

Jiadong Sun
Jingnan Liu
Shiwei Fan
Xiaochun Lu
Editors

China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III



Lecture Notes in Electrical Engineering

Volume 342

Board of Series editors

Leopoldo Angrisani, Napoli, Italy
Marco Arteaga, Coyoacán, México
Samarjit Chakraborty, München, Germany
Jiming Chen, Hangzhou, P.R. China
Tan Kay Chen, Singapore, Singapore
Rüdiger Dillmann, Karlsruhe, Germany
Haibin Duan, Beijing, China
Gianluigi Ferrari, Parma, Italy
Manuel Ferre, Madrid, Spain
Sandra Hirche, München, Germany
Faryar Jabbari, Irvine, USA
Janusz Kacprzyk, Warsaw, Poland
Alaa Khamis, New Cairo City, Egypt
Torsten Kroeger, Stanford, USA
Tan Cher Ming, Singapore, Singapore
Wolfgang Minker, Ulm, Germany
Pradeep Misra, Dayton, USA
Sebastian Möller, Berlin, Germany
Subhas Mukhopadhyay, Palmerston, New Zealand
Cun-Zheng Ning, Tempe, USA
Toyoaki Nishida, Sakyo-ku, Japan
Bijaya Ketan Panigrahi, New Delhi, India
Federica Pascucci, Roma, Italy
Tariq Samad, Minneapolis, USA
Gan Woon Seng, Nanyang Avenue, Singapore
Germano Veiga, Porto, Portugal
Haitao Wu, Beijing, China
Junjie James Zhang, Charlotte, USA

About this Series

“Lecture Notes in Electrical Engineering (LNEE)” is a book series which reports the latest research and developments in Electrical Engineering, namely:

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering

LNEE publishes authored monographs and contributed volumes which present cutting edge research information as well as new perspectives on classical fields, while maintaining Springer’s high standards of academic excellence. Also considered for publication are lecture materials, proceedings, and other related materials of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of electrical engineering.

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer’s other Lecture Notes series, LNEE will be distributed through Springer’s print and electronic publishing channels.

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

Jiadong Sun · Jingnan Liu
Shiwei Fan · Xiaochun Lu
Editors

China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III



Editors

Jiadong Sun
China Aerospace Science
and Technology Corporation
Chinese Academy of Sciences
Beijing
China

Jingnan Liu
Wuhan University
Wuhan
China

Shiwei Fan
China Satellite Navigation Office
Beijing
China

Xiaochun Lu
Chinese Academy of Sciences
Beijing
China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-3-662-46631-5 ISBN 978-3-662-46632-2 (eBook)
DOI 10.1007/978-3-662-46632-2

Library of Congress Control Number: 2014937269

Springer Heidelberg New York Dordrecht London
© Springer-Verlag Berlin Heidelberg 2015

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

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

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

Printed on acid-free paper

Springer-Verlag GmbH Berlin Heidelberg is part of Springer Science+Business Media
(www.springer.com)

Editorial Board

Topic 1: BDS/GNSS Navigation Applications

Yamin Dang, Chinese Academy of Surveying and Mapping, China

Jianping Cao, Research Institute of Air Force Equipment, China

Jing Li, Communication and Information Center of Ministry of Transport, China

Shuanggen Jin, Shanghai Astronomical Observatory, Chinese Academy of Sciences, China

Yanming Feng, Queensland University of Technology Brisbane, Australia

Topic 2: Satellite Navigation Signal System, Compatibility and Interoperability

Feixue Wang, National University of Defense Technology, China

Xingqun Zhan, Shanghai Jiaotong University, China

Guangxia Li, PLA University of Science and Technology

Baowang Lian, Northwest University

Topic 3: Precise Orbit Determination and Positioning

Qile Zhao, Wuhan University, China

Xiaogong Hu, Shanghai Astronomical Observatory, Chinese Academy of Sciences, China

Geshi Tang, Beijing Aerospace Control Center (BACC), China

Maorong Ge, Geological Research Center in Germany

Chris Rizos, University of New South Wales International Association of Geodesy International GNSS Service, Australia

Topic 4: Atomic Clock Technique and Time-Frequency System

Lianshan Gao, Beijing Institute of Radio Metrology and Measurement, China

Chunhao Han, Beijing Satellite Navigation Center, China

Jinjun Zheng, China Academy of Space Technology (CAST), China

Xiaohui Li, National Time Service Center, Chinese Academy of Sciences, China

Zhiheng Jiang, Time Department Bureau International Des Poids et Mesures, France

Topic 5: Satellite Navigation Augmentation and Integrity Monitoring

Jinping Chen, Beijing Satellite Navigation Center, China

Junlin Zhang, OLinkStar Co., Ltd., China

Jianwen Li, Surveying and Mapping Institute of Zhengzhou, China

Rongzhi Zhang, Xi'an Satellite Control Center, China

Yang Gao, University of Calgary, Canada

Topic 6: BDS/GNSS Test and Assessment Technology

Jun Yang, National University of Defense Technology, China

Henglin Chu, Beijing Satellite Navigation Center, China

Baoguo Yu, The 54th Research Institute of China Electronics Technology Group Corporation, China

Wenxian Yu, Shanghai Jiaotong University, China

Jinling Wang, University of New South Wales, Australia

Topic 7: BDS/GNSS User Terminal Technology

Haibo He, Beijing Satellite Navigation Center, China

Hong Li, Tsinghua University, China

Yongchao Geng, The 20th Research Institute of China Electronics Technology Group Corporation, China

Sam Pullen, Stanford University

Yan Liu, Guangzhou Haige Communications Group Incorporated Company, China

Topic 8: Satellite Navigation Model and Method

Qin Zhang, Chang'an University, China

Luping Xu, Xidian University

Xiaolin Jia, Xi'an Institute of Surveying and Mapping, China

Kefei Zhang, RMIT University, Australia

Shaojun Feng, Imperial College London, England

Topic 9: PNT System and New Technologies of Navigation

Mingquan Lu, Tsinghua University, China

Zhongliang Deng, Beijing University of Posts and Telecommunications, China

Hong Yuan, Navigation Headquarters, Chinese Academy of Sciences, China

Dangwei Wang, The 20th Research Institute of China Electronics Technology Group Corporation, China

Yongbin Zhou, National University of Defense Technology, China

Ping Shuai, China Academy of Space Technology, China

Topic 10: Policies and Regulations, Standard and Intellectual Property

Daiping Zhang, China Defence Science and Technology Information Center, China

Yonggang Wei, China Academy of Aerospace Standardization and Product Assurance, China

Ping Zhang, Peking University, China

Haibo Liu, Institute of Policy and Management, Chinese Academy of Science, China

Preface

BeiDou Navigation Satellite System (BDS) is China's global navigation satellite system which has been developed independently. BDS is similar in principle to the global positioning system (GPS) and compatible with other global satellite navigation systems (GNSS) worldwide. The BDS will provide highly reliable and precise positioning, navigation and timing (PNT) services as well as short-message communication for all users under all-weather, all-time and worldwide conditions.

China Satellite Navigation Conference (CSNC) is an open platform for academic exchanges in the field of satellite navigation. It aims to encourage technological innovation, accelerate GNSS engineering and boost the development of the satellite navigation industry in China and in the world.

The 6th China Satellite Navigation Conference (CSNC 2015) was held during May 13–15, 2015, Xian, China. The theme of CSNC2015 is Opening-up, Connectivity, Win-win, which covers a wide range of activities, including technical seminars, academic exchanges, forums, exhibitions and lectures. The main topics are as follows:

1. BDS/GNSS Navigation Applications
2. Satellite Navigation Signal System, Compatibility and Interoperability
3. Precise Orbit Determination and Positioning
4. Atomic Clock Technique and Time-Frequency System
5. Satellite Navigation Augmentation and Integrity Monitoring
6. BDS/GNSS Test and Assessment Technology
7. BDS/GNSS User Terminal Technology
8. Satellite Navigation Models and Methods
9. PNT System and New Technologies of Navigation
10. Policies and Regulations, Standard and Intellectual Property

The proceedings (Lecture Notes in Electrical Engineering) have 197 papers in ten topics of the conference, which were selected through a strict peer-review process from 513 papers presented at CSNC2015. In addition, another 251 papers were selected as the electronic proceedings of CSNC2015, which are also indexed

by “China Proceedings of Conferences Full-text Database (CPCD)” of CNKI and Wan Fang Data.

We thank the contribution of each author and extend our gratitude to the 215 referees and 49 session chairs who are listed as members of the editorial board. The assistance of the organizing committees of CNSC2015 and the Springer editorial office is highly appreciated.

Jiadong Sun
Chair of CSNC2015

The 6th China Satellite Navigation Conference (CSNC 2015)

Scientific Committee

Chairman

Jiadong Sun, China Aerospace Science and Technology Corporation

Vice-Chairman

Rongjun Shen, China

Jisheng Li, China

Qisheng Sui, China

Zuhong Li, China Academy of Space Technology

Shusen Tan, Beijing Satellite Navigation Center, China

Executive Chairman

Jingnan Liu, Wuhan University

Yuanxi Yang, China National Administration of GNSS and Applications

Shiwei Fan, China

Committee Members: (By Surnames Stroke Order)

Xiancheng Ding, China Electronics Technology Group Corporation

Qingjun Bu, China

Liheng Wang, China Aerospace Science and Technology Corporation

Yuzhu Wang, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences

Guoxiang Ai, National Astronomical Observatories, Chinese Academy of Sciences

Shuhua Ye, Shanghai Astronomical Observatories, Chinese Academy of Sciences

Zhaowen Zhuang, National University of Defense Technology
Qifeng Xu, PLA Information Engineering University
Houze Xu, Institute of Geodesy and Geophysics, Chinese Academy of Sciences
Guirong Min, China Academy of Space Technology
Xixiang Zhang, China Electronics Technology Group Corporation
Lvqian Zhang, China Aerospace Science and Technology Corporation
Junyong Chen, National Administration of Surveying, Mapping
and Geoinformation
Benyao Fan, China Academy of Space Technology
Dongjin Luo, China
Guohong Xia, China Aerospace Science and Industry Corporation
Chong Cao, China Research Institute of Radio Wave Propagation (CETC 22)
Faren Qi, China Academy of Space Technology
Sili Liang, China Aerospace Science and Technology Corporation
Shancheng Tu, China Academy of Space Technology
Rongsheng Su, China
Zhipeng Tong, China Electronics Technology Group Corporation
Ziqing Wei, Xi'an Institute of Surveying and mapping

Organizing Committee

Secretary General

Haitao Wu, Navigation Headquarters, Chinese Academy of Sciences

Vice-Secretary General

Wenhai Jiao, China Satellite Navigation Office Engineering Center
Yuehu Hu, The 20th Research Institute of China Electronics Technology Group
Corporation
Weina Hao, Navigation Headquarters, Chinese Academy of Sciences

Committee Members: (By Surnames Stroke Order)

Qun Ding, Xi'an Institute of Navigation Technology
Miao Yu, Beijing Institute of Space Science and Technology Information
Jun Lu, China Satellite Navigation Office Engineering Center
Yintang Yang, Xidian University
Xiuwan Chen, Peking University
Mingquan Lu, Tsinghua University

Xiang'an Zhao, China Defense Science and Technology Information Center

Jing Zhao, Ministry of Science and Technology National Remote Sensing Center

Guangzhou Ouyang, Academy of Opto-Electronics, Chinese Academy of Science

Gang Hu, Beijing Unicore Communications Inc.

Yamin Dang, Chinese Academy of Surveying and Mapping, China

Zhong Dou, National Time Service Center of Chinese Academy of Sciences

Contents

Part I Precise Orbit Determination and Positioning

1 Parametric Study of Solar Radiation Pressure Model Applying to Navigation Satellite Orbit Determination for Long Arc.	3
Qiuli Chen, Haihong Wang, Hui Yang and Zhonggui Chen	
2 BDS/GNSS Real-Time Kinematic Precise Point Positioning with Un-differenced Ambiguity Resolution	13
Lizhong Qu, Qile Zhao, Jing Guo, Guangxing Wang, Xiangxin Guo, Qiang Zhang, Kecai Jiang and Liang Luo	
3 Characteristics Analysis of BeiDou Melbourne-Wübbena Combination.	31
Xiyang He and Xiaohong Zhang	
4 Refining of BDS Differential Code Bias Model	47
Qiankun Liu, Lifan Sui, Guorui Xiao, Yu Gan, Guobin Qi and Tian Zheng	
5 Estimation Strategy and Accuracy Analysis of GNSS Real-Time Precise Satellite Clock Error.	57
Liang Chen, Changjiang Geng, Quan Zhou and Wenhai Jiao	
6 Integrating BDS and GPS to Accelerate Convergence and Initialization Time of Precise Point Positioning	67
Zongpeng Pan, Hongzhou Chai, Zehui Liu, Kefan Yang, Yang Chong and Yangyin Xu	
7 Study on BeiDou Navigation Satellite Precise Orbit Determination Based on the Extended Kalman Filtering.	81
Yan Wang, Chuanding Zhang and Lijie Song	

8	Analysis of Multi-frequency BDS/GPS RTK Positioning	95
	Yijun Tian, Dongqing Zhao, Shuangna Zhang, Zhiyong Huang and Hang Dong	
9	Upper Atmospheric Density Retrieval from Accelerometer on Board GRACE Mission	105
	Runjing Chen and Bibo Peng	
10	Dynamic GPS Precise Point Positioning for Deformation Monitoring Using Prior Information	117
	Zhiping Liu, Ziqiang Zhao and Qiuzhao Zhang	
11	Positioning Accuracy Analysis of Beidou Continuous Operation Tracking Stations	129
	Xiangxin Guo, Qile Zhao, Shenghua Jiang and Min Li	
12	A Parallel Processing Strategy of Large GNSS Data Based on Precise Point Positioning Model	139
	Yang Cui, Zhiping Lu, Hao Lu, Jian Li, Yupu Wang and Lingyong Huang	
13	New Results of Multi-GNSS Orbits Validation Based on SLR Observations	149
	Jinchao Xia, Geshi Tang, Chao Han, Jianfeng Cao, Hongzheng Cui and Xie Li	
14	Study in BDS Uncombined PPP Ionospheric Delay Estimation and Differential Code Biases	161
	Huarun Wang, Hongzhou Chai, Min Wang, Zongpeng Pan and Yang Chong	
15	Research of Key Technology on the Combination of GPS, VLBI, SLR and DORIS Solution for Station Coordinates and ERPs	173
	Min Li, Tian-he Xu and Hang Yu	
16	Cycle-Slip Detection and Correction Based on Polynomial-Fitting Ionosphere-Free Combination and Ionospheric Total Electron Contents Rate.	191
	Ming Liu, Hongzhou Chai and Di Li	
17	BDS and GPS Ultra-Short Baseline Measurement Performance Comparison and Analysis	201
	Kang Zhang, Jinming Hao, Yu Zhang, Qiang Li and Tuansheng Yang	

18 The iGMAS Combined Products and the Analysis of Their Consistency 213
 Hongliang Cai, Kangkang Chen, Tianhe Xu and Guo Chen

19 Enhanced RTK Integer Ambiguity Resolution with BeiDou Triple-Frequency Observations 227
 Tao Li, Kongzhe Chen and Jinling Wang

20 GPS/BDS One-Step Combined Precise Orbit Determination Based on Double-Differenced Mode 239
 Yao Kong, Baoqi Sun, Xuhai Yang and Xiaozhen Zhang

21 Research on Ranging Algorithm Based on Combined-Process Method Using Frequency-Difference, Time-Difference and Relative-Velocity 249
 Rui-Qiang Yang, Peng Zhang, Xing-Wang Zhong, Deng-Feng Wang and Nian-Ke Zong

22 Multipath Effect on Phase Center Calibration of GNSS Antenna 259
 Lixun Li, Baiyu Li, Huaming Chen and Feixue Wang

23 An Improved TCAR Based on the Optimal Combination of Carrier Phase and Pseudo-range Observations. 269
 Xing Wang, Wenxiang Liu, Yangbo Huang and Guangfu Sun

24 A Hybrid Navigation Constellation Inter-satellite Link Assignment Algorithm for the Integrated Optimization of the Inter-satellite Observing and Communication Performance 283
 Bo Xu, Donghui Wang, Wenxiang Liu and Guangfu Sun

25 Distributed Orbit Determination Based on Increased Measurement Covariance EKF for Global Navigation Satellite System with Inter-satellite Link 297
 Kai Xue, Yuanlan Wen, Ying Liao, Yisheng Song, Tianxiang Su and Zhi Zhang

26 Strategy and Accuracy Analysis of Space-Borne GPS Single-Frequency Real-Time Orbit Determination 311
 Fuhong Wang, Lei Guo and Xuewen Gong

27 Orbit Determination Using Combined GPS + Beidou Observations for Low Earth Cubesats: Software Validation in Ground Testbed 321
 Yang Yang, Xiaokui Yue, Geshi Tang, Hongzheng Cui and Baiyan Song

28 Preliminary Analysis of Positioning Performance with BDS Virtual Reference Station Technology 335
 Dongfeng Yu, Pengbo Li, Guangxing Wang, Mingzhi Zhou and Zhigang Hu

Part II Atomic Clock Technique and Time-Frequency System

29 Development of Space Mini Passive Hydrogen Maser 343
 Yonghui Xie, Pengfei Chen, Shanmin Liu, Yuxian Pei, Shuai Tao and Chuanfu Lin

30 Development of High Frequency-Temperature Stability of OCXO for Aerospace Applications 351
 Lei Yang, Yuhao Qin, Jian Huang, Xiaoqiang Zhang, Wei Jiang, Bing Cheng, Zhifu Feng and Fanghong Guo

31 Realization and Influencing Factors Analysis of ACES (Atomic Clock Ensemble in Space) Management 359
 Jun Xie, Yunfeng Sun, Yongsheng Qu, Dong He and Ming Zhao

32 Research of Navigation Constellation Independent Punctuality Based on Different Configuration of Satellite Clocks. 369
 Yang Yang, Fan Jian jun, Hong Yuan, Jin Shu xin and Yang Yu fei

