institute of Neuroinformatics staffs for assistance with data collection. This work was supported by National Natural Science Foundation of China 60971096, 2012CB518200.

## References

1. Seung HS (2011) Neuroscience: towards functional connectomics. *Nature* 471:170–172
2. Costa LdF, Oliveira ON, Travieso G, Rodrigues FA, Villas Boas PR, Antiqueira L, Viana MP, Correa Rocha LE (2011) Analyzing and modeling real-world phenomena with complex networks: a survey of applications. *Adv Phys* 60:329–412
3. Zhou C, Zemanová L, Zamora-Lopez G, Hilgetag CC, Kurths J (2007) Structure–function relationship in complex brain networks expressed by hierarchical synchronization. *New J Phys* 9:178
4. Liu Y, Liang M, Zhou Y, He Y, Hao Y, Song M, Yu C, Liu H, Liu Z, Jiang T (2008) Disrupted small-world networks in schizophrenia. *Brain* 131:945–961
5. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289
6. Kruskal JB (1956) On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc Am Math Soc* 7:48–50
7. Kamada T, Kawai S (1989) An algorithm for drawing general undirected graphs. *Inf Process Lett* 31:7–15
8. Newman MEJ (2006) Modularity and community structure in networks. *Proc Natl Acad Sci U S A* 103:8577–8582
9. Freeman L (1977) A set of measures of centrality based on betweenness. *Sociometry* 40:35–41
10. Tomasi D, Volkow ND (2010) Functional connectivity density mapping. *Proc Natl Acad Sci U S A* 107:9885–9890
11. Fox MD, Corbetta M, Snyder AZ, Vincent JL, Raichle ME (2006) Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc Natl Acad Sci U S A* 103:10046–10051
12. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10:186–198
13. Kashtan N, Alon U (2005) Spontaneous evolution of modularity and network motifs. *Proc Natl Acad Sci U S A* 102:13773–13778
14. Zhou S, Mondragón RJ (2004) The rich-club phenomenon in the internet topology. *IEEE Commun Lett* 8:180–182
15. Bullmore E, Sporns O (2012) The economy of brain network organization. *Nat Rev Neurosci* 13:336–349

# Reconfigurable Control Allocation of Multi-Surfaces Aircraft Based on Improved Fixed Point Iteration

Kejun Bi, Weiguo Zhang, Chengzhi Chi and Jingkai Zhang

**Abstract** For the real-time requirement of reconfigurable control allocation problem in the field of multi-surfaces aircraft, the control allocation scheme based on max direction derivative increment (MDDI) fixed point (FXP) iteration is proposed. The increment update for current iteration along the MDD and the design steps are given. Moreover, the convergence of the improved method is also proved. Comparisons of different methods are simulated in multi-surfaces aircraft model. The simulation results show the rapidity of MDDIFXP method compared with the original one and the effectiveness of the method in solving reconfigurable control allocation problem of multi-surfaces aircraft.

**Keywords** Multi-surfaces aircraft · Control allocation · Fixed point arithmetic · Pseudo-inverse method · Reconfiguration · Improvement

## 1 Introduction

Due to the increasing requirements on the reliability, maneuverability, and survivability of modern aircraft, control surfaces are no longer limited to three conventional ones: aileron, elevator, and rudder, and many more control surfaces have been introduced. With these redundant control surfaces, the problem of allocating these controls to achieve the desired moments becomes non-unique and far more complex [1]. So effective control allocation schemes has been studied following the work of Durham [2] to distribute the required control moments over the control surfaces [3]. In particular, in the case of control surface failures or damages [4], an effective and rapid reallocation of the control surface deflections

---

K. Bi (✉) · W. Zhang · C. Chi · J. Zhang  
School of Automation, Northwestern Polytechnical University, Xian 710072, China  
e-mail: chengzhi.chi@gmail.com

with the remaining healthy control surfaces is needed in order to maintain acceptable performance, and this requirement asks for the so-called reconfigurable control allocation or control reallocation technique, which is a necessary part of the reconfigurable (or fault tolerant) flight control systems, and that is important for continuing flight mission or safe landing.

Control allocation is the problem of distributing the control requirements among redundant control surfaces for satisfying the optimized objectives within their range of position and rate limits [5–7]. A comparison of different control allocation methods are documented in [8]. However, the fault-tolerant control reallocation problem has not been well investigated except a few notable works presented in [1, 9–11].

In this paper, we propose control allocation scheme based on an improved fixed point (FXP) algorithm, which is evaluated in ADMIRE aircraft model. The paper is organized as follows: The improved FXP algorithm max direction derivative increment fixed point (MDDIFXP) is presented in Sect. 2. MDD can be achieved by computing the derivative along all directions of current iteration, and the increments of the former and current iteration are taken as the update for current iteration along the MDD, also, the convergence of the method is proved. Simulations of the control allocation scheme compared to different methods in the normal situation and fault-tolerant reconfiguration in the presence of different partial control surface faults are presented in Sect. 3. Finally, conclusions and future work are described in Sect. 4.

## 2 MDDIFXP Based Quadratic Programming

The control allocation problem studied in the paper is achieved by solving a quadratic programming (QP) problem which involves the minimization of a quadratic cost function subject to both equality and inequality constraints. In this section, we will first discuss the establishment of FXP and MDDIFXP method for QP problem. We then prove the convergence of MDDIFXP method.

### 2.1 FXP Method

The optimal control input is given by the solution to a weighted optimization problem [8]:

$$\min_u J = \|W_u(u - u_d)\|_2^2 + \gamma \|W_v(Gu - v)\|_2^2 \quad (1a)$$

$$\text{Subject to } \underline{u} \leq u \leq \bar{u} \quad (1b)$$

Here  $u_d$  is the desired control input and  $v$  is virtual control input,  $W_u$  and  $W_v$  are weighting matrices.  $G$  is control effectiveness matrix,  $\gamma > 0$  is the weighting factor. Equation (1a) can be rewritten by:

$$\begin{aligned}
J &= (u - u_d)^T W_u (u - u_d) + \gamma (Gu - v)^T W_v (Gu - v) \\
&= u^T W_u u - u^T W_u u_d - u_d^T W_u u + u_d^T W_u u_d + \gamma (u^T G^T W_v G u \\
&\quad - u^T G^T W_v v - v^T W_v G u + v^T W_v v) \\
&= u^T (W_u + \gamma G^T W_v G) u + u^T (-2W_u u_d - 2\gamma G^T W_v v) \\
&\quad + u_d^T W_u u_d + \gamma v^T W_v v \\
&= \frac{1}{2} u^T T u + u^T d + r
\end{aligned}$$

Here  $T = 2W_u + 2\gamma G^T W_v G$ ,  $d = -2W_u u_d - 2\gamma G^T W_v v$ ,  $r = u_d^T W_u u_d + \gamma v^T W_v v$ .

Because  $r$  is a constant value in a sampled cycle, so the solution will be the same if we remove  $r$  from the cost function. Then, the control allocation problem based on QP can be reduced to

$$\min_u J = \frac{1}{2} u^T T u + u^T d \quad (2a)$$

$$\text{Subject to } \underline{u} \leq u \leq \bar{u} \quad (2b)$$

When  $u_d = 0$ , Eq. (1) can be rewritten by [8, 9]:

$$J = \frac{1}{2} [(1 - \varepsilon)(Gu - v)^T Q_1 (Gu - v) + \varepsilon u^T Q_2 u] \quad (3a)$$

$$\text{Subject to } \underline{u} \leq u \leq \bar{u} \quad (3b)$$

Here  $Q_1 = W_v^T W_v > 0$ ,  $Q_2 = W_u^T W_u > 0$ ,  $\varepsilon = (1 + \gamma)^{-1}$ .