33 Fiber Based Radio Frequency Dissemination Scheme to Multiple Users 379
 Wei Chen, Dan Xu, Nan Cheng, Qin Liu, Fei Yang, Youzhen Z. Gui and Haiwen W. Cai

34 Influence of Lamp Spectral Profile on Short-Term Stability and Light Shift of a Rubidium Atomic Clock. 387
 Qiang Hao, Shengguo He, Feng Xu, Feng Zhao and Ganghua Mei

35 Evaluation and Monitoring on the Single Station Time Difference Based on the BDS, GPS and GLONASS Data 399
 Guang Sun, Hua Lu, Lirong Shen, Xiaolin Jia, Meijun Guo, Yijun Mo and Yingjie Hong

36 Prediction of Navigation Satellite Clock Bias by Gaussian Process Regression 411
 Yu Lei, Danning Zhao, Zhaopeng Hu and Hongbing Cai

Part III PNT System and New Technologies of Navigation

37 Distributed GNSS Collaborative Positioning Algorithms and Performance Analysis 427
 Bin Huang, Zheng Yao, Xiaowei Cui and Mingquan Lu

38 Indoor Location Fingerprinting System Using DTMB Signal 439
 Qiqi Wang, Shuai Li, Yingxin Zhao, Kun Chen, Bing Liu, Xiyan Xu and Hong Wu

39 Simulation Analysis and Research on Key Technology for a Full-Time-Running Stellar Refraction Autonomous Navigation 449
 Qin Lin, Zhi Li, Huaifeng Li, Liang Song and Bo Meng

40 Experimental Investigation on CNS/SINS Integrated Navigation Using Star Tracker 463
 Dangwei Wang and Lei Zhou

41 Simultaneous Ranging and Communication Based on X-Ray Communications 475
 Shibin Song, Luping Xu, Hua Zhang, Yuanjie Bai and Liyan Luo

42 Cross Correlation Mitigation Algorithm for Indoor Positioning Receiver 485
 Yang Hu, Zhongliang Deng, Jichao Jiao, Jie Xia, Yuezhou Hu and Zhuang Yuan

43 An Optimal Data Fusion Algorithm Based on the Triple Integration of PPP-GNSS, INS and Terrestrial Ranging System 493
 Wei Jiang, Yong Li and Chris Rizos

44 Pulsar Navigation Profile Folding and Measurement Based on Loop Tracking. 507
Xinyuan Zhang, Ping Shuai and Liangwei Huang

45 Application of an Improved K Nearest Neighbor Algorithm in WiFi Indoor Positioning 517
Yeqing Fang, Zhongliang Deng, Chen Xue, Jichao Jiao, Hui Zeng, Ruoyu Zheng and Shunbao Lu

46 A New Pulse Time-of-Arrival Estimation Method for X-Ray Pulsar Navigation 525
Qingqing Lin, Ping Shuai and Liangwei Huang

47 A Research of Code Tracking Loop for Navigation Signal Based on DS/FH Modulation. 535
Zhuxi Yu, Jiaolong Wei, ZuPing Tang, Zhihui Zhou and Yuan Xue

48 Method to Construct Database of X-ray Pulsar-Based Navigation Using Space-Based Observation Data 545
Xiao-long Hao, Qiang-wen Yang and Liang-wei Huang

49 Physical Analysis on Pulsar-Based Navigation System: Preliminary Designs of Timing Model and a New Prototype of X-Ray Detector. 557
Yuanjie Du

50 GNSS P2P Cooperative Positioning System for Multiple Search-and-Rescue Robots 571
Yulong Song and Baowang Lian

51 A Position Approach Based on the Special Sub-frame in TDD LTE-A System 583
Dongyan Wei, Zhili He, Xuping Gong, Ying Xu and Hong Yuan

52 Research on the Pulsar Optimizing Method and the Database Construction. 595
Xiao-Ming Bei, Ping Shuai, Liang-wei Hang and Xin-yuan Zhang

53 Study on the Transportation of Electrons in the Graphene-Based X-Ray Detector. 603
Yaojun Wu, Ping Shuai, Hengbin Zhang, Qian Zhang and Lingzhong Fu

54 Concepts and Perspectives on Navigation Satellite Autonomous Health Management System Based on Cognitive Technology 613
 Jun Xie, Jianjun Zhang and Ming Xue

55 An X-Ray Pulsar TOA Estimation Method Considering Spacecraft Orbit Motion 625
 Liangwei Huang, Ping Shuai and Xinyuan Zhang

56 The Research of X-Ray Pulsar Signals Simulation Method 635
 Lingzhong Fu, Ping Shuai, Mengfan Xue, Haifeng Sun and Haiyan Fang

57 Multi-frame Visual Odometry in Image-Aided Inertial Navigation System 649
 N.S. Gopaul, J.G. Wang and B. Hu

58 Integration of GNSS and MEMS-Based Rotary INS for Bridging GNSS Outages 659
 Shuang Du, Wei Sun and Yang Gao

59 An Integrated PDR/GNSS Pedestrian Navigation System 677
 Haiyu Lan, Chunyang Yu and Naser El-Sheimy

60 Moving Target Autonomous Positioning Based on Vision for UAV 691
 Long Zhao and Pengfeng Chen

61 An Orbit Determination Algorithm for Spacecrafts Navigated by a Single X-Ray Pulsar 703
 Rong Jiao, Luping Xu, Hua Zhang and Yan Ma

62 Research on the Architecture of Cloud GNSS Based on Hadoop 717
 Linyang Li, Zhiping Lu, Lihui Fan and Jian Li

63 Wi-Fi Fingerprint Positioning Updated by Pedestrian Dead Reckoning for Mobile Phone Indoor Localization 729
 Qiang Chang, Samuel Van de Velde, Weiping Wang, Qun Li, Hongtao Hou and Steendam Heidi

64 Sequence-Based Motion Recognition Assisted Pedestrian Dead Reckoning Using a Smartphone. 741
 Chengxuan Liu, Ling Pei, Jiuchao Qian, Lin Wang, Peilin Liu and Wenxian Yu

65	An Unconventional Full Tightly-Coupled Multi-Sensor Integration for Kinematic Positioning and Navigation	753
	Jian-Guo Wang, Kun Qian and Baoxin Hu	
66	A WIFI/INS Indoor Pedestrian Navigation System Augmented by Context Feature.	767
	Ling Yang, Yong Li and Chris Rizos	
67	A Pedestrian Movement Direction Recognition Method Based on Inertial Sensors	781
	Shunbao Lu, Zhongliang Deng, Chen Xue, Yeqing Fang, Ruoyu Zheng and Hui Zeng	
68	High-Precision Simulator for Strapdown Inertial Navigation Systems Based on Real Dynamics from GNSS and IMU Integration	789
	Gongmin Yan, Jinling Wang and Xinyi Zhou	
69	A Research on All Source Navigation and Positioning and Its Critical Technology.	801
	YongBin Zhou, Jun Lai, XiYe Guo and Jun Yang	
70	Mixed Interacting Filter for Tracking with Multiple System and Model	809
	Xiaoguang Zhang, Dongyan Wei, Ying Xu and Hong Yuan	
71	Research on Multi-Source Fusion Based Seamless Indoor/Outdoor Positioning Technology.	819
	Ying Xu, Hong Yuan, Dongyan Wei, Qifeng Lai, Xiaoguang Zhang and Weina Hao	
72	Study on Quantum Inter-satellite Link Technology	839
	Xian'an Zheng, Ying Wang, Yansong Meng and Zhe Su	
 Part IV Policies and Regulations, Standard and Intellectual Property		
73	The Urgency and Principals of the Industry Policy Optimization for Beidou Satellite Navigation System.	853
	Junlin Yang and Xiangming Hu	

Part I
Precise Orbit Determination
and Positioning

Chapter 1

Parametric Study of Solar Radiation Pressure Model Applying to Navigation Satellite Orbit Determination for Long Arc

Qiuli Chen, Haihong Wang, Hui Yang and Zhonggui Chen

Abstract The precision of broadcast ephemeris of navigation satellite is one of the important factors, which determines the performance indicator of navigation satellite system. However, the result of precise orbit determination and orbit prediction affected the precision of broadcast ephemeris. It's a very complex task to calculate the precise orbit of satellite. The in-orbit satellite is effected by several kinds of perturbations, that are Earth's non-spherical gravitation, Sun and Moon gravitation, solar radiation pressure (SRP), tidal acceleration and other little perturbing acceleration. Compared with the other pressure, SRP is the most important error source for medium-earth-orbit satellites. On the basis of engineering parameter and attitude control modes of a certain navigation satellite, this paper had constructed the composite SRP model including direct SRP, thermal radiation pressure (TRP), Earth radiation pressure (ERP). The influence of SRP model parameters had been analyzed in precise orbit determination by long arc orbit calculation. According to the analyzed result, the adjust strategy of the parameters of SRP model had been given for long arc orbit determination, thus the result of orbit determination could be improved further.

Keywords Navigation satellite · Orbit determination for long arc · Solar radiation pressure · Parameter analysis

1.1 Introduction

The precision of broadcast ephemeris of navigation satellite is one of the important factors, which determines the performance indicator of navigation satellite system. However, the result of precise orbit determination and orbit prediction affected the precision of broadcast ephemeris. It's a very complex task to calculate the precise

Q. Chen (✉) · H. Wang · H. Yang · Z. Chen
Beijing Institute of Spacecraft System Engineering, Beijing 100094, China
e-mail: cql_hong@163.com

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_1

orbit of satellite. The in-orbit satellite is effected by several kinds of perturbations, that are Earth's non-spherical gravitation, Sun and Moon gravitation, solar radiation pressure (SRP), tidal acceleration and other little perturbing acceleration [1]. Compared with the other pressure, SRP is the most important error source for medium-earth-orbit satellites. On the basis of engineering parameter and attitude control modes of a certain navigation satellite, this paper had constructed the composite SRP model including direct SRP, thermal radiation pressure (TRP), Earth radiation pressure (ERP). The influence of SRP model parameters had been analyzed in precise orbit determination by long arc orbit calculation. According to the analyzed result, the adjust strategy of the parameters in SRP model had been given for long arc orbit determination.

1.2 Dynamic Model for Orbit Determination

In general, orbit determination includes initial orbit determination and orbit improvement. The desired time Orbit parameters are calculated according to measuring data of given arc orbit, that is orbit determination. Conventionally, initial orbit determination is that six orbital elements be calculated by six unattached measuring parameters, which is a two-body problem without any perturbation elements. On the basis of initial orbit, orbital elements of given epoch could be refined by orbit improvement using much measuring data [2].

Dynamic model for satellite initial orbit determination as follow [3],

$$\begin{cases} x = x(\sigma_i, t) \\ y = y(\sigma_i, t) \\ z = z(\sigma_i, t) \\ x' = x'(\sigma_i, t) \\ y' = y'(\sigma_i, t) \\ z' = z'(\sigma_i, t) \end{cases} \quad (1.1)$$

where, x, y, z is satellite position, x', y', z' being velocity of satellite, and $\sigma_i (i = 1, \dots, 6)$ being six unattached integral constants. On the basis of x, y, z and x', y', z' of given time, instantaneous position and velocity of satellite could be solved, which is the process of initial orbit determination.

In precise orbit determination, several perturbed forces must be calculated. And the precision of perturbed force model would affect the precision of satellite dynamic model. In theory, the precision of orbit would be better with much precise perturbed force model. The magnitude of perturbing acceleration of navigation satellite with geosynchronous orbit altitude had been shown in Table 1.1. And whether the perturbed force would affected sub-meter orbit had also be provided in this table [4].

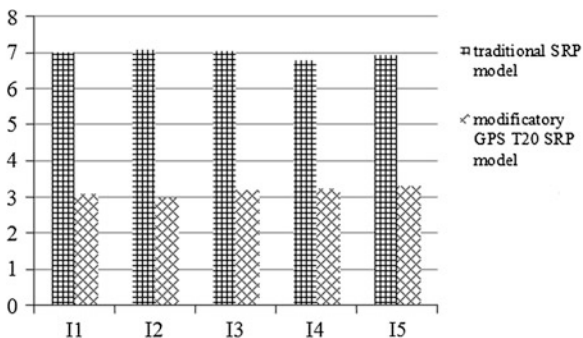
Table 1.1 Magnitude of perturbing acceleration of navigation satellite

Perturbing force	Perturbing acceleration (m/s ²)	Effect precision (10 ⁻¹ m)
Earth's non-spherical gravitation	3.7×10^{-5}	✓
Sun gravitation	3.7×10^{-6}	✓
Moon gravitation	3.7×10^{-6}	✓
Planet gravitation	3.5×10^{-12}	×
Direct SRP	2×10^{-6} (S/m = 0.1 m ² /Kg)	✓
Tide	8.7×10^{-12}	×
Oceanic	1.4×10^{-12}	×
ERP	1.5×10^{-10}	✓
Relativity	5×10^{-10}	✓

According to the magnitude of the perturbing elements in Table 1.1, selection principle of different perturbed force models had been designed in this paper's orbit determination. RKF7(8) integral method had been used in orbit integral calculation. At first, several perturbing elements used the existed model which hasn't relationship with satellite engineering parameters. The above perturbed forces include 70 × 70 Earth's gravitation, Sun and Moon gravitation, and relativity. To SRP model, traditional fixed area-mass radio model and modificatory GPS T20 model had been chosen respectively in orbit determination. Precise orbit had been determined for 3 days arc. The Root Mean Square (RMS) of three-dimensional position error of 5 navigation satellites had been compared in Fig. 1.1.

The result of Fig. 1.1 showed that, the result of precise orbit determination using traditional SRP model and GPS SRP model was unideal. Because the SRP model is related with satellite surface configuration dimension, surface material characteristic, and attitude control law. Thus, on the basis of factual navigation satellite engineering parameters constructing SRP model is necessary.

Fig. 1.1 RMS of 5 navigation satellites for precise orbit determination (m)



1.3 Solar Radiation Pressure Modeled

Surface elements of the navigation satellite include regular and un-regular polygon, column, taper, paraboloid, and so on. Complex three-dimensional surface configuration drawing of the satellite was shown in Fig. 1.2, which had been studied in this paper.

The mission requirements of the Navigation satellite determine its attitude control mode, that is navigation antenna which coincides with the +Z axis of body fixed coordinate system pointing geocentric center of the Earth, to transmitted the navigation signals. In order to get enough solar panels energy, ensure the entire star power, avoid that single degree rotating of solar panels can't be guaranteed the precision of tracking the Sun, The addition of a yaw control about Z axis could make the solar panels to track and perpendicular to the solar panels by rotating round Y axis of body fixed coordinate system. When the angle between the sun vector and the orbital plane is smaller than a certain value B_0 , satellite attitude control system no longer work, that is yaw-fixed regime.

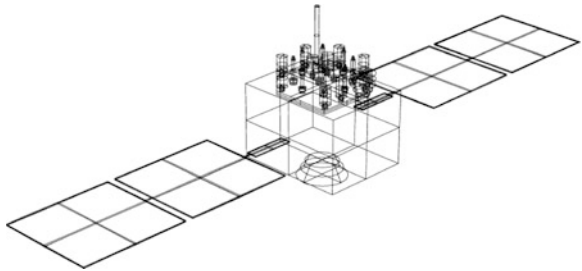
Direct solar is the most important source of SRP. Solar could be equivalent to parallel light due to that the distance between Sun and satellite is so far, and the area of Sun is much bigger than that of satellite. A hank of Sun light hitting satellite surface could be divide into three parts, that are absorbed, specular reflected, diffuse reflected. On the basis of Einstein's special theory of relativity and synthesis calculation, solar radiation pressure in the normal and tangential on the surface element caused by a beam of light are respectively F_n and F_s :

$$F_n = -(AE \cos \theta / c) \left[(1 + v\mu) \cos \theta + \frac{2}{3} v(1 - \mu) \right] \quad (1.2)$$

$$F_s = (AE \cos \theta / c) (1 - v\mu) \sin \theta \quad (1.3)$$

where, A be irradiated surface area, $E = 1367 \text{ W/m}^2$ be solar radiation intensity, θ be incidence angle which is the angle between the incident light and the surface element normal, lightspeed $c = 3 \times 10^8 \text{ m/s}^2$, v be reflectivity of the surface material, μ be specular coefficient of surface of the material. To the in-orbit satellite, the radiated relationship between surface parts and Sun were determined by attitude

Fig. 1.2 Complex three-dimensional surface configuration of the satellite



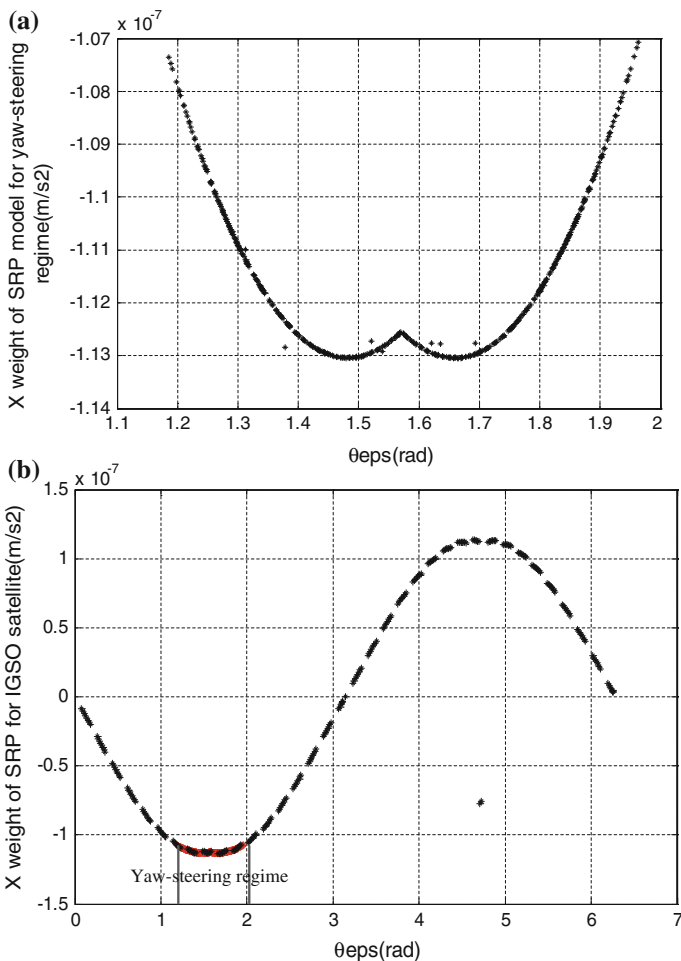


Fig. 1.3 a X weight of SRP acceleration for yaw-steering regime. b X weight of SRP acceleration for yaw-steering and yaw-fixed regime

control law. During yaw-steering regime, Sun vector moved always in XOZ surface of body-fixed coordinate system. For yaw-fixed regime, there was an angle between Sun vector and XOZ surface of body-fixed coordinate system, named β . Theoretically, direct SRP model has a Y weight during yaw-fixed regime.