Suppose  $u = (u_1, \dots, u_m)^T \in R^m$ , which satisfies

$$s_i(u) = \begin{cases} \underline{u}, & u_i \leq \underline{u} \\ u_i, & \underline{u} < u_i < \bar{u}, \\ \bar{u}, & u_i \geq \bar{u} \end{cases} \quad (4)$$

$s(\cdot)$  is the vector saturator, then the algorithm becomes

$$u = s[(1 - \varepsilon)\omega G^T Q_1 v - (\omega T - I)u] \triangleq f(u) \quad (5)$$

Here  $\omega = \|T\|_F^{-1} = (tr(T^T T))^{-\frac{1}{2}}$ , which decides the step length.  $T = (1 - \varepsilon)G^T Q_1 G + \varepsilon Q_2$ ,  $\varepsilon \in (0, 1)$ .

So the iterative formula of FXP method is listed as follows:

$$u^{k+1} = f(u^k) (k = 0, 1, \dots, N) \quad (6)$$

## 2.2 MDDIFXP Method

The principle of MDDIFXP is improvement of increment update for current iteration along the MDD, in order to fasten the process of iteration.

**Definition 1** Suppose that  $f : R^n \rightarrow R^n$  is a continuous function, and for the unit vector  $e_i = (\delta_1^i, \dots, \delta_n^i)^T$ ,  $\delta_i^i = 1$ ,  $\delta_j^i = 0$  ( $j \neq i, i = 1, 2, \dots, n$ ), the derivative of function  $f(u0 + te_i)$  at point  $t = 0$  (if it exists) is named as the first partial derivative of  $f$  at point  $u0$  with respect to  $u_i$  ( $i = 1, 2, \dots, n$ ). For  $\forall i = 1, 2, \dots, n$ , if the first partial derivative of  $f$  at point  $u$  with respect to  $u_i$  exists, then the gradient of  $f(u)$  at point  $u$  is defined as follows:

$$\nabla f(u) = \left( \frac{\partial f(u)}{\partial u_1}, \frac{\partial f(u)}{\partial u_2}, \dots, \frac{\partial f(u)}{\partial u_n} \right)^T \quad (7)$$

**Definition 2** For  $u0 \in R^n$ ,  $d \in R^n$ , the directional derivative of  $f$  at point  $u0$  with respect to direction  $d$  is defined as follows:

$$\frac{\partial f(u0)}{\partial d} = \lim_{t \rightarrow 0^+} \frac{f(u0 + td) - f(u0)}{t} \quad (8)$$

Suppose  $Df(u0; d)$  is the directional derivative of  $f$  at point  $u0$  with respect to direction  $d$ . When the first partial derivative of  $f$  is continuously differentiable, the directional derivative can be computed by:

$$Df(u0; d) = \nabla f(u0)^T d. \quad (9)$$

Then, the gradient of FXP function  $f(u)$  at point  $u$  is defined as:

$$\overset{f}{\nabla} f(u) = \left( \frac{f(u + \Delta te_1) - f(u)}{\Delta t}, \dots, \frac{f(u + \Delta te_n) - f(u)}{\Delta t} \right)^T. \quad (10)$$

Here, the  $f$  at the top left corner of the left side of the equation means FXP.  $\Delta t$  is the iteration interval, and which is set to 1 normally. So the directional derivative of  $f$  at point  $u0$  with respect to direction  $d$  can be computed by:

$$Df(u0; d) = \overset{f}{\nabla} f(u0)^T d \quad (11)$$

Then, MDD is the unit vector  $e_i$  satisfies

$$\text{Max} \left[ \frac{f(u0 + \Delta te) - f(u0)}{\Delta t} \right] \quad (i = 1, 2, \dots, n) \quad (12)$$

**Definition 3** For  $u^k, u^{k+1} \in R^n$  and unit vector  $e_i$ , the increment of  $f$  at point  $u^{k+1}$  with respect to direction  $e_i$  is defined as:

$$\Delta f(u^{k+1}; e_i) = \frac{f(u^{k+1} + \Delta t e_i) - f(u^{k+1})}{\Delta t} - \frac{f(u^k + \Delta t e_i) - f(u^k)}{\Delta t} \quad (13)$$

MDDIFXP method takes the increments of the former and current iteration as the update of current iteration along the MDD, and the design steps are:

1. Compute  $f(u^k), f(u^{k+1})$ ;
2. Compute  $\frac{f}{\nabla} f(u^{k+1})^T e_i (i = 1, 2, \dots, n)$ , take the unit vector  $e_i$  as the MDD vector  $e_M$ , which satisfies  $\text{Max} \left[ \frac{f(u^{k+1} + \Delta t e_i) - f(u^{k+1})}{\Delta t} \right]$ ;
3. Compute  $\Delta f(u^{k+1}; e_M)$  and update  $f(u^{k+1}) = f(u^{k+1}) + \Delta f(u^{k+1}; e_M) e_M$ ;
4. Take  $u^{k+2} = f(u^{k+1})$  as the new iteration point, and compute  $f(u^{k+2}), f(u^{k+3})$ ;
5. If  $k < K (K \in N)$  and  $\left\| \frac{f}{\nabla} f(u^{k+1}) \right\| > \xi (\xi > 0)$ , repeat step 1–4, or turn to basic FXP algorithm.

Note that the updated  $f(u^{k+1})$  should also satisfies constraint condition.

### 2.3 The Convergence of MDDIFXP Method

Before discussing the convergence of MDDIFXP algorithm, we introduce the convergence of basic FXP algorithm.

When the cost function is convex, every optimization algorithm should meet Kuhn-Tucker condition [12]. For convex programming, Kuhn-Tucker condition is the necessary and sufficient condition. And the necessary and sufficient condition of  $u^*$  is the solution to problem of positive definite quadratic programming is that  $u^*$  is the Kuhn-Tucker point [12]. From [13], we can see the solution to Eq. (5) is also the solution to Eq. (2), so FXP algorithm can solve the QP problem accurately [14]. Then, the convergence of basic FXP algorithm is proved.

Hence,  $u^*$  is the globally optimal solution to the control allocation QP problem. Suppose  $\{u^n\}$  is the sequence of iterations for basic FXP algorithm, then the sequence should converge to  $u^*$ . According to the design steps of MDDIFXP algorithm, the improved algorithm changes only finite term of the sequence  $\{u^n\}$ ; therefore, the convergence of the sequence is not changed. Further, due to the limit of sequence is unique, the iteration sequence of MDDIFXP converge to  $u^*$ .

### 3 Simulation and Performance Evaluation

#### 3.1 MDDIFXP in Normal Control Allocation

ADMIRE aircraft model is used in the simulation, and the main data of simulation are listed as follow:

$$G = \begin{bmatrix} 0.7984 & -0.7984 & -4.5787 & -3.9413 & 3.9413 & 4.5787 & 2.6919 \\ 1.3841 & 1.3841 & -1.0906 & -1.7433 & -1.7433 & -1.0906 & 0.0046 \\ -0.3970 & 0.3970 & -0.2014 & -0.4256 & 0.4256 & 0.2014 & -1.6265 \end{bmatrix}$$

$$Wv = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 12 \end{bmatrix}, \quad Wu = I_7, \quad \varepsilon = 1e - 3, \quad K = 3, \quad \zeta = 1e - 3$$

The control surfaces are

$$u = [\delta_{c_l} \quad \delta_{c_r} \quad \delta_{ro_e} \quad \delta_{ri_e} \quad \delta_{li_e} \quad \delta_{lo_e} \quad \delta_r]^T$$

And the ranges of limits are listed in Table 1.

Two series of virtual inputs are generated in the simulation, as showed in black line in Figs. 1 and 2. The algorithms such as Pseudo-inverse (PINV), Linear Programming (LP), Interior point (IP), FXP, and MDDIFXP are compared. All algorithms are simulated 100 times, and the average computation time is listed in Table 2.