TRP is caused by temperature grads between satellite surface parts and space environment. Stefan-Boltzmann law [5] is the theoretical basis of TRP simulate modeled. The input of TRP model were complex three-dimensional surface configuration and surface temperature of in-orbit navigation satellite. Similarly to direct solar light, the satellite surface also radiated by Earth reflected light. Earth reflected light were divided into optical radiation and infrared radiation. Optical radiation

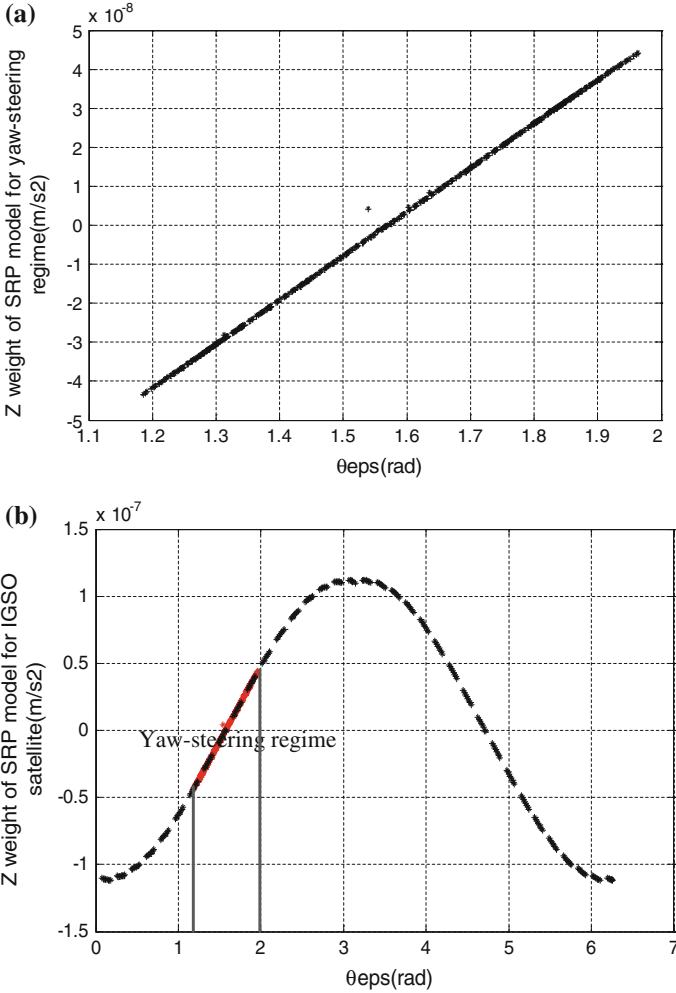


Fig. 1.4 **a** Z weight of SRP acceleration for yaw-steering regime, **b** Z weight of SRP acceleration for yaw-steering and yaw-fixed regime

was exited on the hemisphere of Earth radiated by Sun, which was determined by the relative geometry relationship of Sun, Earth, and the satellite. The intensity of infrared radiation was related with orbital altitude and clouds change.

On the basis of complex three-dimensional surface configuration in Fig. 1.2, other engineering parameters, and two attitude control modes, direct SRP, TRP, and ERP were modeled in body-fixed coordinate system. X, Z, and Y solar radiation perturbation acceleration is shown in Figs. 1.3, 1.4 and 1.5 respectively.

In Figs. 1.3 and 1.4, the angle between Sun vector and +Z axis of body-fixed coordinate system was chosen as independent variable for X weight and Z weight

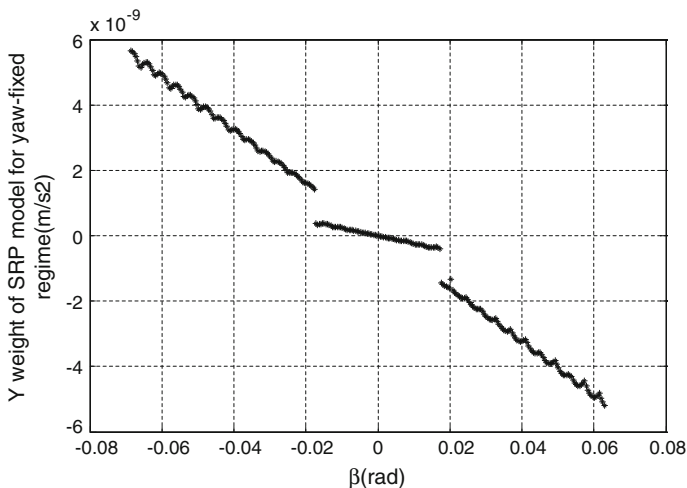


Fig. 1.5 Y weight of SRP acceleration for yaw-fixed regime

of SRP model, named θ_{eps} . Theoretically, Y weight of direct SRP model was exited only during yaw-fixed regime. The simulate result show that, Y weight of SRP model had perfect pertinency with β . Thus, independent variable of Y weight model was β .

1.4 Simulation Result

In order to validate the accuracy of simulate SRP mode, precise orbit determination had been calculated using precise ephemeris of October 2013 published by Wuhan University. Besides SRP, orbit determination algorithm and perturbation elements were same to the Fig. 1.1. The arc orbit was 3 days. The sampling alternation was 60 s. Precise orbit determination had been calculated for five navigation satellites. Three results of precise orbit determination had been stated in Fig. 1.6, where the Root Mean Square (RMS) of three-dimensional position error of 5 navigation satellites had been compared again.

RMS of three-dimensional position error of precise orbit determination using three different SRP models had been compared in Fig. 1.6. The comparative result shown that, the precision of orbit determination using simulation SRP model was better than traditional SRP model and modificatory GPS T20 SRP model. Compared with traditional SRP model, ameliorative scope of simulation SRP model was 64 %, which is 25 % for modificatory GPS T20 SRP model.

Whereas, the precision of satellite engineering parameters used in Sect. 1.3 for simulation SRP model couldn't be very perfect. Besides, there would be error between actual in-orbit estate and ideal estate for SRP model. Least square method

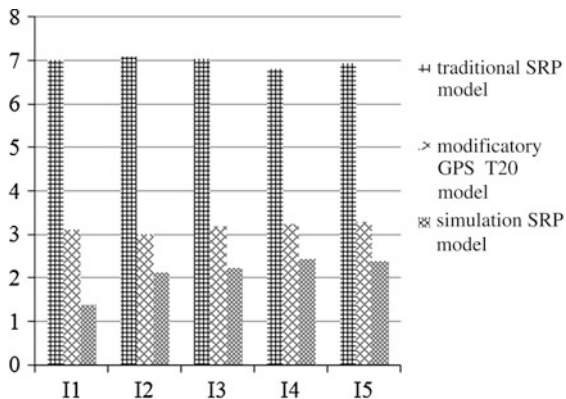


Fig. 1.6 RMS of 5 navigation satellites for precise orbit determination (m)

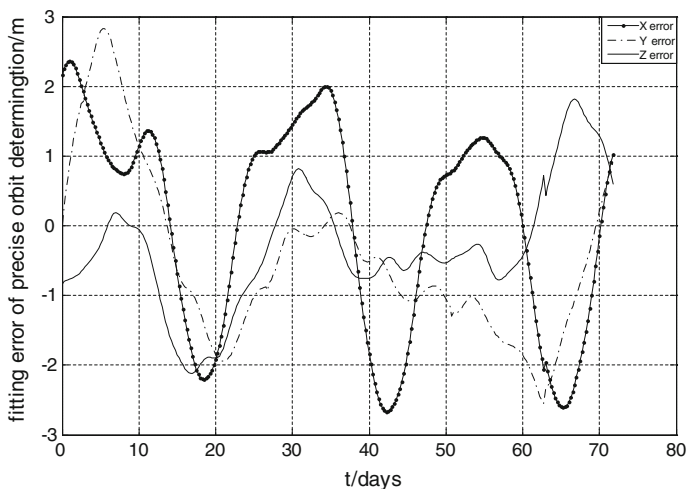


Fig. 1.7 RMS of 3 days arc precise orbit determination for the satellites

had been used to adjust the parameters of simulation SRP model. And optimize SRP model had been used in precise orbit determination for a navigation satellite.

The change of RMS of 3 days arc had been shown in Fig. 1.7. The parameters of simulation SRP model were adjusted by least square method. And the RMS of precise orbit determination for the satellite was 1.09 m, that were 0.25 m better than simulation SRP model. RMS of precise orbit determination for the other satellites had also been improved using the same method.

1.5 Conclusion

Simulation calculation had proved that SRP model are related with satellite engineering parameters in this paper. Thus, it's necessary to construct SRP model according to satellite actual parameters for high precise navigation satellite. Compared with traditional SRP model with fixed area-mass ratio and modificatory GPS T20 model, RMS of three-dimensional position error of precise orbit determination using simulation SRP model had been advanced 64, 25 % respectively. On the basis of simulation SRP model and published observation data, SRP model parameters could be optimized ulteriorly, which could improve the RMS of precise orbit determination farther.

References

1. Qiuli C, Zhonggui C, Wang H (2013) Method of modeling solar radiation pressure based on attitude control law of navigation satellites. In: CSNC2013, Wuhan
2. Liu L, Wang H, Hu S (2005) Summary on satellite orbit determination. *J Spacecraft TT&C Technol* 24(2)
3. Yang J, Fan J (2009) *Orbit dynamic and control of spacecraft*. China Aerospace Press, Beijing
4. Wen Y *Analysis and simulation technology of navigation satellite system*
5. Matpack Version 1.9.9. <http://WWW.matpack.de/>
6. Tom K, Moriba J (2007) Analysis of orbit prediction sensitivity to thermal emissions acceleration modeling for high area-to-mass ratio (HAMR) objects. Boeing LTs Inc: 6200 Uptown Blvd. Suite 240 Albuquerque

Chapter 2

BDS/GNSS Real-Time Kinematic Precise Point Positioning with Un-differenced Ambiguity Resolution

Lizhong Qu, Qile Zhao, Jing Guo, Guangxing Wang, Xiangxin Guo, Qiang Zhang, Kecai Jiang and Liang Luo

Abstract BeiDou Navigation Satellite System (BDS) is constructed and operated by China independently. It has important significance in researching the real-time kinematic precise point positioning (PPP-RTK) based on BDS. The fusion of multi-GNSS in data processing can increase satellites observed, improve the geometric configuration of satellites constellation, enhance the accuracy, continuity and reliability in real-time kinematic positioning. This paper deduced the mathematical model of BDS/GNSS PPP-RTK, studied the theories and methods of BDS/GNSS un-difference ambiguity resolution, and realized multi-GNSS PPP-RTK by fixing the un-differenced ambiguities of BDS-only, GPS-only, both BDS and GPS satellites, respectively. Multi-GNSS data of reference stations from Crustal Monitor of Network of China (CMONOC) were processed. The experiment results showed that the values of fractional cycle bias (FCB) of BDS IGSO and MEO satellites kept stable; The precision were less than 1 cm in plane direction and 3 cm in vertical direction after the fusion of BDS, GPS and GLONASS even without ambiguity resolution, respectively; It not only raised the precision of positioning but also quicken the convergence speed after fixing the un-differenced ambiguities of BDS IGSO and MEO satellites; The performance of multi-GNSS PPP-RTK improved further after fixing the un-differenced ambiguities of both BDS and GPS satellites.

Keywords BDS · GNSS · Real-time kinematic precise point positioning · Un-differenced ambiguity resolution · Fractional cycle bias

L. Qu (✉) · Q. Zhao · J. Guo · G. Wang · X. Guo · Q. Zhang · K. Jiang · L. Luo
GNSS Research Center, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
e-mail: qulizhong@whu.edu.cn

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_2

2.1 Introduction

Beidou Navigation Satellite System (BDS) is constructed and operated by China independently, which is the third mature Global Navigation Satellite System (GNSS) by following the US GPS and the Russian GLONASS. From December 27, 2012 on, BDS began to provide regional service officially. Nowadays, BDS satellite constellation is composed of 5 GEO, 5 MEO and 4 IGSO satellites [1]. Domestic and foreign scholars have carried out a large number of studies on the precise orbit determination and precise positioning of BDS [2–6]. The accuracy of BDS satellites' orbit radial direction reached 10 cm by using Wuhan University PANDA software package [5, 6]. The fusion of multi-GNSS data effectively increase the number of satellites observed, improve satellite spatial geometry construction. Thus, the integration of multi-GNSS data can improve the positioning accuracy, continuity and reliability [7–10].

Traditional precise point positioning using un-differenced and ionosphere-free pseudo-range and phase combination measurements can obtain positioning results of centimetre level. But the fractional bias of the hardware delay result in the loss of integrity character of un-differenced integer ambiguity and the float solutions are acquired [11]. Currently, the un-differenced ambiguity resolution researches focused on the GPS system. There were three main methods: decoupled clock model, single-difference between satellites model and integer phase clock model [12–16]. By ignoring the effect of receiver and satellite terminal pseudo-range hardware delay on phase ambiguity, Geng et al. proved the equivalence of single-difference between-satellites model and integer phase clock model [17]. Without any assumption, Shi et al. proved the three methods equivalent by deducing the rigorous formulas and comparing computational efficiency of the three methods [18]. BDS adopt Code Division Multiple Access (CDMA) signal, which is the same as GPS. So the GPS ambiguity resolution methods were able to be used on BDS. Based on the single-difference between satellites ambiguity resolution method, we realized GNSS real-time PPP by fixing the BDS-only, GPS-only and combined BDS/GPS ambiguity. As we all known, the signal establishment system of GLONASS is Frequency Division Multiple Access (FDMA). The hardware delay in receiver terminal between satellites cannot be eliminated directly by using single-difference between satellites. As a result, the ambiguity resolution of GLONASS satellites is more complicated [19]. In this manuscript, we just consider the ambiguity resolution of BDS and GPS satellites.

2.2 Data Collection

24 stations with Multi-GNSS data from Crustal Monitor of Network of China (CMONOC) were chosen as reference stations in this paper. The stations distribution is shown in Fig. 2.1. All stations just contain GPS, BDS and GLONASS observations. Figure 2.2 shows the sky plots of GPS, BDS and GLONASS satellites tracked at WUHN station. As we can see that, the distribution of GPS satellites tracked is

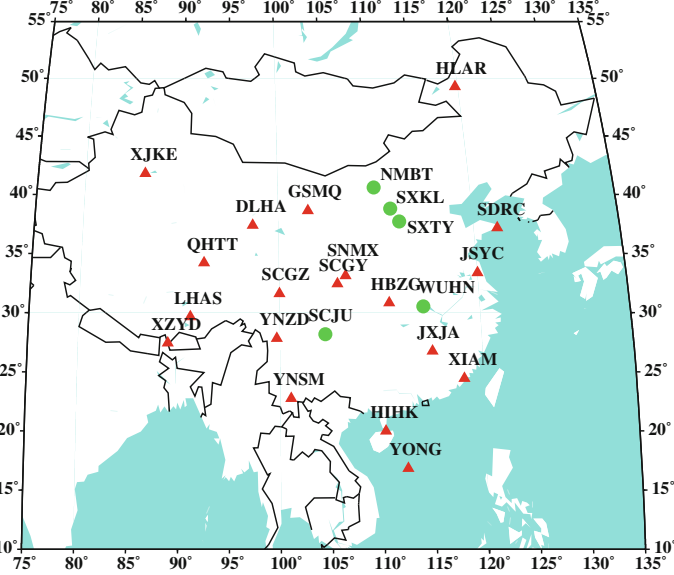


Fig. 2.1 CMONOC GNSS stations distribution for FCB estimation and ambiguity resolution

more balanced than that of GLONASS and BDS and the satellites distribution after the integration of multi-GNSS is more intensive. Figure 2.3 shows the time series of DOP values and number of satellites tracked at WUHN station. The integration of multi-GNSS improved the geometric configuration of satellite constellation, increased the number of observable satellites, which can improve the accuracy, enhance the continuity and the reliability in navigation and positioning service.

2.3 Processing Strategy

2.3.1 Mathematical Model

Precise point positioning mathematical model based on original dual-frequency ionosphere-free combination is:

$$P_r^s = \rho_0^s + c(dt_r - dt^s) + m_r^s T_r + c(B_r - B^s) + e_s \quad (2.1)$$

$$L^s = \rho_0^s + c(dt_r - dt^s) - \lambda n^s + m_r^s T_r + c(b_r - b^s) + \varepsilon_s \quad (2.2)$$

where ρ_0^s is the geometric distance between station r and satellite s . dt_r is the receiver clock error. T_r is the residual error of wet troposphere delay. m_r^s is the mapping function of wet troposphere delay. n^s is integer ambiguity. B_r , b_r are receiver terminal hardware delay of pseudo-range and phase, respectively. B^s , b^s are satellite terminal

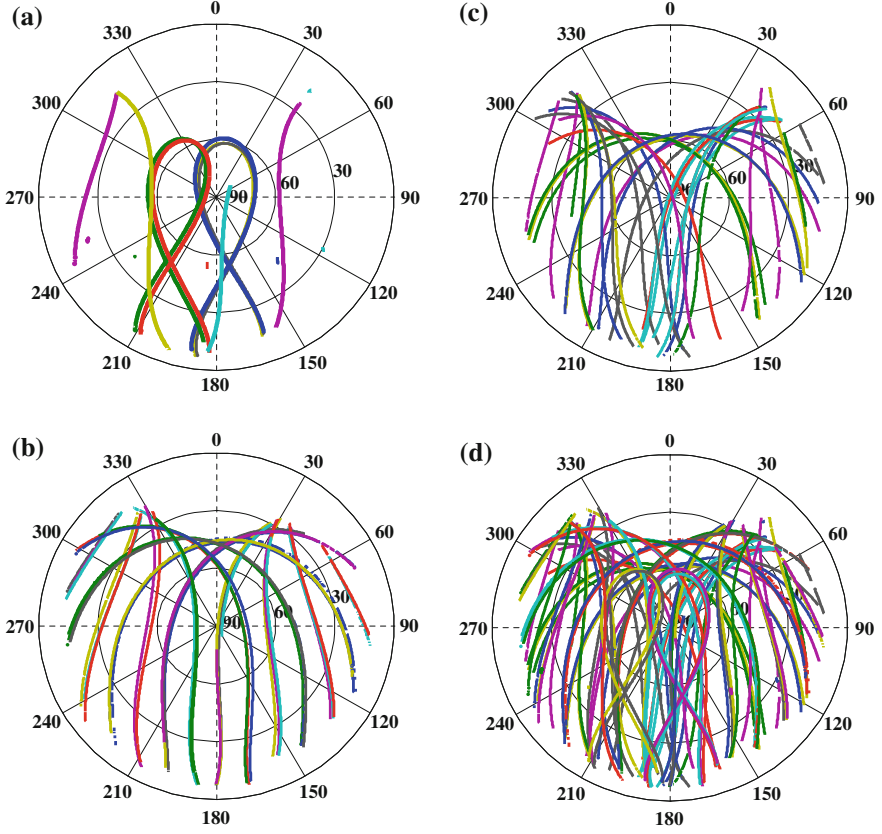


Fig. 2.2 Sky Plot of BDS, GPS and GLONASS at WUHN. **a** BDS. **b** GLONASS. **c** GPS. **d** BDS/GPS/GLONASS

hardware delay of pseudo-range and phase, respectively. e_s , ε_s are measurement noise of pseudo-range and phase, respectively, which includes the multipath errors. During parameter estimation, pseudo-range hardware delay B_r is absorbed by the receiver clock error. Formulas 2.1 and 2.2 are usually written as:

$$P^s = \rho_0^s + c \cdot \delta t_r + m_r^s T_r + e_s \quad (2.3)$$

$$L^s = \rho_0^s + c \cdot \delta t_r - \lambda n^s + m_r^s T_r + \varepsilon_s \quad (2.4)$$

where,

$$\delta t_r = dt_r + B_r \quad (2.5)$$

$$N^s = n^s + c(b_r - b^s - B_r + B^s)/\lambda \quad (2.6)$$

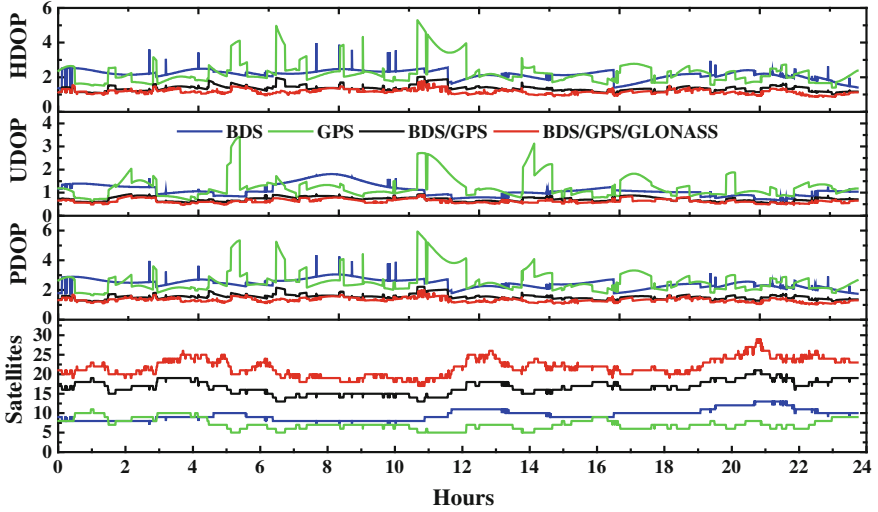


Fig. 2.3 Time series of DOP and satellites tracked at WUHN

Multi-GNSS PPP mathematical model is written as:

$$\begin{cases} P^G = \rho_0^G + c \cdot \delta t_r^G + m_r^G T_r + e_G \\ L^G = \rho_0^G + c \cdot \delta t_r^G - \lambda^G N^G + m_r^G T_r + \varepsilon_G \\ P^B = \rho_0^B + c \cdot \delta t_r^G + c \delta t_r^{B-G} + m_r^B T_r + e_B \\ L^B = \rho_0^B + c \cdot \delta t_r^G + c \delta t_r^{B-G} - \lambda^B N^B + m_r^B T_r + \varepsilon_B \\ P^R = \rho_0^R + c \cdot \delta t_r^G + c \delta t_r^{R-G} + m_r^R T_r + e_R \\ L^R = \rho_0^R + c \cdot \delta t_r^G + c_r^{R-G} - \lambda^R N^R + m_r^R T_r + \varepsilon_R \end{cases} \quad (2.7)$$

where,

$$\delta t_r^{B-G} = \delta t_r^B - \delta t_r^G \quad (2.8)$$

$$\delta t_r^{R-G} = \delta t_r^R - \delta t_r^G \quad (2.9)$$

Compared with GPS and BDS, the signal structure of GLONASS is FDMA, which results in internal frequency bias (IFB) between satellites. In this paper, the IFBs were not estimated that were absorbed into the pseudo-range residual and phase ambiguity.

2.3.2 Ambiguity Resolution Process

Real-time un-differenced ambiguity resolution contains two steps: (a) estimate the satellite FCB of wide-lane and narrow-lane; (b) ambiguity resolution at rovers.

2.3.2.1 FCB Estimation

As Formula 2.6 shown, the initial phase and hardware delay are strongly correlated with the ambiguity. After merging the hardware delay of the phase and code in the receiver terminal with those in the satellite terminal, Formula 2.6 can be written as:

$$N^s = \tilde{N}^s + \beta_r - \beta^s \quad (2.10)$$

The integer part of the hardware delay in the receiver and satellite terminal are difficult to separate from the ambiguity and the fractional cycle bias (FCB) of them broke the integer character of ambiguity. We can use the classic LAMBDA method to fix the ambiguity as long as FCBs were separated accurately from the integer part. The process to calculate the wide-lane and narrow-lane FCB is shown as follows:

- (1) Firstly, we solve the real wide-lane ambiguity based the Melbourne-Wübbena (MW) combination observation and then separate the integer part N_{mw}^s of the ambiguity from the FCB $(\beta_r - \beta^s)_{mw}$ by rounding directly;
- (2) Secondly, we construct the single-difference between satellites by using wide-lane FCB $(\hat{B}_r - \hat{B}^j)_{mw}$ to eliminate the FCB in receiver terminal. Then, all the same satellite pair FCBs of all the stations are acquired to calculate the average values. Assuming that n satellites were tracked by all the stations in the same time, we could get $n(n-1)/2$ average FCB values. After that, we use the least-square adjust with the quasi-stable datum or gravity datum method to get the wide-lane and narrow-lane FCB of every satellite;
- (3) Thirdly, the GPS daily static station coordinates by using PANDA software package are regard as the “ground truth” and the precise orbit and satellite clock error are used to calculate the un-differenced real ambiguity. The fixed wide-lane ambiguities are subtracted from the real ambiguity to get the real narrow-lane ambiguities. Then, we can get the satellite narrow-lane FCB by following Steps (1) and (2).

2.3.2.2 Ambiguity Resolution at Rovers

The satellite wide-lane and narrow-lane FCBs of the reference network are broadcasted to rovers in real-time. Firstly, we separate the wide-lane FCBs from the wide-lane observations to fix the wide-lane ambiguities in a short initial time. Meanwhile, we acquire the real solutions of un-differenced ambiguities and the real solutions of the narrow-lane ambiguities can be solved out by using Formula 2.11. Then, the satellites narrow-lane FCBs are subtracted from the real solutions of narrow-lane ambiguities. After separating the receiver narrow-lane FCBs, the classic LAMBDA method is used to search and fix the narrow-lane ambiguities.

Finally, the fixed ambiguities are used to update the positions and we can get the integer solutions [20].

$$N_{nl}^s = \frac{f_1 + f_2}{f_1} N^s - \frac{f_2}{f_1 - f_2} N_{wl}^s \quad (2.11)$$

where, N_{nl}^s is the real un-differenced ambiguity, N_{wl}^s is the integer un-differenced wide-lane ambiguity.

2.3.3 Data Processing Strategy

Precise orbit and clock corrections are BDS/GPS/GLS/GALIEO multi-GNSS products generated by Wuhan University PANDA software package [6]. The sampling interval of precise satellite clock corrections is 30 s. The prior standard deviation of PC and LC observations of BDS, GPS and GLONASS satellites are 0.2 and 0.002 m, respectively. During the parameter estimation, GPS and GLONASS satellites antenna PCO and PCV are corrected and BDS satellites antenna PCO are corrected for BDS satellites PCV are unknown. The PCO and PCV of receiver antenna are unknown. Meanwhile, the troposphere delay, the phase winding and the tidal effect are considered. The cutoff elevation angle is 7° . The specific observation model is shown in Table 2.1.

2.4 Results Analysis

Data on September 27, 2014, were processed in this paper (day of year is 260) and the sample interval is 30 s. 19 stations were taken as reference stations, which were the red triangle in Fig. 2.1. The wide-lane and narrow-lane FCB of BDS MEO and IGSO satellites and GPS satellites were estimated and provided to the rovers. When constitute single difference observations between satellites, we selected one satellite from the same system as the reference star. The FCB initialization time was 1200 s and the FCBs were upgraded every 30 s. In addition, 5 stations were taken as rovers to compare the accuracy of PPP-RTK float and fixed solutions, which were the green cycles in Fig. 2.1. We processed the static data to simulate the kinematic condition in BDS-only, GPS-only, combined BDS/GPS and combined BDS/GPS/GLONASS mode. But only the ambiguities of BDS and GPS were fixed. The GPS daily static solutions by using PANDA software package were taken as the “ground truth” for the 5 rovers. By comparing the float and fixed solutions with the “ground truth”, we analyzed the convergence time and the accuracy after convergence in the East, North and Up components for each mode.

Table 2.1 BDS/GNSS PPP-RTK measurement model and parameters estimation strategy

Observation model	
Observables	Un-differenced Ionosphere-free combination (PC and LC)
Elevation mask	7°
Interval	30 s
Precise orbit	PANDA software precise orbit: BDS/GPS/GLONASS/GALIEO [6]
Precise satellite clock error	PANDA software 30 s precise satellite clock correction: BDS/GPS/GLONASS/GALIEO
Satellite antenna PCO	GPS, GLS: IGS08.atx, BDS Default
Satellite antenna PCV	GPS, GLS: IGS08.atx, BDS unknown
Phase wrapping	Considered [21]
Receiver antenna PCO and PCV	Unknown
Tropospheric	Hydrostatics and wet-component delay: Saatamoien model, mapping function: VMF1
Ionosphere	Eliminate first order ionosphere by using PC and LC observations
Solid tide, ocean tide, pole motion	IERS conventions 2010
Parameter estimation	
Estimator	Square root information filter [22]
Base station coordinate (for estimating FCB)	GPS Daily Static Solution by using PANDA software, 3D Accuracy 1 cm
Wide-lane and narrow-lane FCB upgrade interval	30 s
FCB initial time	1200 s for both GPS and BDS satellites
Rovers coordinate (PPP)	Random-walking, 5 m, 10 cm $\sqrt{\Delta t}$
Receiver clock error	White noise
Tropospheric	Piece-wise-constant, 2 h interval
System time in receiver terminal	White noise
Integer ambiguity	Estimated as constant. Fix all GPS satellites, fix all BDS IGSO and MEO satellites

2.4.1 FCB Stability Analysis

Figure 2.4 shows the time series of the un-differenced wide-lane and narrow-lane FCBs of GPS satellites. Figure 2.5 shows the time series of the un-differenced wide-lane and narrow-lane FCBs of BDS IGSO and MEO satellites. In the two figures, the upper panel is the wide-lane FCB time series and the lower panel is the narrow-lane FCB time series. As we can see, the un-differenced wide-lane and narrow-lane FCB of GPS satellites are relatively stable in time domain and only a few satellites fluctuate in the initial stage. The changes of all the satellites keep within 0.2 cycles. The un-differenced wide-lane and narrow-lane FCB of BDS MEO satellites are also

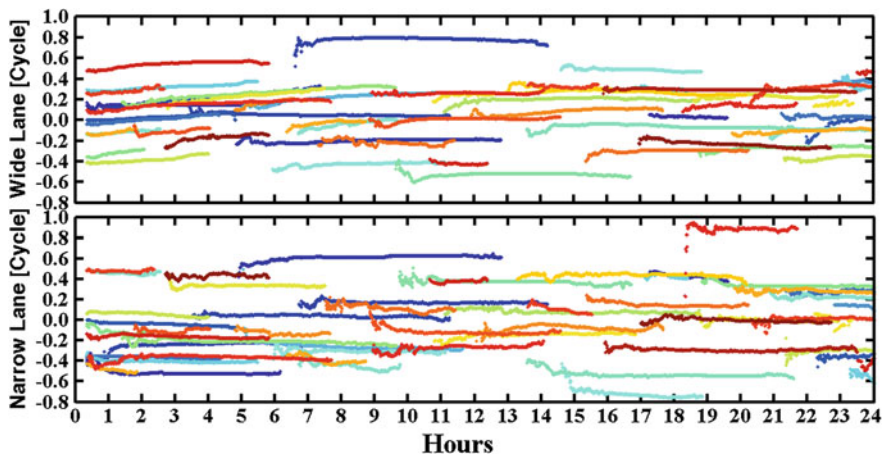


Fig. 2.4 Time series of GPS satellites wide-lane and narrow-lane FCB of CMONOC

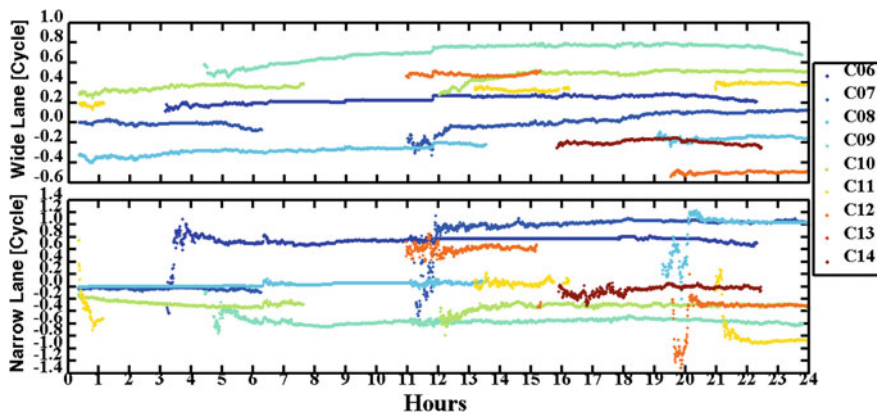


Fig. 2.5 Time Series of BDS IGSO and MEO satellites wide-lane and narrow-lane FCB of CMONOC

relatively stable in time domain and the changes keep within 0.2 cycles, while the wide-lane FCB of BDS IGSO satellites seem less stable than those of GPS satellites and BDS MEO satellites, the changes of the IGSO satellites reach 0.2–0.4 cycles. The narrow-lane FCB of BDS IGSO and MEO satellites are relatively stable after convergence while fluctuate larger than those of GPS satellites in the initial stage. As Formula 2.11 shown, the convergence time of the narrow-lane ambiguities depend on the wide-lane and un-differenced ambiguities. The length of FCB initial time seems enough for GPS and BDS wide-lane FCB estimation. But as Figs. 2.7 and 2.8 shows, BDS PPP needs more convergence time than that of GPS. During the convergence stage, the narrow-lane FCB change with the un-differenced ambiguity.