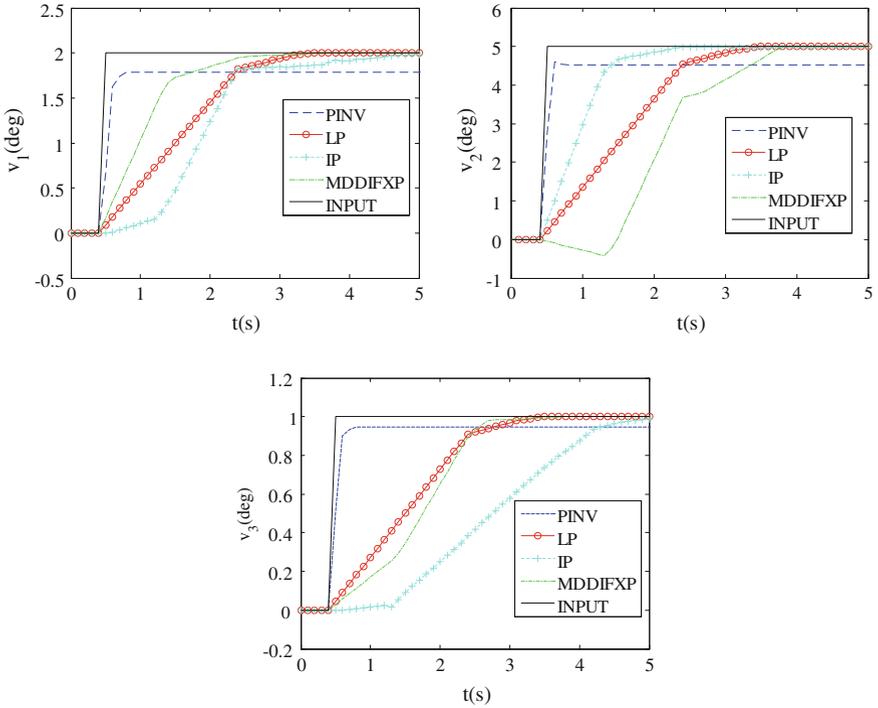
According to Table 2, we can see that among all the methods, MDDIFXP has a minimum computation time for two series of virtual inputs except PINV. Moreover, for two series of virtual inputs, MDDIFXP saves 2.12 and 2.21 % of the computation time, respectively. Figures 1 and 2 show that the outputs can track the corresponding virtual inputs with MDDIFXP, and from Fig. 1, we can see the outputs cannot track the corresponding virtual inputs well with PINV. According to the above analyzing, rapidity of the MDDIFXP method compared with the original FXP is demonstrated.

#### 3.2 MDDIFXP in Fault Reallocation

In the above section, we have verified the rapidity and effectiveness of MDDIFXP in normal control allocation. Then, we apply MDDIFXP method to the reconfigurable control allocation of multi-surfaces aircraft. In the following, two scenarios are simulated: (1) a floating fault to be occurred at 2.3 s in right inner elevon; (2) a 50 % of loss of control effectiveness to be occurred at 2 s in right outer elevon. It is assumed that the FDD information is available for control re-allocation, but with

**Table 1** Ranges of limits of control surfaces

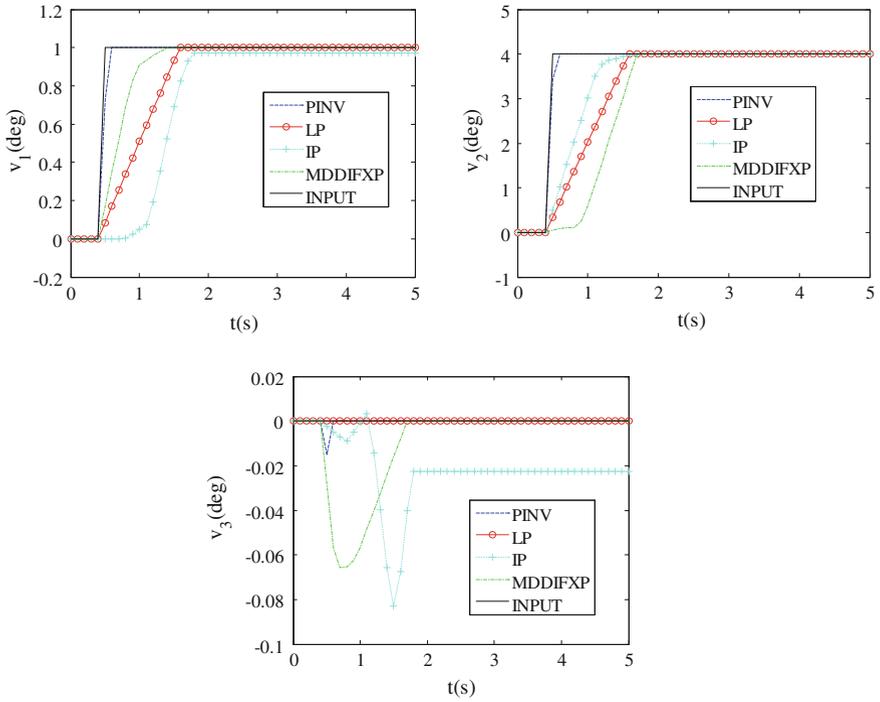
| Control surface | Control variable   | Minimum deflection (°) | Maximum deflection (°) | Maximum rate (°/s) |
|-----------------|--|------------------------|------------------------|--------------------|
| Canards         | $\delta_{c_l}, \delta_{c_r}$                             | -55                    | 25                     | 50                 |
| Elevons         | $\delta_{lie}, \delta_{rie}, \delta_{roe}, \delta_{loe}$ | -30                    | 30                     | 150                |
| Rudder          | $\delta_r$   | -25                    | 25                     | 100                |



**Fig. 1** Comparison of different methods on virtual input  $V_1$

0.3 and 0.5 s FDD time delay simulated for floating fault and loss of control effectiveness fault control reconfiguration, respectively.

As shown in Figs. 3 and 4, without reconfiguration, the outputs cannot track the inputs. With reconfiguration, the reconfigured outputs track the inputs with zero steady-state error. Meanwhile, the effectiveness of the MDDIFXP method in the reconfiguration control allocation problem is indicated.



**Fig. 2** Comparison of different methods on virtual input  $V_2$

**Table 2** Comparison of different algorithms

| No | CPU computation time (s) |        |        |        |         |
|----|--------------------------|--------|--------|--------|---------|
|    | PINV                     | IP     | LP     | FXP    | MDDIFXP |
| 1  | 0.0280                   | 0.0522 | 0.4517 | 0.0519 | 0.0508  |
| 2  | 0.0295                   | 0.0472 | 0.4506 | 0.0452 | 0.0442  |