2.4.2 Positioning Results

Figure 2.6 shows the differences of BDS-only, GPS-only, BDS/GPS and BDS/GPS/GLONASS PPP float and fixed solutions against the “ground truth” in East, North and Up components. Table 2.2 shows the average values of RMS of the differences of BDS-only, GPS-only, BDS/GPS and BDS/GPS/GLONASS PPP of the rovers. We can find that:

- (1) The RMS of the differences of BDS PPP float solutions for all the rovers are better than 5, 2 and 7 cm in East, North and Up components, respectively. The average values of the RMS are 2.8, 1.9 and 5.9 cm, respectively. The RMS of GPS PPP float solutions are better than 3, 2 and 4 cm, respectively. The average values are 2.2, 1.2 and 3.2 cm, respectively. The East and North components of BDS PPP float solutions are very close to those of GPS PPP while the Up component is far from that of GPS, which mainly dues to the worse orbit accuracy of BDS. Compared with BDS-only and GPS-only PPP, the accuracy of combined BDS/GPS positioning has improved significantly. The RMSs are better than 2, 1 and 3 cm, respectively. And the average values are 1.3, 0.7 and 2.6 cm, respectively. The accuracy improved 40.9, 41.7,

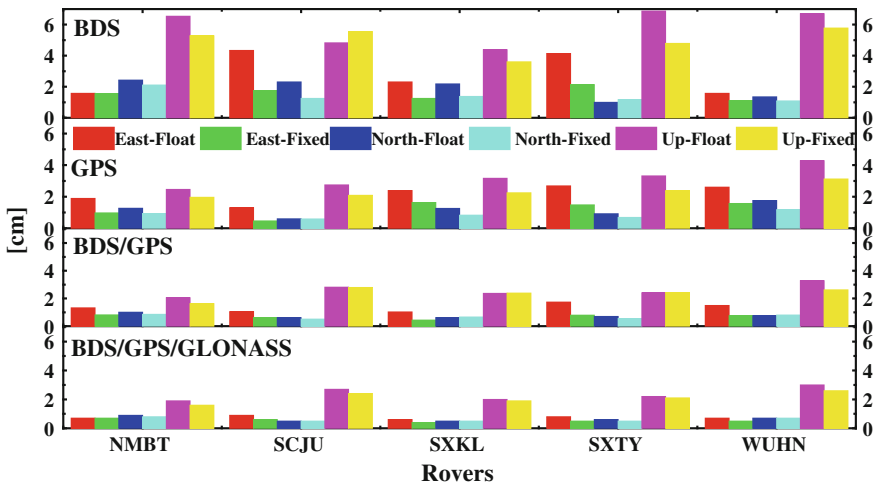


Fig. 2.6 RMS of float and fixed solutions of BDS, GPS, BDS/GPS and BDS/GPS/GLONASS PPP

Table 2.2 Average values of RMS of float and fixed solutions of BDS, GPS, BDS/GPS and BDS/GPS/GLONASS PPP

RMS	East (cm)		North (cm)		Up (cm)	
	Float	Fixed	Float	Fixed	Float	Fixed
BDS	2.8	1.6	1.9	1.4	5.9	5.0
GPS	2.2	1.2	1.2	0.8	3.2	2.4
BDS/GPS	1.3	0.7	0.7	0.7	2.6	2.4
BDS/GPS/GLONASS	0.7	0.5	0.6	0.5	2.3	2.1

18.8 % and 53.6, 63.2, 55.9 % than that of GPS-only and BDS-only PPP. The RMS of BDS/GPS/GLONASS PPP float solutions are better than 1, 1 and 3 cm, respectively. The average values of the RMS are 0.7, 0.6 and 2.3 cm, respectively. The accuracy improved 68.2, 50.0, 50.0 % and 75.0, 68.4, 61.0 % than that of GPS-only and BDS-only PPP and improved 68.4, 61.0, 11.5 % than that of BDS/GPS PPP.

- (2) Compared with float solutions, PPP fixed solutions improve greatly in the East, North and Up components. The average values of the RMS of BDS PPP fixed solutions are 1.6, 1.4 and 5.0 cm in East, North and Up components, respectively, which improve 42.9, 26.3 and 15.3 % than those of float ones. The average values of the RMS of GPS PPP fixed solutions are 1.2, 0.8 and 2.4 cm, respectively, which improve 45.5, 33.3 and 25.0 % than those of float ones. The average values of the RMS of BDS/GPS PPP fixed solutions are 0.7, 0.7 and 2.4 cm, respectively, which improve 43.8, 50.0, 52.0 % and 41.7, 12.5, 0.0 % than those of BDS-only and GPS-only fixed solutions, respectively. The RMS of BDS/GPS/GLONASS fixed solutions are 0.5, 0.5 and 2.1 cm, respectively, which improved 68.8, 64.3, 58.0 % and 58.3, 37.5, 12.5 % than those of BDS-only and GPS-only fixed solutions and improved 28.6, 28.6, 12.5 % than that of BDS/GPS fixed ones. They also improved 28.6, 16.7 and 8.7 % than those of BDS/GPS/GLONASS float ones;
- (3) Combined BDS/GPS PPP float solutions can achieve the accuracy of GPS PPP fixed solutions. BDS/GPS/GLONASS PPP float solutions can achieve the accuracy of BDS/GPS fixed solutions and beyond the accuracy of the GPS fixed solutions.

Figures 2.7 and 2.8 show the time series of the differences of BDS-only, GPS-only, BDS/GPS and BDS/GPS/GLONASS PPP float solutions and fixed solutions against “ground truth” at WUHN and SXKL station. As we can see that:

- (1) The differences of BDS PPP float and fixed solutions against “ground truth” are within ± 2 cm in East and North components after convergence, which are very close to those of GPS PPP float and fixed solutions. The Up component of BDS PPP float and fixed solutions are within ± 10 cm, which were worse than those of GPS. The North component of BDS PPP float solutions has the fastest convergence speed while the East component has the slowest speed. The reason may be that it takes a long time to make the ambiguity parameters converge. BDS PPP fixed solutions significantly reduce the convergence time. This is because phase observations were converted to precise distance measurement after ambiguities resolution. The parameters are estimated quickly and exactly which quickens the convergence speed.
- (2) The East and North components of GPS PPP float and fixed solutions are within ± 2 cm after convergence and the Up component is within ± 5 cm.
- (3) The East and North components of BDS/GPS PPP float and fixed solutions are within ± 1 cm after the convergence and the Up component is within ± 5 cm. The fusion of BDS/GPS PPP float solutions did not reduce the convergence

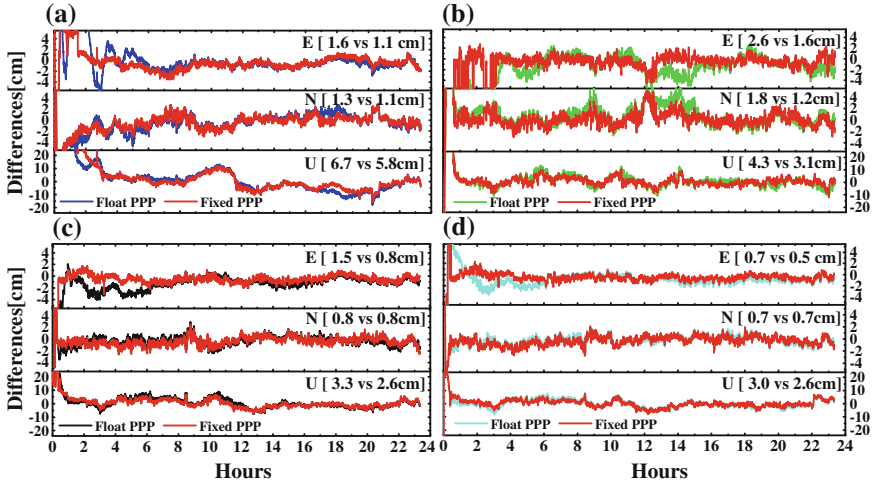


Fig. 2.7 Time series of float and fixed solutions of BDS-only, GPS, BDS/GPS and BDS/GPS/GLONASS PPP at WUHN. **a** BDS. **b** GPS. **c** BDS/GPS. **d** BDS/GPS/GLONASS

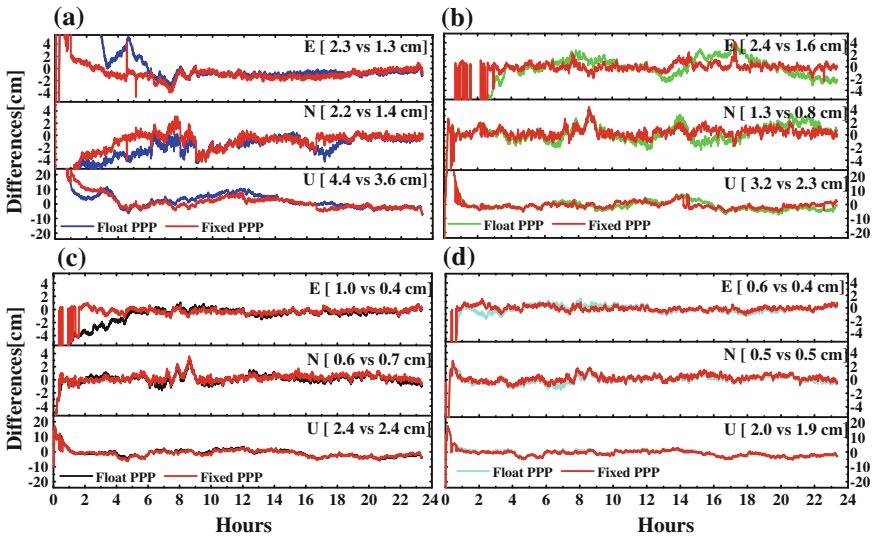


Fig. 2.8 Time series of float and fixed solutions of BDS-only, GPS-only, BDS/GPS and BDS/GPS/GLONASS PPP at SXKL. **a** BDS. **b** GPS. **c** BDS/GPS. **d** BDS/GPS/GLONASS

time significantly, which may due to the equal weighting of BDS and GPS observations.

- (4) The East and North components of BDS/GPS/GLONASS PPP float and fixed solutions are within ± 1 cm after the convergence and the Up component is within ± 5 cm.

- (5) The ambiguities resolution significantly accelerates the convergence speed and improves the positioning accuracy.

2.4.3 Systematic Bias Stability Analysis

Figure 2.9 shows the system time difference between BDS and GPS and between GLONASS and GPS at WUHN and SXKL station. We can see that the system time difference after convergence is relatively stable in the time domain. As Formulas 2.8 and 2.9 shown, the system time difference mainly depends on the pseudo-range hardware delay in the, which is quite stable in a day time.

2.4.4 Residual Analysis

Observation residuals contain measurement noise, multipath errors, orbit errors and dismodelled error such as receiver antenna PCV and PCO, which are important indicator of orbit determination and positioning accuracy [6]. Figure 2.10 shows the time series of BDS, GPS and GLONASS satellites LC and PC of all the 5 rovers.

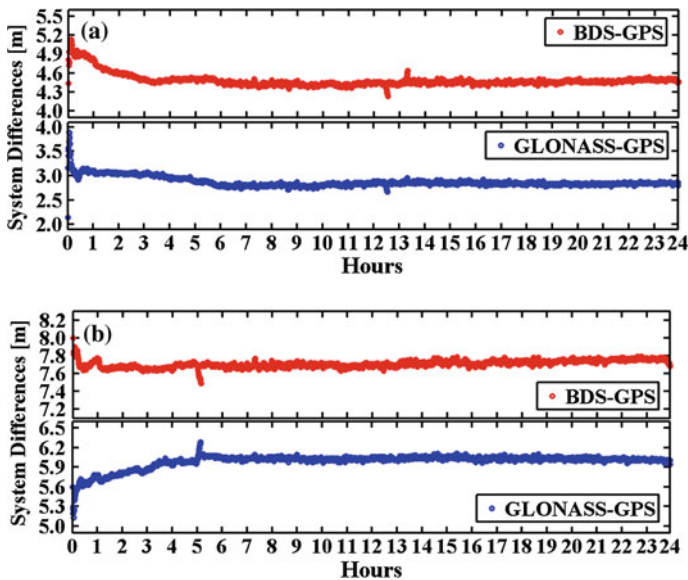


Fig. 2.9 Time series of system time difference between GLONASS and GPS and between BDS and GPS at WUHN and SXKL. **a** WUHN. **b** SXKL

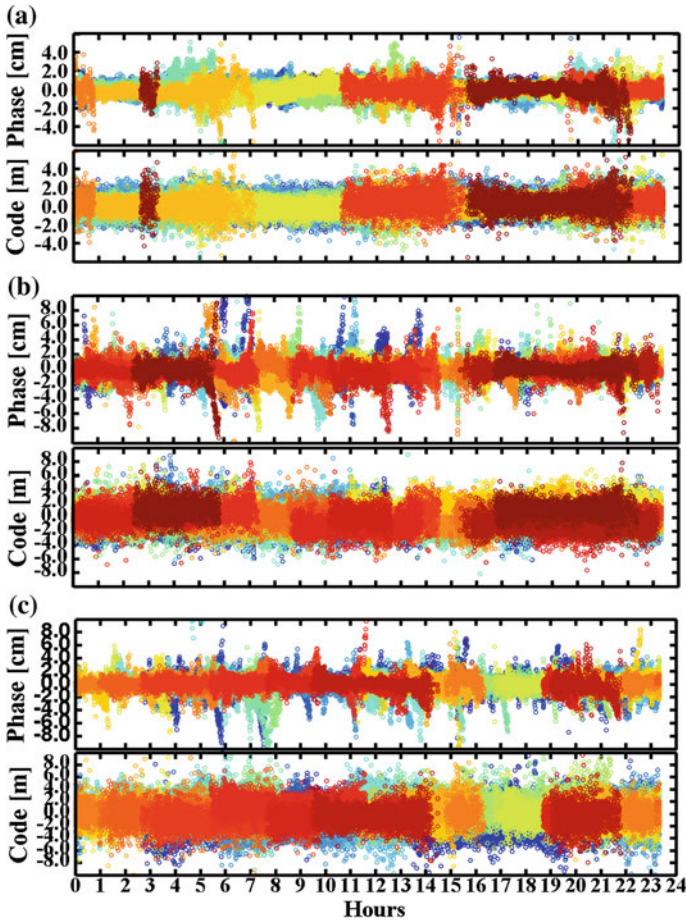


Fig. 2.10 Time series of BDS, GPS and GLONASS phase and code residuals of rovers. **a** BDS Residuals. **b** GPS Residuals. **c** GLONASS Residuals

Figure 2.11 shows the RMS of PC and LC of BDS, GPS and GLONASS satellites. It can be seen that the LC residuals of BDS, GPS and GLONASS are all within the ± 2 cm. The PC residuals of BDS satellite are within ± 2 m while those of GPS and GLONASS satellites are within ± 4 m. The average values of BDS satellites LC and PC residuals RMS are 0.66 and 0.83 cm, respectively, which are less than those of GPS satellites LC and PC residuals RMS while GLONASS satellites have the largest LC and PC RMS values. The reason may be that all of the five rovers locate in China, where satellites observed are mainly IGSO and GEO in a day time, whose ambiguities are more stable compared to MEO satellites, which fit their observations better.

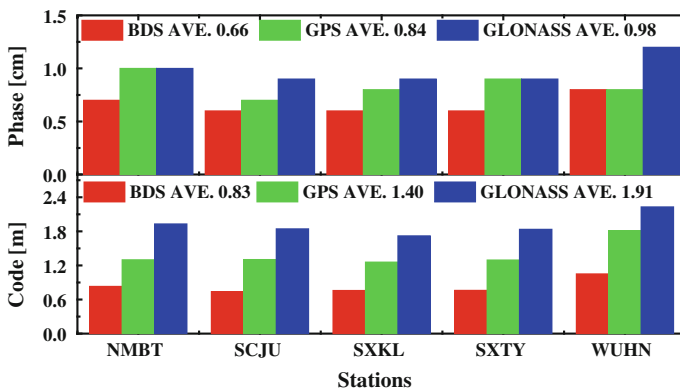


Fig. 2.11 RMS of BDS, GPS and GLONASS phase and code residuals of Rovers

2.5 Conclusions

In this article, we deduced the mathematical model of BDS/GNSS real-time kinematic PPP, studied the process and algorithm of un-differenced ambiguity resolution and realized the GNSS real-time kinematic PPP by fixing BDS-only, GPS-only and combined BDS/GPS ambiguities. Multi-GNSS data of reference stations from CMONOC were processed. We compared the accuracy of PPP float and fixed solution, studied the stability of wide-lane and narrow-lane FCB of BDS IGSO and MEO and GPS satellites, analyzed the stability of system time difference between BDS and GPS, GLONASS and GPS, respectively and compared the LC and PC residuals of BDS, GPS and GLONASS satellites. The following conclusions were acquired:

- (1) The wide-lane FCB of BDS MEO satellites keep stable in time domain within 0.2 cycles while those of IGSO satellites are within 0.2–0.4 cycles. The narrow-lane FCB of BDS MEO and IGSO satellites keep stable within 0.2 cycles after convergence. The wide-lane and narrow-lane FCB of GPS satellites keep stable in time domain and the changes stay within 0.2 cycles;
- (2) The fusion of multi-GNSS positioning raised the precision, reliability and continuity. The results of 5 rovers in this article showed that the precision were less than 1 cm in plane and 3 cm in vertical direction even without ambiguities resolution.
- (3) The precision of GPS-only and BDS-only fixed solutions improved a lot than those of the real solutions. The precision raised 1–2 cm in three-dimension. Fixed the ambiguities of GPS and BDS at the same time improved further positioning performance. The convergence speed was accelerated after ambiguities especially for the east component.
- (4) The system time differences between BDS and GPS and between GLONASS and GPS are very stable in the time domain.

We find that the fusion of multi-GNSS and the ambiguity resolution are the important ways to improve the accuracy of precise positioning and reduce the convergence time.

Acknowledgments We are grateful to the reviewers for their comments and suggestions. This work was supported by the “iGMAS” project of China, the National Nature Science Foundation of China (No. 41374034) and the National “863 Program” of China (Grant No. 2014AA123101).

References

1. China Satellite Navigation Office. BeiDou navigation satellite system signal In space interface control document open service signal B1I (version 1.0); 2012. <http://www.beidou.gov.cn/attach/2012/12/27/201212275f2be9ad57af4cd09c634b08d7bc599e.pdf>
2. Shi C, Zhao Q, Li M et al (2011) Precise orbit determination of Beidou satellites with precise positioning. *China Sci Earth Sci* 42(6):854–861
3. Shi C, Zhao Q, Hu Z, Liu J (2012) Precise relative positioning using real tracking data from COMPASS GEO and IGSO satellites. *GPS Sol*. doi: [10.1007/s10291-012-0264-x](https://doi.org/10.1007/s10291-012-0264-x)
4. Hauschild A, Montenbruck O, Sleewaegen J, Huisman L, Teunissen P (2012) Characterization of compass M-1 signals. *GPS Sol*. doi:[10.1007/s10291-011-0210-3](https://doi.org/10.1007/s10291-011-0210-3)
5. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Teunissen P, Nakamura S (2012) Initial assessment of the COMPASS/BeiDou-2 regional navigation satellite system. *GPS Sol*. doi: [10.1007/s10291-012-0272-x](https://doi.org/10.1007/s10291-012-0272-x)
6. Zhao Q, Guo J, Li M, Qu L, Hu Z, Shi C, Liu J (2013) Initial results of precise orbit and clock determination for COMPASS navigation satellite system. *J Geod*. doi: [10.1007/s00190-013-0622-7](https://doi.org/10.1007/s00190-013-0622-7)
7. Yang Y, Li J, Xu J, Tang J, Guo H, He H (2011) Contribution of the compass satellite navigation system to global PNT users. *Chin Sci Bull*. doi: [10.1007/s11434-011-4627-4](https://doi.org/10.1007/s11434-011-4627-4)
8. Cai CS, Gao Y (2012) Modeling and assessment of combined GPS/GLONASS precise point positioning. *GPS Sol*. doi: [10.1007/s10291-012-0273-9](https://doi.org/10.1007/s10291-012-0273-9)
9. Li M, Qu L, Zhao Q, Guo J, Su X, Li X (2014) Precise point positioning with the BeiDou navigation satellite system. *Sensors* 14(1):927–943. doi:[10.3390/s140100927](https://doi.org/10.3390/s140100927)
10. Qu L, Zhao Q, Li M, Guo J, Su X, Liu J (2013) Precise point positioning using combined BeiDou and GPS observations, vol 245, pp 241–252. doi:[10.1007/978-3-642-37407-4_22](https://doi.org/10.1007/978-3-642-37407-4_22)
11. Zumberge JF, Hefflin MB, Jefferson DC, Watkins MM, Webb FH (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. *J Geophys Res* 102(B3):5005–5017
12. Ge M, Gendt G, Rothacher M, Shi C, Liu J (2008) Resolution of GPS carrier-phase ambiguities in precise point positioning (PPP) with daily observations. *J Geod* 82(7):389–399
13. Collins P, Bisnath S, Francois L, He’roux P (2010) Undifferenced GPS ambiguity resolution using the decoupled clock model and ambiguity datum fixing. *Navigation* 57(2):123–135
14. Laurichesse D, Mercier F, Berthias J, Bijac J (2008) Real time zero difference ambiguities blocking and absolute RTK. In: Proceedings of the ION NTM-2008, Institute of Navigation, San Diego, California, pp 747–755
15. Geng J, Teferle F, Shi C, Meng X, Dodson A, Liu J (2009) Ambiguity resolution in precise point positioning with hourly data. *GPS Sol* 13(4):263–270
16. Li XX, Zhang XH, Li P (2012) PPP for rapid precise positioning and orbit determination with zero-difference integer ambiguity fixing. *Chin J Geophys* 55(3):833–840 (in Chinese) doi: [10.6038/j. ISSN. 0001-5733.2012.03.013](https://doi.org/10.6038/j. ISSN. 0001-5733.2012.03.013)
17. Geng J, Meng X, Dodson A, Teferle F (2010) Integer ambiguity resolution in precise point positioning: method comparison. *J Geod* 84(9):569–581

18. Shi J, Gao Y (2013) A comparison of three PPP integer ambiguity resolution methods. *GPS Solut* 18(4):519–528. doi:[10.1007/s10291-013-0348-2](https://doi.org/10.1007/s10291-013-0348-2)
19. Wanninger L (2011) Carrier-phase inter-frequency biases of GLONASS receivers. *J Geod* 86(2):139–148. doi:[10.1007/s00190-011-0502-y](https://doi.org/10.1007/s00190-011-0502-y)
20. Dong DN, Bock Y (1989) Global positioning system network analysis with phase ambiguity resolution applied to crustal deformation studies in California. *J Geophys Res Solid Earth* 94(B4):3949–3966. doi:[10.1029/JB094iB04p03949](https://doi.org/10.1029/JB094iB04p03949)
21. Wu J, Wu S, Hajj G, Bertiger W, Lichten S (1993) Effects of antenna orientation on GPS carrier phase. *Manuscr Geod* 18:91–98
22. Zhao Q, Liu J, Ge M, Shi C (2006) Applications of square-root information filtering and smoothing on orbit determination of LEO satellites with on-board GPS data. *Wuhan Univ J Geo Inf Sci* 31:12–15

Chapter 3

Characteristics Analysis of BeiDou Melbourne-Wübbena Combination

Xiyang He and Xiaohong Zhang

Abstract Melbourne-Wübbena (MW) combination, which has characteristics of geometry-free and ionosphere-free (first-order), has been often used for cycle slip detection and facilitating ambiguity resolution in GNSS dual-frequency data processing. The constellation of BeiDou Navigation Satellite System (BDS) is different from that of GPS. GPS constellation includes only Medium Earth Orbit (MEO) satellites, while BDS constellation is composed of three types of satellites: Geostationary Orbit (GEO), Inclined Geosynchronous Orbit (IGSO) and MEO satellites. We studied the characteristics of BDS MW combinations of different types of satellites with static observations. The results indicated that: all un-differenced MW combinations of BDS observations show obvious systematic biases. These biases cannot be cancelled out by single-differencing between satellites. On the contrary, they are amplified due to the superposition of the systematic biases of two satellites observations. By single-differencing between receivers, different characteristics are shown for different types of satellites and different baseline distances. The systematic biases of IGSO and MEO MW combination series can be cancelled out by single-differencing between receivers over short and medium baselines, however, they cannot be cancelled out over long baselines. For GEO MW combinations series, their systematic biases cannot be cancelled out over both short and long baselines. The further analysis of the DD (Double-differenced) MW combinations indicated that the systematic biases which cannot be cancelled out by single-differencing between receivers also cannot be cancelled out by double-differencing between both receivers and satellites. By summing up these results, we analyzed the sources of the systematic biases of BDS MW combination: the systematic biases of IGSO and MEO MW combination series mainly originate from multipath errors of satellite; the systematic biases of GEO MW combination series also originate from multipath errors, however, it is not sure that they originate from satellite multipath, receiver multipath, or the combined influence of both satellite and receiver multipath.

X. He (✉) · X. Zhang

School of Geodesy and Geomatics, Wuhan University, 129 Luoyu Road, Wuhan 430079,
People's Republic of China
e-mail: hexiyang704814@126.com

Keywords Melbourne-Wübbena (MW) combination · Beidou navigation satellite system (BDS) · Un-differencing · Single-differencing · Double-differencing · Numerical characteristics

3.1 Introduction

Chinese BDS has been built up at the end of 2012, which can provide PNT services. The current BDS is a regional navigation satellite system, including 5 GEO satellites, 5 IGSO satellites and 4 MEO satellites.

Some recent studies [5, 9–11] have discovered that there are systematic code-carrier divergences for BDS signals. These divergences can reach up to 0.4 m for high elevations, and even 1 m for low elevations. Other studies [13] analysed the characteristics of MW combination. They show that the fractional parts of mean values of IGSO MW combination for different passes remain stable while the scenarios for GEO and MEO satellites of BDS are different. Daily fluctuations can be seen in the time series of GEO MW combinations. Their results indicated that MW combination also suffer from the affection of these systematic biases. In order to address this problem, Wanninger and Beer [14] proposed an elevation-dependent correction model, which can be used for the correction of BDS IGSO and MEO code measurements. Their tests showed that remarkable improvement was observed after correction for single-frequency PPP (Precise Point Positioning) which is based on single-frequency ionosphere-free combination. However, this model provides only relative correction values, the real biases are still unknown.

For traditional GNSS positioning, one of the methods to eliminate the errors is implementing single-differencing between satellites, single-differencing between receivers or double-differencing between both receivers and satellites. With regarding to BDS code systematic errors, what characteristics will be shown after single-differencing or double-differencing? Whether the characteristics remain unchanged for different baseline distances? These problems are still unknown for us which are worth deep investigating.

Melbourne-Wübbena (MW) [7, 12] combinations eliminates both geometrical terms and first-order ionospheric delay, which plays an important role in GNSS precise positioning. In TurboEdit algorithm [2], MW combination is used for cycle slip detection. With regarding to long baseline relative positioning or PPP AR (ambiguity resolution), it is also used for WL (Wide-Lane) ambiguity resolution [1, 3, 4]. For these reasons, we analyze the characteristics of MW combinations which are un-differenced, single-differenced between receivers, single-differenced between satellites and double-differenced between both satellites and receivers respectively, thus providing guidance for GNSS data processing of precise positioning.

3.2 GNSS MW Combination Considering the Effects of Multipath and Hardware Delays of Satellite and Receiver

Considering the multipath and hardware delays, un-differenced GNSS phase and pseudorange observations of dual-frequency signals can be expressed as

$$L_1 = \lambda_1 \phi_1 = \rho - If_2^2 / (f_1^2 - f_2^2) + \lambda_1 (n_1 + \delta\phi_{1,r} - \delta\phi_1^s) + m_{L_1} + \varepsilon_L \quad (3.1a)$$

$$L_2 = \lambda_2 \phi_2 = \rho - If_1^2 / (f_1^2 - f_2^2) + \lambda_2 (n_2 + \delta\phi_{2,r} - \delta\phi_2^s) - \Delta\rho + m_{L_2} + \varepsilon_L \quad (3.1b)$$

$$P_1 = \rho + If_2^2 / (f_1^2 - f_2^2) + \delta P_{1,r} - \delta P_1^s + m_{P_1} + \varepsilon_P \quad (3.1c)$$

$$P_2 = \rho + If_1^2 / (f_1^2 - f_2^2) - \Delta\rho + \delta P_{2,r} - \delta P_2^s + m_{P_2} + \varepsilon_P \quad (3.1d)$$

Where ϕ_1 and ϕ_2 are the raw phases, L_1 and L_2 are the carrier phase ranges, P_1 and P_2 are the code pseudoranges, f_1 and f_2 are the signal frequencies, c is the speed of light, λ_1 and λ_2 are the signal wavelengths, I is the difference in ionospheric delay between L_1 and L_2 , ρ is the nondispersive delay, including the effects of geometric delay, tropospheric delay, clock signatures, and any other delay with affects all four observables identically, n_1 and n_2 are the phase ambiguities, $\delta\phi_{1,r}$ and $\delta\phi_{2,r}$ are the hardware delays of phase observables originating in the receiver, $\delta\phi_1^s$ and $\delta\phi_2^s$ are the hardware delays of phase observables originating in the satellite transmitter, $\delta P_{1,r}$ and $\delta P_{2,r}$ are the hardware delays of pseudorange observables originating in the receiver, δP_1^s and δP_2^s are the hardware delays of pseudorange observables originating in the satellite transmitter. $\Delta\rho$ is the differential delay between the L_1 and L_2 phase centers, which is correlated to satellite elevation and azimuth. ε represents the observational noise and the random part of multipath, m represents the systematic part of multipath.

According to (3.1a–d), the wide-lane phase delays can be written as

$$L_w = \frac{(f_1 L_1 - f_2 L_2)}{f_1 - f_2} = \rho + \frac{If_1 f_2}{f_1^2 - f_2^2} + \frac{c((n_1 - n_2) + (\delta\phi_{1,r} - \delta\phi_1^s) - (\delta\phi_{2,r} - \delta\phi_2^s))}{f_1 - f_2} - \frac{f_2 \Delta\rho}{f_1 - f_2} + \frac{(f_1 m_{L_1} - f_2 m_{L_2})}{f_1 - f_2} + \frac{\sqrt{f_1^2 + f_2^2}}{f_1 - f_2} \varepsilon_L \quad (3.2)$$

We also derived the following pseudorange combination

$$P_w = \frac{(f_1 P_1 + f_2 P_2)}{f_1 + f_2} = \rho + \frac{If_1 f_2}{f_1^2 - f_2^2} + \frac{[f_1 (1,r - \delta P_1^s) + f_2 (\delta P_{2,r} - \delta P_2^s)]}{f_1 + f_2} + \frac{f_2 \Delta\rho}{f_1 + f_2} + \frac{(f_1 m_{P_1} + f_2 m_{P_2})}{f_1 + f_2} + \frac{\sqrt{f_1^2 + f_2^2}}{f_1 + f_2} \varepsilon_P \quad (3.3)$$

Then, MW combination in cycles can be expressed as [1]

$$\begin{aligned}
 MW &= (L_w - P_w)/\lambda_w = [(n_1 - n_2) + (\delta\phi_{1,r} - \delta\phi_{2,r}) - (\delta\phi_1^s - \delta\phi_2^s)] \\
 &\quad - \frac{[f_1(\delta P_{1,r} - \delta P_1^s) + f_2(\delta P_{2,r} - \delta P_2^s)]}{(f_1 + f_2)\lambda_w} - \frac{2\Delta\rho f_1 f_2}{(f_1^2 - f_2^2)\lambda_w} \\
 &\quad + m_{MW} + \frac{\sqrt{f_1^2 + f_2^2}}{(f_1 + f_2)\lambda_w} \varepsilon_P
 \end{aligned} \tag{3.4}$$

where

$$\begin{aligned}
 \lambda_w &= \frac{c}{f_1 - f_2}, \\
 m_{MW} &= \frac{1}{(f_1 - f_2)\lambda_w} (f_1 m_{L_1} - f_2 m_{L_2}) - \frac{1}{(f_1 + f_2)\lambda_w} (f_1 m_{P_1} + f_2 m_{P_2}), \\
 2\Delta\rho \frac{f_1 f_2}{f_1^2 - f_2^2} &\approx 3.97\Delta\rho \text{ for GPS,} \\
 2\Delta\rho \frac{f_1 f_2}{f_1^2 - f_2^2} &\approx 3.86\Delta\rho \text{ for BDS.}
 \end{aligned}$$

Both GPS and BDS MW combination have large wavelength, which are 86.2 and 84.7 cm respectively. Because the pseudorange noise is much larger than the phase noise, we ignored the effect of phase noise in (3.4). If $\Delta\rho$ is no more than 1 cm, $2\Delta\rho \frac{f_1 f_2}{f_1^2 - f_2^2}$ is smaller than 4 cm, amounting to being smaller than 0.05 cycles. It implies that $\Delta\rho$ is much smaller than observational noise, therefore, we may safely neglect this term. MW combination, then, can be expressed as

$$\begin{aligned}
 MW &= [(n_1 - n_2) + (\delta\phi_{1,r} - \delta\phi_{2,r}) - (\delta\phi_1^s - \delta\phi_2^s)] \\
 &\quad - \frac{[f_1(\delta P_{1,r} - \delta P_1^s) + f_2(\delta P_{2,r} - \delta P_2^s)]}{(f_1 + f_2)\lambda_w} + \frac{\sqrt{f_1^2 + f_2^2}}{(f_1 + f_2)\lambda_w} \varepsilon_P + m_{MW}
 \end{aligned} \tag{3.5}$$

Based on the above equation, by differencing between receiver we have

$$\begin{aligned}
 \Delta MW &= [(\Delta n_1 - \Delta n_2) + (\Delta\delta\phi_{1,r} - \Delta\delta\phi_{2,r})] - \frac{(f_1 \Delta\delta P_{1,r} + f_2 \Delta\delta P_{2,r})}{(f_1 + f_2)\lambda_w} + \Delta m_{MW} \\
 &\quad + \frac{\sqrt{2(f_1^2 + f_2^2)}}{(f_1 + f_2)\lambda_w} \varepsilon_P
 \end{aligned} \tag{3.6}$$

Single differencing between receivers eliminates the hardware delays originating in the satellite transmitter.

Similarly, by differencing between satellites, we have

$$\begin{aligned} \nabla MW = & [(\nabla n_1 - \nabla n_2) - (\nabla \delta \phi_1^s - \nabla \delta \phi_2^s)] + \frac{1}{(f_1 + f_2)\lambda_w} (f_1 \nabla \delta P_1^s + f_2 \nabla \delta P_2^s) + \nabla m_{MW} \\ & + \frac{\sqrt{2(f_1^2 + f_2^2)}}{(f_1 + f_2)\lambda_w} \varepsilon_P \end{aligned} \quad (3.7)$$

Single differencing between satellites eliminates the hardware delays originating in the receiver.

Subtracting (3.6) with (3.5), we derived the double-differenced MW combination

$$\nabla \Delta MW = (\nabla \Delta n_1 - \nabla \Delta n_2) + \nabla \Delta m_{MW} + \frac{2\sqrt{(f_1^2 + f_2^2)}}{(f_1 + f_2)\lambda_w} \varepsilon_P \quad (3.8)$$

Double differencing between both receivers and satellites eliminates hardware delays originating from both satellites and receivers.

As can be seen from (3.5), both geometrical terms (such as geometry range, clocks and tropospheric delay) and first-order ionospheric delay are cancelled out. The residual terms are observable noise, multipath error and the differenced hardware delays of different signals. If hardware delays are stable and no cycle slips happen, assuming that the observable noise and multipath have white noise characteristic, then the derived MW combination would have white noise characteristic. This characteristic has been used for cycle slip detection and wide-lane ambiguity resolution in GPS data processing. The existing studies have demonstrated that the hardware delays are stable, however, the multipath error may have systematic characteristics.

3.3 Characteristics of BeiDou MW Combination

In order to analyze the characteristics of BDS MW combinations, we use four baselines data for our test. The distances of the four baselines are shown in Table 3.1. Three stations CUT0, CUT2 and CUTA are located in Curtin, Australia. Another two stations WHCD and WHHP are located in Wuhan, China. All the data were collected from February 22, 2014 to February 26, 2014 with a sampling rate 30 s using Trimble NetR9 receivers, which can receive both GPS and BDS

Table 3.1 Baseline information

Baseline	Distance
CUT0-CUT2	0 m
CUTA-CUT2	8.4 m
WHCD-WHHP	47.2 km
CUTA-WHHP	6617.6 km

dual-frequency signals. The characteristics of between-receivers single-differenced (SD) MW combinations, between-satellites single-differenced (SD) MW combinations and double-differenced (DD) MW combinations are analyzed.