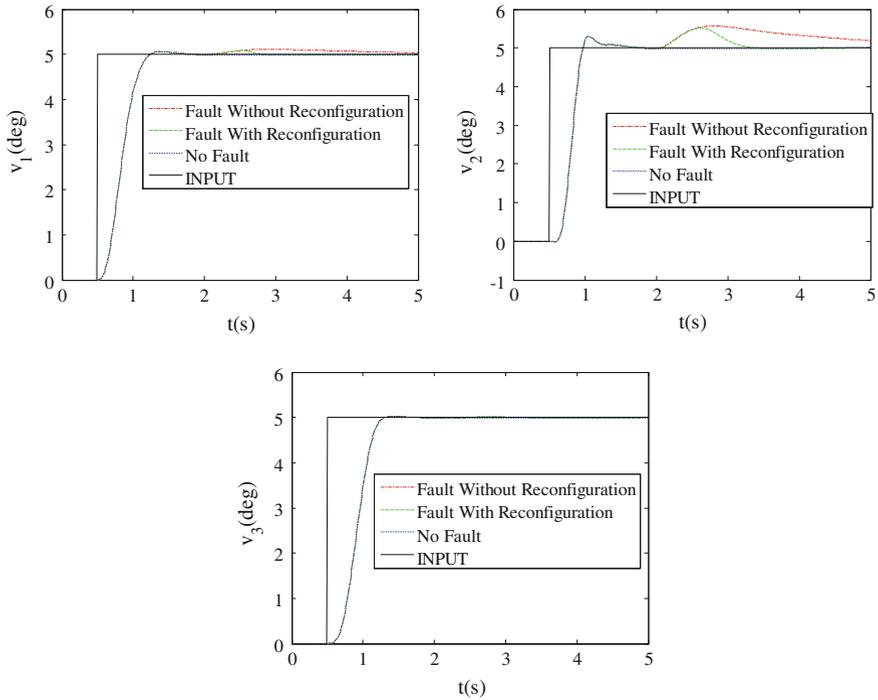


Fig. 3 Responses of virtual inputs with floating fault

### 4 Conclusion and Future Work

In this paper, an improved FXP method MDDIFXP is applied to control allocation and fault reallocation in ADMIRE model. The design steps of MDDIFXP are given, among which, the increment update for current iteration along the MDD is discussed in detail. Moreover, the convergence of the method is also proved. The comparisons of different methods applied to ADMIRE model show the rapidity of the MDDIFXP method, and the effectiveness of MDDIFXP method in the reconfiguration control allocation problem is also verified. Future work includes incorporation of FDD schemes in the ADMIRE model and further investigation on the conditions of step 5 in order to improve the performance of MDDIFXP method.

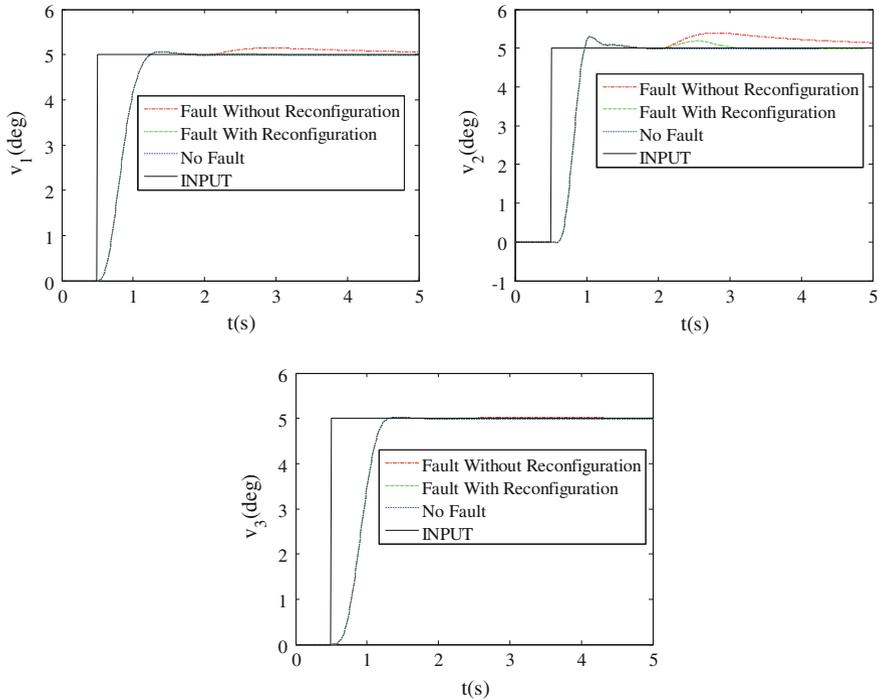


Fig. 4 Responses of virtual inputs with loss of control effectiveness

## References

1. Zhang YM, Suresh VS, Jiang B, Theilliol D (2007) Reconfigurable control allocation against aircraft control effector failures. In: IEEE international conference on control applications. IEEE Press, New York, pp 1197–1202
2. Durham WC (1993) Constrained control allocation. *J Guidance Control Dyn* 16:717–725
3. Shi JP, Zhang WG, Li GW, Liu XX (2010) Research on the allocation efficiency of redistributed pseudo-inverse algorithm. *Scientia Sinica (Informationis)* 40:519–525
4. Lombaerts T, Schravendijk MV, Chu P, Mulder JA (2011) Adaptive nonlinear flight control and control allocation for failure resilience. *Adv Aerosp Guidance, Navig Control I*:41–53
5. Schofield B, Häggglund T (2008) Optimal control allocation in vehicle dynamics control for rollover mitigation. In: American control conference. IEEE Press, New York, pp 3231–3236
6. Johansen TA, Fuglseth TP, Tøndel P, Fossen TI (2008) Optimal constrained control allocation in marine surface vessels with rudders. *Control Eng Pract* 16:457–464
7. Ahmadi J, Khaki-Sedighb A, Ohadia A (2012) Robustification of input redundant feedback systems using robust actuator weighting in the control allocation problem. *Int J Control* 85:1380–1400
8. Härkegård O (2003) Backstepping and control allocation with applications to flight control. Linköping University, Linköping
9. Burken JJ, Lu P, Wu ZL, Bahm C (2001) Two reconfigurable flight-control design methods: robust servomechanism and control allocation. *J Guidance Control Dyn* 24:482–493

10. Davidson JB, Lallman FJ, Bundick WT (2001) Integrated reconfigurable control allocation. In: AIAA guidance, navigation and control conference, American Institute of Aeronautics and Astronautics, Alexander, pp 1–11
11. Omerdic E, Roberts G (2004) Thruster fault diagnosis and accommodation for open-frame underwater. *Control Eng Pract* 12:1575–1598
12. Huang HX, Han JY (2006) *Mathematical programming*. Tsinghua University Press, Beijing
13. Wang J, Longoria RG (2009) Coordinated and reconfigurable vehicle dynamics control. *IEEE Trans Control Syst Technol* 17:723–732
14. Lu P (1996) Constrained tracking control of nonlinear systems. *Syst Control Lett* 27:305–314

# Particle Filter-Based Object Tracking and Handover in Disjoint View Multi-Cameras

Xiaoyan Sun, Faliang Chang and Wenhui Dong

**Abstract** In intelligent video surveillance, multiple cameras, even a distributed network of video sensors, have to be employed to monitor activities over a complex area nowadays. Hence, the continuous object tracking across multiple cameras and object handover between adjacent cameras is urgently needed, in which many appearance cues and spatial–temporal information can be employed. This paper fuses the spatial–temporal cues with appearance cues into a particle filter to handle the camera handover with multiple cameras having non-overlapping view. The spatial–temporal cues, including source and sink regions, their transition probabilities, and transition time among adjacent regions, are learned offline. Then a spatial–temporal progressive matching scheme using particle filter is proposed to deal with camera handover among adjacent cameras. In particle filter matching course, the commonly used appearance cue, i.e. the histogram in HSV color space is used. Once an object enters into sink region, we first continuously scatter particles in source regions related to this sink region according spatial–temporal information until the object emergence detected, and secondly, based on the particle weights of every source region, adjust their particle numbers till the camera handover is successfully completed. Encouraging experiment results show the efficiency of this scheme.