3.3.1 UD MW Combinations

We first computed the un-differenced MW combination series of all visible satellites of the above five stations. The results show that the characteristics of BDS MW combinations differ from those of GPS MW combinations, which indicates that the BDS MW combinations indeed suffer from systematic biases of pseudoranges. Figure 3.1 shows the examples of un-differenced MW combinations of GPS satellites and BDS GEO, IGSO and MEO satellites. For clarity and convenient, only one continuous arc is given for G01 and C06 and C11. For BDS GEO satellites, since they usually can be observed continuously, five days MW series of C01 are given. As can be seen, the un-differenced MW combinations of GPS satellites have white-noise characteristics with a constant averaged value (the value of the noise is dependent

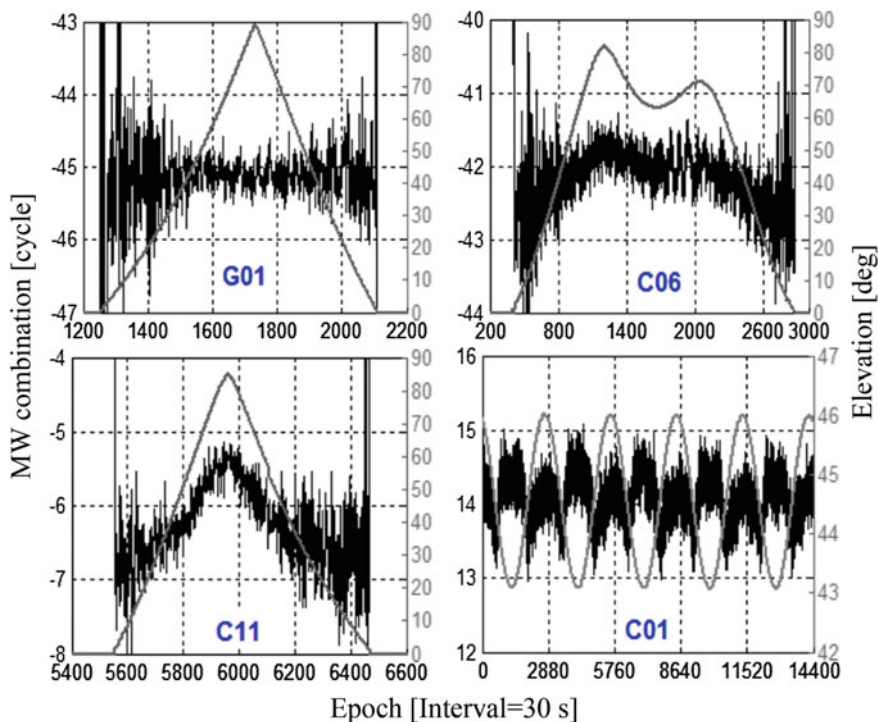


Fig. 3.1 UD MW combinations of station CUTA for different types of satellites

with satellite elevation). However, the un-differenced MW combinations of BDS GEO, IGSO and MEO satellites have systematic biases. The systematic biases of BDS GEO satellites have daily periodical variations. The systematic biases of BDS IGSO and MEO satellites are single-peak type and double-peak type respectively. Moreover, the form of the systematic biases is dependent with the satellite elevation.

From (3.4) we know that the systematic biases of BDS un-differenced MW combinations may originate from hardware delays of satellite or receiver, or multipath errors at satellite or receiver. It is well-known that single-differencing between receivers can eliminate or reduce the common errors from satellite, and single-differencing between satellites can eliminate or reduce the common errors from receiver. Therefore, we implement single-differencing between receivers, single-differencing between satellites and double-differencing between both receivers and satellites to BDS MW combinations, and analyze the characteristics of the generated SD and DD MW combination series.

3.3.2 Between-Satellites SD MW Combinations

From (3.7) we know that single-differencing between satellites can eliminate the effect of hardware delay of receiver. In order to confirm that if the systematic biases of MW combinations originate from receiver hardware delays, we obtained the between-satellites SD MW combinations with observations of same satellites types. Figure 3.2 shows the examples of the between-satellites SD MW combinations series of GPS MEO, BDS IGSO, BDS MEO and BDS GEO satellites of station WHCD. For BDS IGSO and MEO satellites, only a typical continuous arc is given for each satellite as examples. For GPS satellites, there are no obvious systematic biases for between satellites SD MW series, because there no obvious systematic biases for UD MW series. For BDS satellites, the systematic biases of MW combination series cannot be eliminated by single-differencing between satellites due to that there are no similar characteristics for BDS MW combinations of different satellites. Because of the superposition of the systematic biases of UD MW combinations of two BDS satellites, the systematic biases are amplified by single-differencing between satellites.

3.3.3 Between-Receivers SD MW Combinations

From (3.7) we know that single-differencing between receivers can eliminate the effect of hardware delay of satellite. In order to confirm if the systematic biases of MW combinations originate from satellite transmitter, we computed the between-receivers SD MW combination series with observations of three baselines

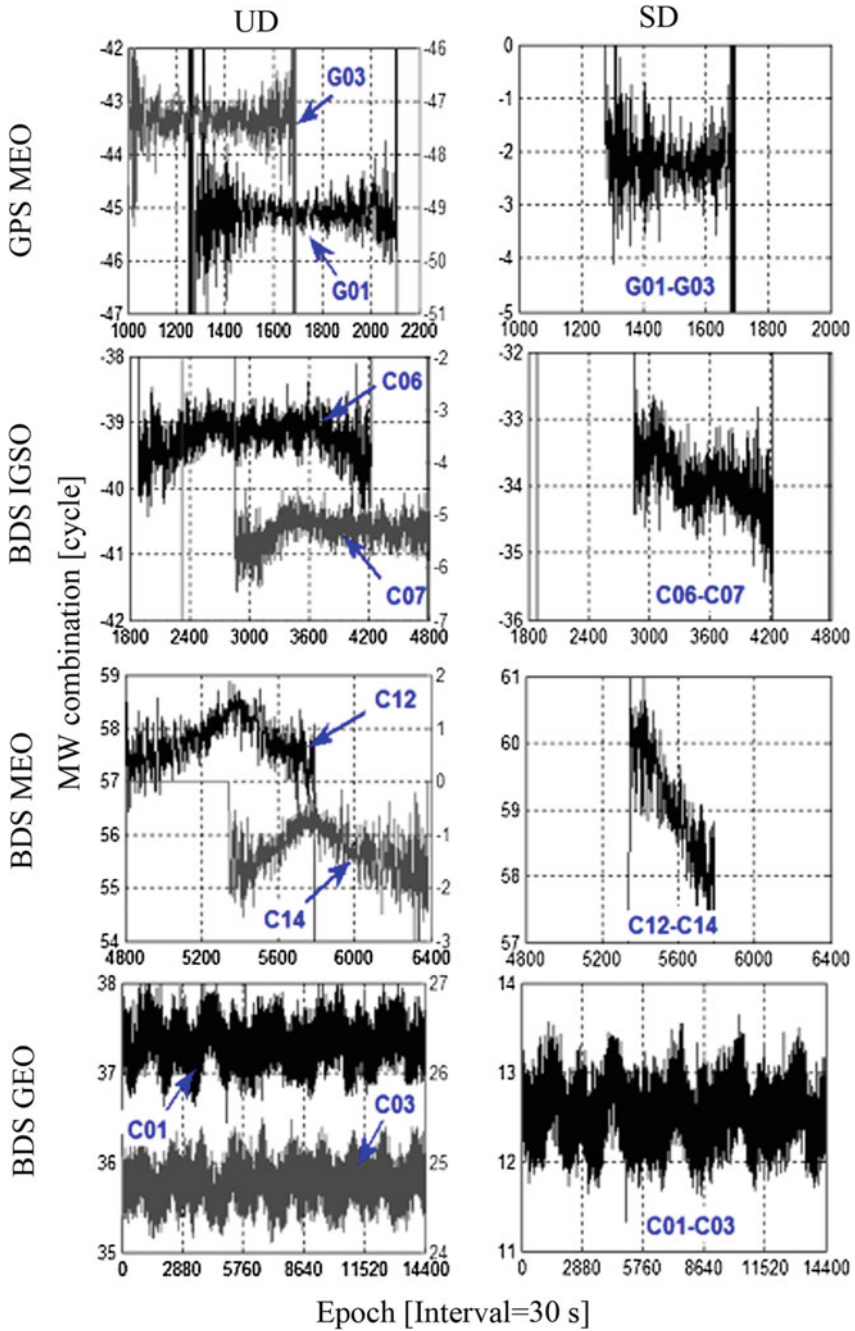


Fig. 3.2 Between-satellites SD MW combinations of station WHCD for different types of satellites

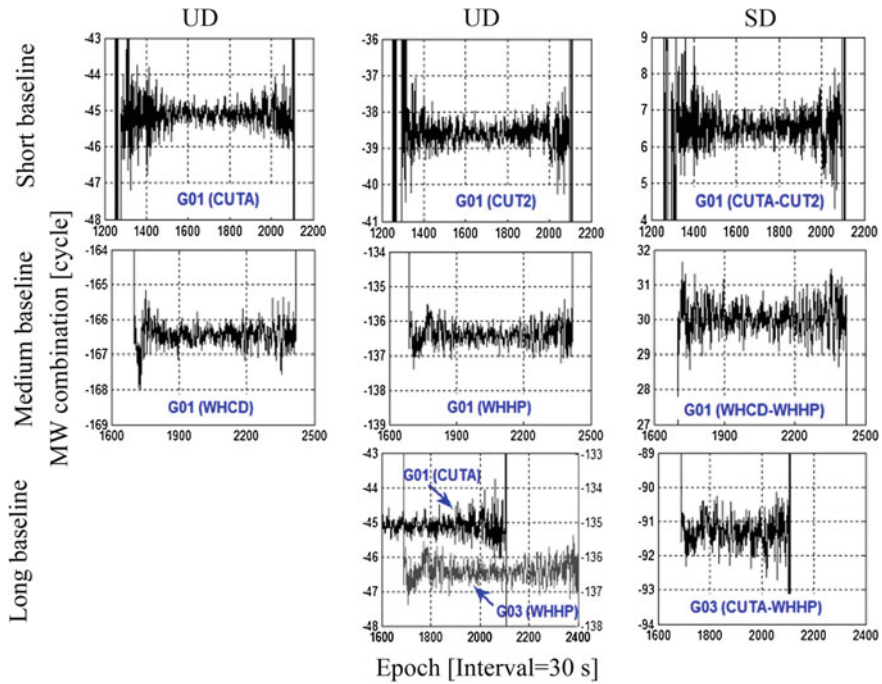


Fig. 3.3 Between-receivers SD MW combinations of GPS MEO satellites for different baseline distances

CUTA-CUT2, WHCD-WHHP and CUTA-WHHP, which are short, medium and long baselines respectively. Figure 3.3 plots the examples of between-receivers SD MW combination series of GPS MEO satellites for different baseline distances. Only one typical continuous arc is given for each satellite as examples for different distance sceneries. As can be seen from Fig. 3.3, there are no obvious systematic biases for between receivers SD MW series no matter how long the baseline distance is, because there no obvious systematic biases for UD MW series. Figures 3.4, 3.5 and 3.6 plots the examples of between-receivers SD MW combination series of BDS IGSO, MEO and GEO satellites for different baseline distances. For GPS MEO, BDS IGSO and MEO satellites, only one typical continuous arc is given for each satellite as examples for different distance sceneries. As can be seen from these figures, the systematic biases of IGSO and MEO MW combination series can be cancelled out by single-differencing between receivers over short and medium baselines, however, they cannot be cancelled out over long baselines. For GEO MW combinations series, their systematic biases cannot be cancelled out over both short and long baselines.

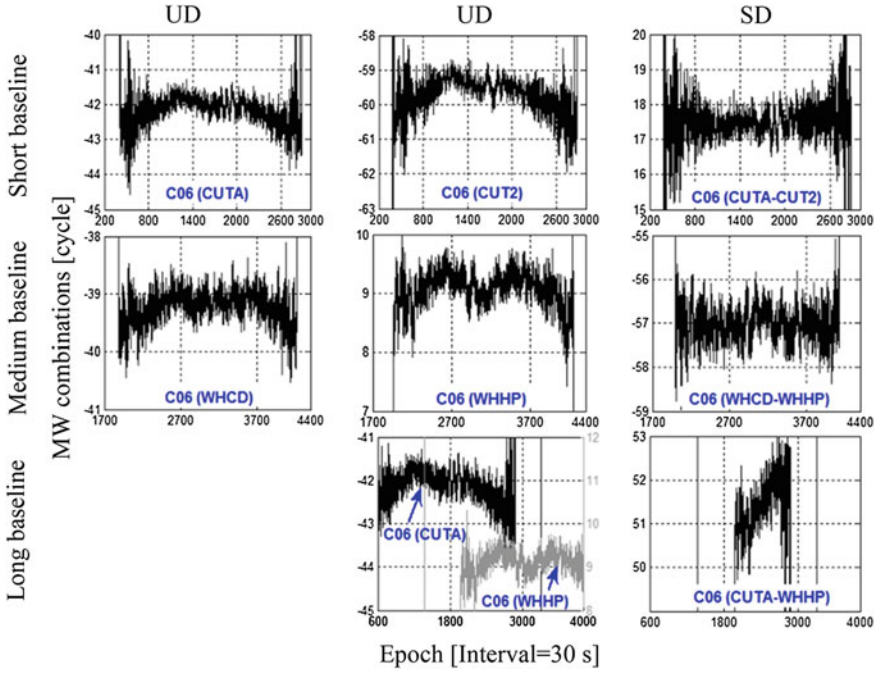


Fig. 3.4 Between-receivers SD MW combinations of BDS IGSO satellites for different baseline distances

3.4 Sources Analysis of the Systematic Biases of BDS Code Pseudoranges

We have confirmed that the systematic biases of BDS MW combination series are amplified by single-differencing between satellites. Since single-differencing between receivers can eliminate the effect of hardware delays originated in receiver, we can deduce that the systematic biases of MW combination series must be not related to the hardware delays originated in receiver. However, we still cannot deduce that the systematic biases of MW combination series are also not related to multipath originated in receiver, because the satellite elevation and azimuth of the two satellites used for differencing may be different, therefore the generated multipath even for a same receiver may be different. Hence, the possible sources of the systematic biases of MW combination series now include satellites hardware delays, satellite multipath or receiver multipath.

We also confirmed that the systematic biases of BDS IGSO and BDS MEO MW combination series can be cancelled out by single-differencing between receivers over short and medium baselines. It indicates that the systematic biases of BDS IGSO and BDS MEO MW combination series mainly originate from systematic errors of satellite. Moreover, these errors are dependent with the baseline distance

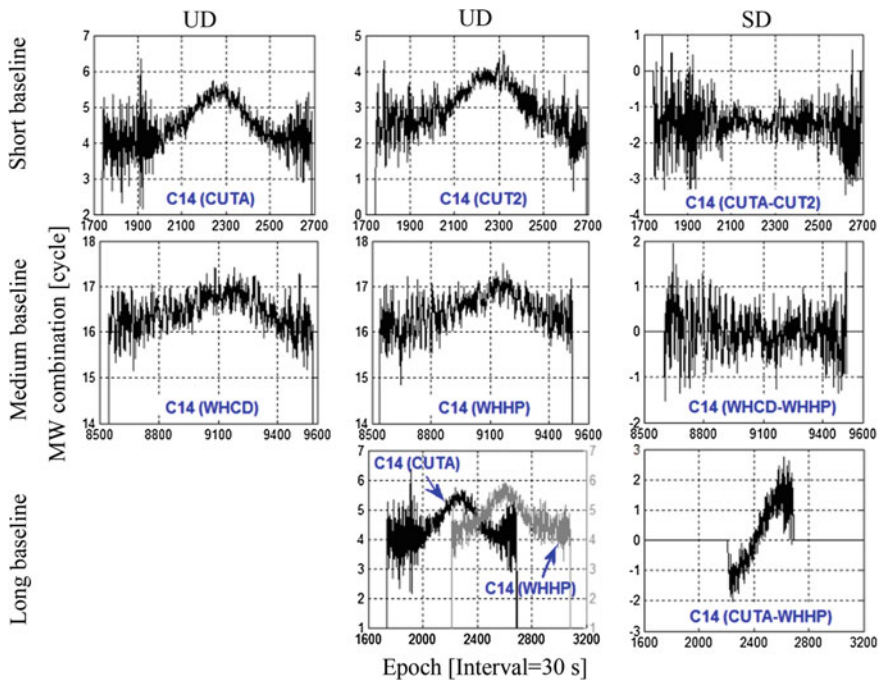


Fig. 3.5 Between-receivers SD MW combinations of BDS MEO satellites for different baseline distances

since the systematic biases of BDS IGSO and BDS MEO MW combination series cannot be cancelled out by single-differencing between receivers over long baselines. Considering that the satellite hardware delays are independent with the baseline distance, we can deduce that these systematic errors must not originate from satellite hardware delays. Summing up these deductions, we can conclude that the systematic biases of BDS IGSO and BDS MEO MW combination series mainly originate from multipath errors of satellite. For BDS GEO satellites, since their systematic errors cannot be cancelled out or reduced by differencing between receivers or satellites, we can deduce that these systematic errors must not originate from satellite or receiver hardware delays. Therefore, we can conclude that the systematic biases of BDS GEO MW combination series are also originate from multipath errors. Being different from BDS IGSO and BDS MEO satellites, we are not sure they originate from satellite multipath, receiver multipath, or the combined influence of both satellite and receiver multipath.