**Keywords** Multi-camera surveillance · Camera handover · Particle filter · Spatial–temporal information

---

X. Sun · F. Chang (✉) · W. Dong  
School of Control Science and Engineering, Shandong University, Shandong, China  
e-mail: flchang@sdu.edu.cn

X. Sun  
e-mail: sxy@sdjzu.edu.cn

X. Sun  
School of Computer Science and Technology, Shandong Jianzhu University, Jinan, China

## 1 Introduction

Intelligent video surveillance has been one of the most active research areas in computer vision, with a wide variety of applications both in public and private environments, such as homeland security, crime prevention, traffic control, accident prediction and detection, and monitoring patients, elderly and children at home. These applications require monitoring indoor and outdoor scenes of airports, train stations, highways, parking lots, stores, shopping malls, and offices [1, 2]. Since a single camera has a limited field of view, multiple cameras, even a distributed network of video sensors, have to be employed to monitor activities over a complex area nowadays. Hence the continuous object tracking across multiple cameras and consistent labeling of objects between adjacent cameras is urgently needed. Continuously tracking objects across cameras and consistent labeling of objects between adjacent cameras is usually termed as “object handover.” The objective of object handover is to maintain the identity of moving objects when they are traveling from one camera to another. More specifically, when an object appears in one camera, we need to determine whether it has previously appeared before in other cameras or is a new object [3]. Moreover, it is usually not feasible to completely cover large areas with cameras having overlapping views due to economic and/or computational reasons. Thus, in realistic scenarios, the system should be able to handle multiple cameras with non-overlapping fields of view. In this paper, we use particle filter to solve object handover with multiple cameras having non-overlapping fields of view.

Possible cues for tracking across cameras include appearance information and spatial–temporal information. Thus, existing algorithms of camera handoff can be classified as: (1) Camera handover based on environment information, such as dealing with the handoff problem based on 3D environment model and calibrated cameras [4], tracking objects on a single global ground plane [5], or based on the homography between camera views [6], the limits of fields of view to establish correspondence between objects in multiple cameras [7]. All these handoff methods need either the common fields of view or camera calibration, which restricts their applications. (2) Camera handoff based on object model. Appearance is the inherent feature of object, so using one cue or multiple cues of object to construct the object model is a more popular method of camera handoff, especially in cases where spatial–temporal reasoning is not feasible or accurate [8]. The most widely used cues are spatial position, shape, color, intensity and motion. A lot of research work has been done on this field only using visual information of appearance to associate objects, which can be named as object re-identification [9, 10]. However, the appearance cues are easily fouled by many factors, such as illumination, view angles, object deformation, etc. (3) Hybrid methods. Combining appearance information with spatial–temporal information into Bayesian theory or other fusion methods is more common since these two cues have their own defects and merits. [11] integrates spatial position, shape, and color information to track object blobs in single camera tracking, calibrates two fixed cameras using five

coplanar control points, then tracks objects across them using Extended Kalman Filter (EKF) to handle occlusion. [12, 13] learn the spatial–temporal relationship of non-overlapping cameras and the appearance relationship of objects, then track them, while the computation speed is slow for learning [14] using a map of the surveillance area, which provides information about possible path trajectories across the non-overlapping region in particle filter to help track prediction when the subject leaves the field of view of a camera and enters another camera.

In order to solve the camera handoff of multiple cameras with disjoint view, a camera handoff method using particle filter is proposed in this paper. We integrated appearance information with spatial–temporal information to label consistent objects in multiple cameras. The spatial–temporal information includes the source, sink regions among cameras, their transition probabilities, and transition time. In this paper, these spatial–temporal cues are learned offline. Based on these spatial–temporal cues particles are sampled and adapted till the correct camera handover is found. Histogram in HSV color space is used to compute the fitness of particle. This algorithm is called particle filter handover. The work most similar with ours is [14], who also use particle filter to resolve the problem of within field of view visual tracking and track prediction when targets leave the field of view of any camera; however, they do not involve the spatial–temporal cues, and their applied environment is restricted in in-house scene.

The remainder of the paper is organized as follows: [Section 2](#) introduces our tracking initialization and intra-camera object detection and tracking, [Sect. 3](#) discusses the camera handover scheme based on particle filter, experiments results are given in [Sects. 4, 5](#) conclude the paper.

## 2 Intra-camera Object Detection and Tracking

Tracking objects continually across cameras involves mainly two parts: Intra-camera object tracking and inter-camera tracking. The first, tracking objects within a camera view is the basic following process, so we discuss our tracking initialization and intra-camera object detection and tracking scheme in this section.

### 2.1 Tracking Initialization

Tracking initialization has been little investigated although it is very important. Most works set the initial target by drawing a rectangle to enclose the object in the first frame. Our initialization can realize an automatic target setting, in which one can directly input the description of target, for example, “a person dressed in red shirt and black pants,” or else one can set the target by friendly man–machine interface, only using drawing or filling. The reason we do this is because we first detect a moving foreground, and then find the assigned object in all moving blocks

using color matching. When comparing the color of moving objects with that of assigned target, we quantified the HSV color space into 53 pieces based on fussy logic [15] and named them like Joost [16]. Hue slice is quantified into 12 pieces defined as: red, orange, yellow, green and yellow, green, blue, cyan, blue, purple, magenta, pink, and red. Saturation S is divided into three parts, taking green for example, the three parts are: white, light green, and green. Value V is divided into three parts, also taking green for example, the three parts are: black, dark green, green. If saturation S is less than the threshold range of 0.16, all the colors are degenerated to grayscale, so this region is quantified alone to five parts: black, dark gray, gray, light gray, and white. Thus the HSV color space is quantified as  $12 * 2 * 2 + 5 = 53$  sub-regions. Considering the vagueness of color border, we use Trapezium membership function in fuzzy quantification.

## 2.2 Particle Filter-Based Intra-camera Object Tracking

Once the original target is determined, particle filter is used to track the object in single camera. The principle of particle filter is derived from the sequential Monte Carlo method [17] that recursively generates random measurements to approximate the distribution of unknown variables. Particle filter technique has proved to be robust and is widely used in many applications, especially object tracking [1, 18]. The principle of particle filter-based tracker is maintaining a probability distribution over the state (location, scale, etc.) of the object being tracked [18]. Particle filters represent this distribution as a set of weighted particles. In general, tracking using cameras is a nonlinear problem, thus the dynamic model is expressed as (1)

$$\begin{aligned} x_k &= f_k(x_{k-1}, u_k) \\ z_k &= h_k(x_k, v_k). \end{aligned} \quad (1)$$

In (1),  $k$  is time stamp,  $x_k$  is the state sequences, and  $z_k$  represents the observation measurement. The notation  $u_k$  and  $v_k$  are noise matrices. The first equation is the function that predicts the current state given the previous state and a noise vector. The second equation defines a measurement model, i.e., it determines how well the prediction of current state based on current observation. Hence particles can be denoted by  $\left\{ x_{1:k}^{(i)}, w_k^{(i)} \right\}_{i=1}^N$ , where  $i$  is the particle index,  $N$  is the total number of particles,  $x_{1:k}^{(i)}$ , and  $w_k^{(i)}$  represent the predicted state and weight of particle  $i$ . Each particle is a guess representing one possible location of the object being tracked. The set of particles contains more weight at locations where the object being tracked is more likely to be. This weighted distribution is propagated through time using a set of equations known as Bayesian filtering equations, and the trajectory of the tracked object can be drawn from the density  $p(x_{0:k}|z_{1:k})$ , which is expressed in (2)

$$p(x_{0:k}|z_{1:k}) \approx \sum_{i=1}^N w_k^{(i)} \delta(x_{0:k} - x_{0:k}^{(i)}). \quad (2)$$

Then the estimation of the state observation can be determined by taking the weighted mean of the particle set at each time step, denoted by (3), where  $h(\cdot)$  is the observation of statement.

$$E(h(x_{1:k})) = \sum_{i=1}^N w_k^{(i)} h(x_{1:k}^{(i)}). \quad (3)$$

A general framework of particle filter algorithm can be described as:

1. Initialize: a particle set of  $N$  particles.
2. Prediction: for each particle using second order auto regressive dynamics in (4), where  $\omega_{k-1}^{(i)}$  is the noise and  $\omega_{k-1}^{(i)} \sim N(0, \sigma^2)$ .

$$x_k^{(i)} - x_{k-1}^{(i)} = x_{k-1}^{(i)} - x_{k-2}^{(i)} + \omega_{k-1}^{(i)}. \quad (4)$$

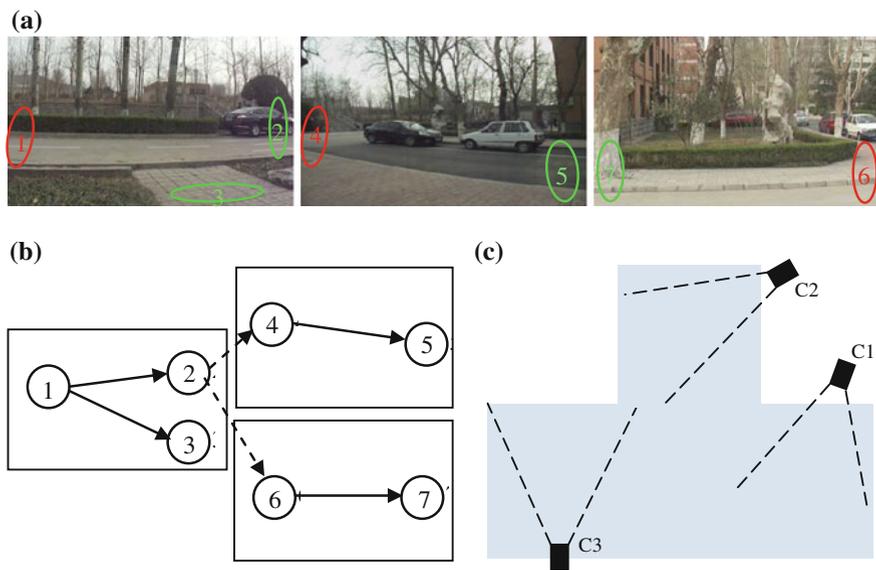
3. Importance: weighting each particle based on current observation by histogram distance.
4. Resample: a process to maintain the diversity of all particles according to their re-normalized weights.

### 3 Camera Handover Based on Particle Filter

The objective of camera handover is to maintain the identity of moving objects when they are traveling from one camera to another, it is parallel with maintaining the identity of moving objects when they are traveling from one frame to another in certain extent, so we consider using particle filter in camera handover as in inter-frame tracking. When the object is traveling from one camera to another, although the sights and illuminations of cameras are quite different, the appearance of object remains consistent. For example, the color and texture of object considered as not changing excludes the impact of light, the velocity of subject remains constant in two adjacent cameras if the angle change of two cameras is small. Thus we employ a spatial-temporal progressive matching scheme in particle filter among adjacent cameras, which extracts the template and velocity of target in prior camera and fuses them into matching the next camera. First we learn the spatial-temporal relationship of camera network in Sects. 3.1, and 3.2 describes the algorithm of camera handover using particle filter.

### 3.1 Spatial–Temporal Information of Inter-camera

We learn the spatial–temporal relationship among adjacent cameras from video sequences of a long time. We group locations where objects appear (source region) and disappear (sink region) by k-means clustering, and the sources and sinks are represented by two-element Gauss distribution. The spatial transition probability and transition time of each pair of source and sink can be calculated by statistics. The spatial transition probability is defined as  $P(S_k^i, S_l^j)$ , where  $S_k^i$  is the  $k$ th source region of camera  $i$  and  $S_l^j$  is the  $l$ th sink region (or exit zone) of camera  $j$ , which can be represented by a digraph in Fig. 1. Figure 1a shows the original sources and sinks, dark and bright ellipses represent source and sink, respectively, in one way. Figure 1b is the digraph, and Fig. 1c is the topology of cameras. In Fig. 1b, vertexes represent source and sink regions, verges represent the transition of source and sink, and the weight of verge is the transition probability of each pair of source and sink. The Transition time is represented by  $P_{S_k^i, S_l^j}(t)$ . Each  $P(t)$  is modeled as a mixture of Gaussian distribution. In this paper, we choose  $K$  as 3, three Gaussian distributions correspond to people walking slowly, at normal speed, and walking quickly. The probability of transition time  $t$  is as follows



**Fig. 1** Spatial–temporal constraint of inter-camera. **a** Original source and sinks of three cameras C1, C2 and C3. **b** Digraph representation. **c** Topology of C1, C2 and C3

$$P(t) = \sum_{j=1}^3 \omega_j * n_j(t, \mu_j, \sigma_j) \quad (5)$$

$$n_j(t, \mu_j, \sigma_j) = \frac{1}{\sqrt{2\pi} \cdot \sigma_j} e^{-\frac{(t-\mu_j)^2}{2\sigma_j^2}}$$

where  $\omega_j$  is the weight of the  $j$ th Gaussian in the mixture, which can be interpreted as a prior probability of the random variable generated by the  $j$ th Gaussian distribution,  $\mu_j$  and  $\sigma_j$  are the mean value and standard deviation of the  $j$ th Gaussian,  $n_j$  is the Gauss probability density function.

### 3.2 Description of Camera Handoff Based on Particle Filter

In order to get effective camera handover, we employ a spatial-temporal progressive matching scheme in particle filter among adjacent cameras. Taking camera handover in Fig. 1 as example, the steps of camera handover are described below in detail.

- S1. When the target goes into the sink region of C1, extract the object model  $H$  and speed  $V_h$ , push the object model  $H$  into targetlist, start the handover judgment. For maximize repression the influence of illumination changes, histogram in HSV space is used in the object model.
  - S1.1 according to the probability of transition time  $P_{S_k^i S_l^j}(t)$  and spatial transition probability  $P(S_k^i S_l^j)$ , scatter particles in source regions of other cameras corresponding to the sink region of C1 (that is, region 4 in C2 and 6 in C3).
  - S1.2 calculates the similarity of all particles with the object model  $H$  according to histogram Intersection in (6).

$$w_k^{(i)} = d[h_k^{(i)}, H_k] = \sum_{j=1}^m \min(h_{kj}^{(i)}, H_{kj}) \quad (6)$$

where  $h_k^{(i)}$  is the histogram of  $i$ th particle in HSV color space of  $t$  time,  $h_k$  is the object template of  $k$  moment.

- S1.3 If all similarities are less than a certain threshold  $T_S$ , judge the target that does not emerge in other cameras, in the next time step go S1.1 continue scatter particles based on the spatial-temporal probability. Once the similarity of some particle appears larger than threshold  $T_S$ , which means target appears, go S2 particle filter handover.
- S2 Particle filter handover.
  - S2.1 motion predictions for all particles in regions 4 and 6 separately at the same time. Motion model is shown in (7) and (8).

$$v_k = \alpha \cdot V_h + (1 - \alpha) \cdot v_{k-1} + G_{k-1} \quad (7)$$

$$x_k = x_{k-1} + v_k \quad (8)$$

where  $V_k$  is the speed of  $k$  time step,  $V_h$  is the object speed in prior camera,  $G_{k-1}$  is two-element Gauss random noise, and  $\alpha$  is fusion coefficient.

- S2.2 computer the weights  $w_k^{(i)}$  of all particles according to histogram Intersection coefficient in (6) and normalized according to (9).

$$w_k^{(i)} = \frac{1}{\sum_i^N w_k^{(i)}} \quad (9)$$

In order to improve the accuracy, we use (10) to renew the template  $H_t$  when computing the weights.

$$H_k = \beta \cdot H + (1 - \beta) \cdot H_{k-1} \quad (10)$$

where  $H$  is the object model of the prior camera,  $H_k$  is the template of  $k$  moment, and  $\beta$  is fusion coefficient.

- S2.3 calculate the weight sums of each source, if the sum of some source is larger than others, increase the particle number of this source and decrease particle number of other sources correlated with the sink. Taking Fig. 1 for example, if the weight sum of source region 4 in C2 is larger than that of region 6 in C3, increase particle number of region 4, decrease number of region 6 at the same time, so that the total particle number is constant.
- S2.4 resample particles according the weights and numbers of each sources.
- S2.5 go to S2.1 recursive motion prediction and increase/decrease the particles. If the number of some source decreases to less than 5 % of the total number, set it to 0. If the particle number of some source is more than 95 % of the total number, or particle numbers of all the other sources are 0, the object in current source is judged as the correct handover, and camera handover is completed.
- S2.6 After particle filter handover is started for a long time  $T$ , there is no particle's similarity larger than threshold  $T_S$  or the case that one particle's similarity larger than  $T_S$  remains short time, we judged the camera handover as failed. Object detection in subsequent cameras is restarted. If some object is detected, compare it with targetlist, judge it as new subject if its model is not similar to any target in targetlist, and then add it to targetlist.
- S3 In the next process, object tracking in single camera view keep going on.

In the above steps, fusion coefficients  $\alpha$  and  $\beta$  are two important parameters, which determine to what extent is the template consistent with prior camera and how fast it adapts to the current scene. The impact of prior template is slowly

reduced with time, so we name this scheme as spatial–temporal progressive matching.

### 3.3 Mathematical Description of Particle Filter Handover

In (2), the trajectory of the tracked object can be drawn from the density  $p(x_{0:k}|z_{1:k})$ . In fact, the density in multiple cameras can be expressed by

$$p(x_{0:k}|z_{1:k}) \approx \sum_{j=1}^C \sum_{i=1}^{N_j} w_k^{(ji)} \delta(x_{0:k} - x_{0:k}^{(ji)}) \quad (11)$$

in which,  $C$  is the number of source regions related to the sink region. If we denote density of each source region as  $p_j(x_{0:k}|z_{1:k})$ , S2.3 in Sect. 3.2 increases the particle number of region with  $\max(p_j(x_{0:k}|z_{1:k}))$ , that is, increases the density of maximal likelihood; decreases that of other source regions, until it finds the correct handover.

## 4 Experiments

In this part, the particle filter handover scheme above is tested by an outdoor scenario taken by us. The experimental environment is: CPU, Intel Pentium 2.93G Hz, and memory, 3.25 GB, developing tools are Visual Studio 2008 and OpenCV 2.3.

In our outdoor scenario, the topology of cameras is shown in Fig. 1, in which the three cameras have no or little overlap view, so the homography-based or limits of fields of view-based methods cannot be used. Some frames of C1, C2, and C3 are shown in Fig. 2, from which one can see that the illuminations of three cameras are quite different for different camera types and views of angles. So the appearance-based method is not reliable. The only effective way is combining appearance information with spatial–temporal information.

The space–time information of three cameras is first learned using a 2 h dataset. In our test scenario, a person passes through C1, then exits from C1 and enters into C2 after 266 frames, not C3. The types of these three cameras are different, and the



Fig. 2 Some frames of three cameras



Fig. 3 Particle distribution in frames target first emerged

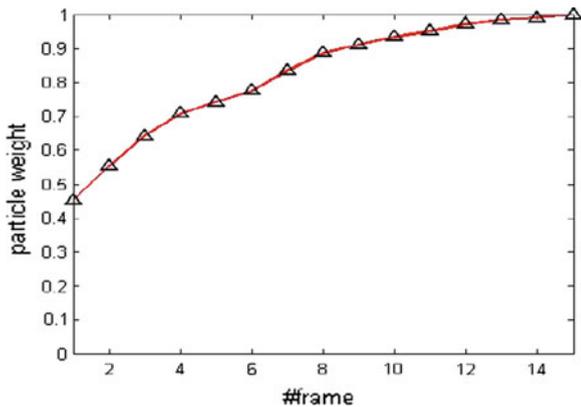
resolutions are also different: Resolutions of C1 and C2 are 320\*240, and that of C3 is 640\*480, so normalization should be done first.

When the target enters into sink region 2 in C1, the camera handover judgment is started, and particles are sampled according to the spatial-temporal constraints. In the first few frames of C2 and C3 into which target did not enter, the camera handover is not started as the similarities of particles with object model  $H$  in prior camera are small. Continue scatter particles in next frame. Once the similarity of some particle appears larger than threshold, it means the target goes in (in this experiment, first similarity larger emerges in frame 266 of C2), does not scatter

Fig. 4 Camera handover successfully completed in frame 271 of C2



Fig. 5 Weight change of particles in C2



particles in the next frame, and switches to the particle filter camera handover. The particle state of frame where target first emerged in source regions 4 and 6 is shown in Fig. 3 where the dots represent a particle. The motion prediction, resample based on weights and number adjustment of particles are carried out with the time. Another parameter setting is:  $T_S = 16$ ,  $\alpha = 0.9$ ,  $\beta = 0.95$ . Experiment results show that in 15 frames after the target first emerged, all particles are transferred into C2 and focus on the target, which indicates the handover is successfully completed. Figure 4 is the handover ending in frame 271 of C2, in which the pixels are particle positions and the rectangle is the estimation of object location. The weight sum change of particles in 15 frames of C2 is given in Fig. 5.

## 5 Conclusion

In multiple camera intelligent surveillance, continuously object tracking across cameras, i.e., object handover is an unresolved problem and has attracted many research works in recent years. This paper proposed a particle filter camera tracking and handover method to solve the continuous tracking across multiple cameras with non-overlapping fields of view. The spatial-temporal constraints, including source and sink regions, transition probabilities, and transition time of each pair of source and sink are learned offline. Then a spatial-temporal progressive matching scheme is used to scatter particles and renew the number of particles in source regions according to the spatial-temporal information among adjacent cameras. Experiment shows that this particle filter camera handover scheme is effective. Further and deeper employment should be done to extend this scheme to camera handover and camera management in more complicate camera networks.

**Acknowledgments** We would like to acknowledge the support from China Natural Science Foundation Committee (No. 60975025, No. 61273277), the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry (20101174), Shandong province Natural Science Foundation Committee (No. ZR2011FM032), and the University Innovation Program from Jinan Science and Technology Bureau (No. 201004002).

## References

1. Yilmaz A, Javed O, Shah M (2006) Object tracking: a survey. *ACM Compu Surv* 38(4):45
2. Wang X (2012) Intelligent multi-camera video surveillance: a review. *Pattern Recogn Lett* pp 1–25
3. Cai Y, Chen W, Huang K, Tan T (2007) Continuously tracking objects across multiple widely separated cameras ACCV. *LNCS* 4843:843–852
4. Focken D, Stiefelhagen R (2002) towards vision-based 3d people tracking in a smart room. In: *Proceedings of IEEE international conference multimodal interfaces*

5. Chang T-H, Gong S (2001) Tracking multiple people with a multi-camera system. In: Proceedings of IEEE workshop on multi-object tracking
6. Khan SM, Shah M (2006) A multiview approach to tracking people in crowded scenes using a planar homography constraint. ECCV 2006
7. Omar J, Sohaib K, Zeeshan R, Mubarak S (2000) Camera handoff: tracking in multiple uncalibrated stationary cameras. In: Proceedings workshop on human motion 2000
8. Nicholas J, Redding J, Fabian O, Judd K, Tristrom C (2008) Cross-matching via feature matching for camera handover with non-overlapping fields of view. In: Proceedings of digital image computing: techniques and applications
9. Omar J, Khurram S, Mubarak S (2005) Appearance modeling for tracking in multiple non-overlapping cameras. In: Proceedings of IEEE international conference computer vision and pattern recognition (CVPR)
10. Doretto G, Thomas S, Peter T, Jens R (2011) Appearance-based person reidentification in camera networks: problem overview and current approaches. *J Ambient Intell Humaniz Comput* 2:127–151
11. Zhou Q, Aggarwal JK (2006) Object tracking in an outdoor environment using fusion of features and cameras. *Image Vis Comput* 4(11):1244–1255
12. Kuo C-H, Huang C, Nevatia R (2010) Inter-camera association of multi-target tracks by on-line learned appearance affinity models. ECCV 2010 Part I, LNCS 6311
13. Chen K-W, Lai C-C, Lee P-J, Chen C-S, Hung Y-P (2011) Adaptive learning for target tracking and true linking discovering across multiple non-overlapping cameras. *IEEE Trans Multimed* 13(4):625–638
14. Lim FL, Leputra W, Tan T (2007) Non-overlapping distributed tracking system utilizing particle filter. *J VLSI Sig Proc* 49:343–362
15. Cheng M-Y, Wang C-K (2007) Dynamic visual tracking based on multi-cue matching. In: Proceedings of international conference on mechatronics
16. Weijer Jvd, Schmid C, Verbeek J (2007) Learning color names from real-world images. In: Proceedings of IEEE international conference computer vision and pattern recognition (CVPR)
17. Bolic M (2004) Architectures for efficient implementation of particle filters. Ph.D. thesis, electrical engineering, Stony Brook University, August 2004. <http://www.sigproc.eng.cam.ac.uk/smc/smcpapers.html>
18. Isard M, Blake A (1998) Condensation-conditional density propagation for visual tracking. *Int J Comput Vision* 29(1):5–28

# Analyzing Effects of Pressing Radio Button on Driver's Visual Cognition

Huacai Xian, Lisheng Jin, Haijing Hou, Qingning Niu  
and Huanhuan Lv

**Abstract** An approach is presented based on driver simulator and SmarteyeII eye tracking system to examine the effects of pressing in-vehicle radio button on driver's visual cognition. Parameters of glance frequency, glance duration, eye movement speed, and visual line moving in different regions of interest (ROIs) in task of pressing the radio button, closely related with driver's visual cognition, were collected and analyzed. Based on the experimental data, driver's visualization model with secondary tasks was built by CogTool. Driver's vision, eye movement, cognition, and hand motion were tracked and recorded by the model. Results of experiment and running model show that pressing the in-vehicle radio button while driving has adverse influence on driver's visual cognition and occupies a lot of the driver's visual resources.

**Keywords** In-vehicle radio · Secondary task · Visual cognition · Visualization model · CogTool

## 1 Introduction

Amelia Stenson reports that 90 % of causes of traffic accidents are related to driver's factor [1]. Some researchers consider that 80 % of the driver's perception of information origins from vision. Thus, it is put in a significant place while driving and accident controlling [2, 3]. Thus, it is of remarkable significance to study the driver's visual cognition with the secondary task which is an important factor for driver distraction. Secondary tasks are tasks that the driver voluntarily or involuntarily engages in which do not directly pertain to the primary task of safe vehicle operation [4], such as taking a look at the navigation system, listening to

---

H. Xian (✉) · L. Jin · H. Hou · Q. Niu · H. Lv  
Transportation and Traffic College, Jilin University, ChangChun 130024, China  
e-mail: xianhuacai1983@163.com

music or radio, talking on the phone, etc. These tasks occupy driver's visual resource, cognitive resource, and motion resource in different degrees and distract the driver's attention simultaneously as well. As the most important factor for safe driving, the driver's visual cognition remains largely unexplored, which is the focus of this study.

Most previous studies paid close attention to eye movements and driving performance while using cell phone [5–7], music and music tempo [8–10], text messaging, and device position while driving [11, 12]. However, few have made detailed and thorough analysis on driver's visual cognition in secondary tasks while driving. Therefore, it is of important significance to study the driver's visual cognition in secondary task.

With this motivation in mind, the overall goal of this paper is to develop a method of examining the effects of pressing radio button on driving performance. In particular, it has the following purposes:

- (1) Collect parameters of drivers' visual cognition behavior, taking an in-vehicle radio as the secondary task while driving and coordinating Driving Simulator with SmarteyeII system, which is separated from driver's head and thus relieves discomfort.
- (2) Use SPSS and Matlab for analyzing parameters of driver's glance frequency, glance duration, eye movement speed, and visual line moving.
- (3) Build driver's visualization model in task and achieve prediction of driver's vision, eye movement, cognition, and hands motion using CogTool.

## 2 Description of Experiment

### 2.1 Experiment Equipments

The experiment equipment mainly includes the SmarteyeII system, an in-vehicle radio, and a platform composed of a BESTURN B50 car and a Driving Simulator.

#### (1) Secondary task

The secondary task was performed on the in-vehicle radio button. The center of radio-button coordinate was (29, -23, -13.5) and  $23^\circ$  away from the steering wheel center (see Fig. 1).

#### (2) Eyes tracking system

Eyes tracking system of the Smarteye was mainly responsible for capturing driver's eye movement during the whole driving procedure. The Smarteye used four cameras (shown in Fig. 1) mounted in front of the windscreen towards driver's face to capture driver's eye movement. Besides, to pick driving environment out of the vehicle, three panoramic cameras were also mounted between two back seats.

**Fig. 1** In-vehicle radio and SmarteyeII



### (3) Driving simulator

Driving simulator is composed of a BESTURN B50 car, modules of dynamics simulation, image simulation, and monitor simulation (see Fig. 2).

## 2.2 Participants

Eight participants (six male and two female) between the ages of 20 and 23 (mean = 21.5, SD = 1.18) were recruited from the University of Jilin. All participants were required to hold a valid class C driver's license for more than 2 years and drive a minimum of 5,000 km per year. Drivers were also required to be in good health.

## 2.3 Experimental Design

The performing course was divided into 13 steps. A four-lane divided highway with an 80–120 km/h velocity limit was chosen. Each lane of the highway was 3.75 m wide and the middle belt was 3 m wide. The experiment was performed in multi-vehicles and non-vehicle environment, respectively, in the following steps:

**Fig. 2** Driving environment



- (1) Participants were taught to operate the radio and then practice performing the task and driving the simulator under multi-vehicles and non-vehicle road conditions repeatedly, first separately and then together.
- (2) Lead driver to complete secondary tasks at any time they feel safe while driving. Each driver completed a drive without performing the in-vehicle radio task and a drive with in-vehicle radio task under different highway road conditions, after which parameters of driver's eye behavior were recorded.
- (3) Driver's visual data was collected while driving and then analyzed.

### 3 Analysis of Drivers' Visual Cognition Behavior

Eye movement was collected by SmarteyeII Recorder. A glance is defined as consecutive focuses on an area of interest, not including saccade transition and blinking behavior. As four regions of interest (ROI), the road, in-vehicle radio, and two rearview mirrors were extracted in the experiment for further research.

Figure 3 shows the tracking of eye movement. From Fig. 4, four ROIs are divided by rectangles of different colors. ROI-0 (blue area) and ROI-3 (green area) are areas of on-road and in-vehicle radio conditions. ROI-1 (red area) and ROI-2 (light yellow area) stand for the right and left rear mirrors respectively.

#### 3.1 Mean Glance Frequency

Glance frequency is defined as the number of glances at a target during the task where each glance is separated by at least one glance to a different target [13]. Apart from glances at the lane, the more intensive the driver's gaze points is, the more visual resources the area occupies, resulting in driver's visual distraction.

The glance frequency distribution without and with secondary tasks on multi-vehicles road is shown in Figs. 5 and 6 respectively. Comparing the ROI0 and ROI1, it can be perceived that the number of glances on rearview mirrors with radio task is

**Fig. 3** Eye movement tracking