To further confirm our deduction, we computed the DD MW combination series as it can exclude the effect of satellite and receiver hardware delays simultaneously. Again three baselines CUTA-CUT2, WHCD-WHHP and CUTA-WHHP are used for our test. Figure 3.7 shows the examples of GPS MEO, BDS IGSO, BDS MEO and BDS GEO double-differenced MW combinations. For GPS MEO, BDS IGSO and BDS MEO satellites, only one typical continuous arc is given for each satellite

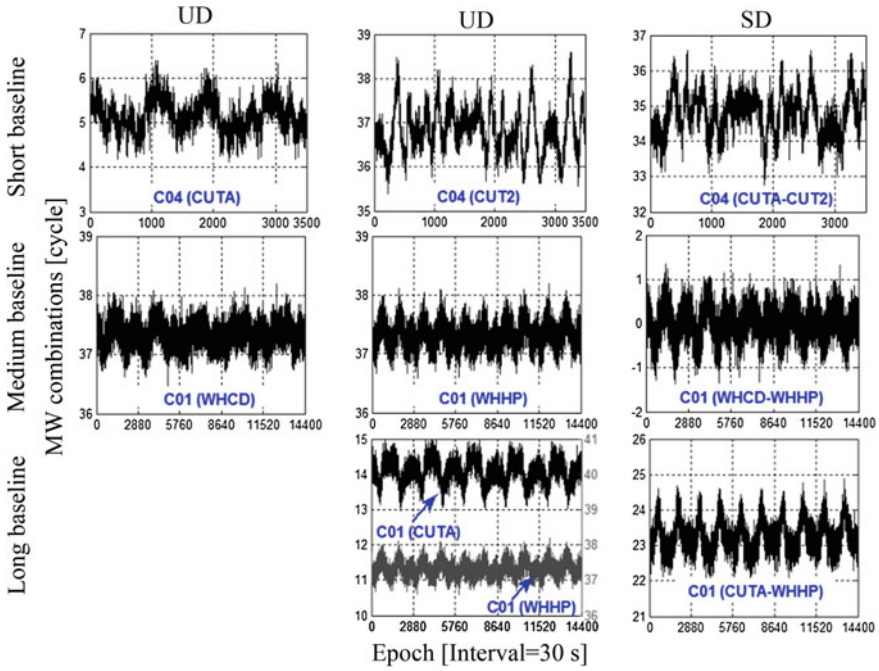


Fig. 3.6 Between-receivers SD MW combinations of BDS GEO satellites for different baseline distances

as examples for different sceneries. As can be seen, there are no systematic biases for GPS DD MW combination series no matter how long the baseline distance is. There are also no systematic biases for BDS IGSO and BDS MEO DD MW combination series over short and medium baselines. However, obvious systematic biases can be observed for BDS IGSO and BDS MEO DD MW combination series over long baselines and BDS GEO DD MW combination series over all three baselines. It indicates that there are similar characteristics for between-receivers SD MW combination series and DD MW combination series: the systematic biases which cannot be cancelled out by single-differencing between receivers also cannot be cancelled out by double-differencing between both receivers and satellites. Based on these results, we further confirmed that the systematic biases of BDS IGSO and BDS MEO MW combination series mainly originate from multipath errors of satellite. The correlations of these errors between receivers are weakening when the distance of the receivers are increasing. The results also further confirmed that the systematic biases of BDS GEO MW combination series originate from multipath errors. These biases do not show obvious spatial correlation.

The above analysis has demonstrated that the systematic biases of BDS GEO MW combination series originate from the multipath. In order to further verify this deduction, we tested it with the data of zero baseline CUT0-CUT2. Because a common antenna is used for two receivers of a zero baseline, the multipath

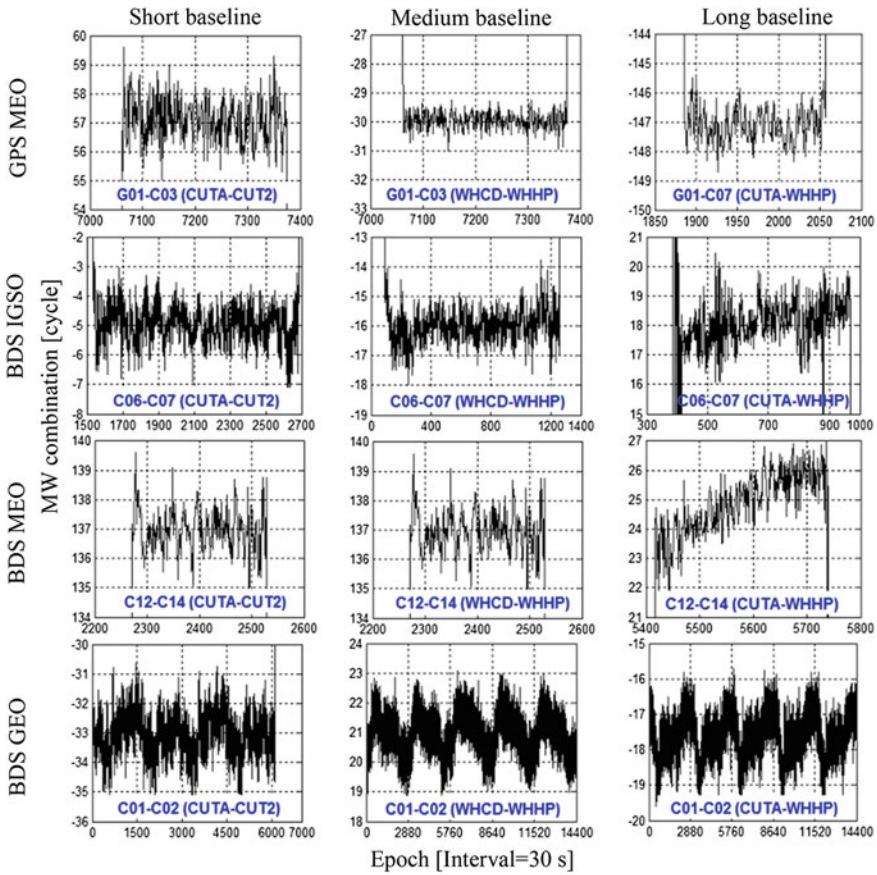


Fig. 3.7 DD MW combinations for different types of satellites and different baseline distances

generated in signals received by two receivers are completely the same. It means that if the systematic biases of MW combination originate from multipath, they will be cancelled out by single-differencing between receivers or double-differencing

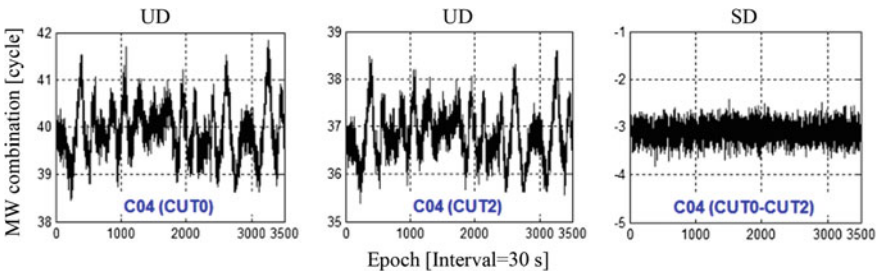


Fig. 3.8 Between-receivers SD MW combination series of BDS GEO satellite C04 of zero baseline CUT0-CUT2

Table 3.2 Existential state of the systematic biases of BDS code pseudoranges, ‘ $\sqrt{\quad}$ ’ indicates the systematic biases exist

Operation		IGSO	MEO	GEO
Single-differencing between satellites		$\sqrt{\quad}$	$\sqrt{\quad}$	$\sqrt{\quad}$
Single-differencing between receivers or double differencing	Zero baseline	\times	\times	\times
	Short baseline	\times	\times	$\sqrt{\quad}$
	Medium baseline	\times	\times	$\sqrt{\quad}$
	Long baseline	$\sqrt{\quad}$	$\sqrt{\quad}$	$\sqrt{\quad}$

‘ \times ’ indicates the systematic biases do not exist

between both satellites and receivers. Figure 3.8 shows the BDS GEO SD MW combination series computed using data of zero baseline CUT0-CUT2. As can be seen, the systematic biases are completely cancelled out by single-differencing between receivers. These results indicated that our deduction is correct.

Table 3.2 summarized the above analysis.

The characteristics of the BDS MW combination series indicated that: if BDS IGSO or BDS MEO code pseudorange observables are used for relative positioning, their systematic biases without having to be considered over short baselines; however, they must be considered over long baselines. For BDS GEO code pseudorange observables, their systematic biases must be considered over both short and long baselines. If one conducts un-differenced positioning or single-differenced positioning between satellites, systematic biases of all types of BDS satellites have to be considered. Since there is no appropriate correction model for BDS GEO satellites at present, it will be a future topic.

3.5 Conclusions

We computed MW combinations of different types of BDS satellites with static observations of four baselines. SD MW combinations and DD MW combinations are generated by single-differencing between satellites, single-differencing between receivers and double-differencing between both satellites and receivers. By analyzing the time series of these combinations, we draw the following conclusions:

- (1) Being different from GPS MW combinations, un-differenced MW combinations of all three types of BDS satellites show obvious systematic biases.
- (2) These biases cannot be cancelled out by single-differencing between satellites. On the contrary, they are amplified due to the superposition of the systematic biases of two satellites.
- (3) The systematic biases of BDS IGSO and BDS MEO MW combination series can be cancelled out by single-differencing between receivers over short and medium baselines, however, they cannot be cancelled out over long baselines. For BDS

GEO MW combinations series, their systematic biases cannot be cancelled out by single-differencing between receivers over both short and long baselines.

- (4) The further analysis of the DD MW combinations indicated that the systematic biases which cannot be cancelled out by single-differencing between receivers also cannot be cancelled out by double-differencing between both receivers and satellites.
- (5) By summing up these results, we analyzed the source of the systematic biases of BDS MW combination: the systematic biases of BDS IGSO and BDS MEO MW combination series mainly originate from multipath errors of satellite; the systematic biases of BDS GEO MW combination series also originate from the multipath errors, however, it is not sure that they originate from satellite multipath, receiver multipath, or the combined influence of both satellite and receiver multipath.

As the former studies and our work all demonstrated that BDS code measurements suffer from the severe affection of multipath, how this effect are produced and how to avoid this effect for the later BDS satellite platforms' designing would be an important issue needed to be paid attention to in the future.

References

1. Blewitt G (1989) Carrier phase ambiguity resolution for the global positioning system applied to geodetic baselines up to 2000 km. *J Geo Phys Res* 94:135–151
2. Blewitt G (1990) An automatic editing algorithm for GPS data. *Geophys Res Lett* 17(3):199–202
3. Bock Y, Gourewitch SA, Counselmann CC, King RW, Abbot RI (1986) Interferometric analysis of GPS phase observations. *Manuscripta Geodaet* 11(4):282–288
4. Ge M, Gendt G, Rothacher M, Shi C, Liu J (2008) Resolution of GPS carrier-phase ambiguities in precise point positioning (PPP) with daily observations. *J Geod* 82(7):389–399
5. Hauschild A, Montenbruck O, Sleewaegen JM, Huisman L, Teunissen P (2012) Characterization of compass M-1 signals. *GPS Solut* 16:117–126. doi:10.1007/s10291-011-0210-3
6. Melbourne W (1985) The case for ranging in GPS-based geodetic systems. In: *First international symposium on precise positioning with the global positioning system*, Rockville, pp 373–386
7. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Riley S (2012) A compass for Asia: first experience with the BeiDou-2 regional navigation system. Poster at IGS workshop in Olsztyn, Poland 2012. <http://www.igs.org/presents/poland2012/posters.html>. Accessed 10 July 2014
8. Montenbruck O, Rizos C, Weber R, Weber G, Neilan R, Hugentobler U (2013) Getting a grip on multi-GNSS—the international GNSS service MGEX campaign. *GPS World* 24(7):44–49
9. Perello Gisbert JV, Batzilis N, Risueño GL, Rubio JA (2012) GNSS payload and signal characterization using a 3 m dish antenna. In: *Proceedings ION GNSS 2012*, Nashville, 347–356
10. Wang (2014) Analysis of Characteristics of BDS Observable Combinations for wide-Lane integer ambiguity resolution In: Sun J, Jiao W, Wu H, Shi C (eds) *Proceedings of China satellite navigation conference (CSNC) 2014*
11. Wübbena G (1985) Software developments for geodetic positioning with GPS using TI-4100 code and carrier measurements. In: *First international symposium on precise positioning with the global positioning system*, Rockville, pp 403–412
12. Wanninger L, Beer S (2014) BeiDou satellite-induced code pseudorange variations: diagnosis and therapy. *GPS Solut*. doi: 10.1007/s10291-014-0423-3

Chapter 4

Refining of BDS Differential Code Bias Model

Qiankun Liu, Lifen Sui, Guorui Xiao, Yu Gan, Guobin Qi
and Tian Zheng

Abstract The differential code bias (DCB) in Global Navigation Satellite System (GNSS) satellites should be precisely determined using real ground GNSS data when designing certain applications, such as precise positioning, navigation and timing. This paper analyzed the relationship between satellite elevation angle with DSNR (difference signal to noise ratio) of BDS (BeiDou Navigation Satellite System) different orbital satellite. The experimental results showed that the DSNR of IGSO satellite was less than 5dBHz and more stable, much better than the other satellites; the change of GEO satellite DSNR was small, but the absolute value was large; MEO satellite DSNR changed frequently and big. A new algorithm to get the BDS observation weight is introduced by considering the distance between the satellite and the Earth and DSNR. The results indicated that the new weight algorithm improved the DCB accuracy and enhanced the stabilities, which improved GEO satellites 1–15 %, IGSO satellites 4.5–16 %, MEO satellites 20–22 %.

Keywords DCB · Quality analysis · BDS · Stochastic model · SNR

4.1 Introduction

Different frequencies navigation information transmit link is different, which lead that different channel delay is different from each other. Describing the difference between the channel delay parameter is called differential code bias (DCB). Studies have shown that DCB is an important error source affected the total electron content (TEC) calculation and precision positioning [1]. Currently DCB solver function model is a hot, but stochastic model is rare [1–5]. Satellite elevation angle and signal to noise ratio (SNR) can reflect the quality of the observations. The stochastic

Q. Liu (✉) · L. Sui · G. Xiao · Y. Gan · G. Qi · T. Zheng
School of Surveying and Mapping, Information Engineering University,
Zhengzhou 450000, China
e-mail: lqk@whu.edu.cn; 137764581@qq.com

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015
Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_4

model based on satellite elevation angle and SNR can effectively weaken the atmospheric residual error and multi-path error. What's more, the stochastic model based on SNR can also weaken the influence of diffraction [6]. GPS is composed of 24 Medium Earth Orbit (MEO) satellites, while BDS consists of three kinds of satellites that are Geostationary Orbit (GEO) satellites, Inclined Geosynchronous Satellite Orbit (IGSO) satellites and Medium Earth Orbit (MEO) satellite [7]. It is a problem that cannot be ignored in the DCB solver that different orbits satellite has different properties and signal quality of different frequencies is differences [8, 9]. This paper analysed the differences between the three BeiDou satellite signal qualities, then proposed a weight algorithm that is suitable for BDS DCB resolution, finally compared the effects of different stochastic model for DCB solver and calculated the BeiDou satellite differential code biases preliminary.

4.2 DCB Estimation Method

GNSS pseudorange observations of frequency i is expressed as:

$$P_{i,r}^s = \rho_{i,r}^s - c \cdot \delta t_r + c \cdot \delta t^s + \frac{40.28}{f_i^2} \cdot STEC + T^s + d_{hd(i,P),r} + d_{hd(i,P)}^s + d_{mult(i,P)}^s + \varepsilon_{i,P}^s \quad (4.1)$$

where s is sequence number of satellites, r is sequence number of receiver, $\rho_{i,r}^s$ is the geometric distance between the satellite s and the receiver r , STEC is slant total electron content, c is speed of light in vacuum (m/s), δt_r is clock bias for station r , δt^s is clock bias for satellite s , T^s is tropospheric delay, $d_{hd(i,P)}^s$, $d_{hd(i,P),r}$ are the satellite and receiver pseudorange code hardware delays of frequency i , $d_{mult(i,P)}^s$ is pseudorange multipath error of frequency i , $\varepsilon_{i,P}^s$ is pseudorange measurement noise.

Difference the pseudorange observation equations at B1 and B2 frequencies; it will get a value of no geometric combination P_4 .

$$P_{4,r}^s = F \cdot STEC + c \cdot (DCB^s + DCB_r) + \varepsilon_{P,r}^s \quad (4.2)$$

where $F = 40.28 \cdot \left(\frac{1}{f_1^2} - \frac{1}{f_2^2}\right)$, $DCB^s = d_{hd(1,P)}^s - d_{hd(2,P)}^s$, $DCB_r = d_{hd(1,P),r} - d_{hd(2,P),r}$ are satellite and receiver differential code bias.

As the pseudorange measurements have larger noise, usually carrier phase smoothed pseudorange measurements are used as data preprocessing strategies to improve pseudorange accuracy. Before pseudorange measurements are smoothed, cycle slip detection and gross errors elimination should be done.

According to the spherical harmonic function model to describes the ionospheric delay [10]. Make the difference between ionospheric residuals combination P_4 and

ionospheric delays as pseudorange observations in least square equations to get the satellite and receiver DCB, there have the observation equations as follow [1]:

$$\begin{cases} L + V = A \cdot X_{DCB} \\ X_{DCB} = [X_{DCB}^s, X_{DCB,r}] \\ L = \tilde{P}_{i,j}^s - F \cdot STEC \end{cases} \quad (4.3)$$

where V are the residuals of pseudo observations, A is design matrix, X_{DCB} are the estimated parameters including satellites and receivers DCB, $\tilde{P}_{i,j}^s$ is the smoothed pseudorange combination.

Since the design matrix A is rank-deficient of one, Eq. (4.3) is underestimated if no priori satellite or receiver DCB reference value is introduced. Under different constraints, satellites or receivers DCB are different, but the combination of receiver and satellite DCB don't change [11]. This paper adopts the method that the sum of all BeiDou satellites DCB is 0, as follows:

$$\begin{cases} S \cdot X_{DCB}^s = 0 \\ S = [1, \dots, 1]_{1 \times n} \end{cases} \quad (4.4)$$

According to the basic principle of rank deficient least squares with constrained conditions, we can get the following equations:

$$\begin{cases} X_{DCB} = (N + S^T S)^{-1} U \\ N = A^T P A \\ U = A^T P L \\ \sigma_0^2 = \frac{V^T P V}{n - n_1 - n_2} \\ Q_{XX} = \sigma_0^2 \cdot (N + S^T S)^{-1} \\ P = \text{diag}(p_1, \dots, p_{n^s}) \end{cases} \quad (4.5)$$

where σ_0^2 is posteriori unit-weighted mean error of the least squares adjustment, Q_{XX} is the covariance matrix of satellites and receivers DCB, n is the observation numbers, n^s is number of observed satellites at station r , n_r is the number of stations involved in solution, P weight matrix for pseudo observations.

4.3 DCB Estimation Stochastic Model Refining

The constellation of BDS is composed of three kinds of satellite. GEO satellites orbital height is 35,786 km [13]; the range accuracy is the highest [9, 12]; the orbit accuracy is the lowest [14]; the DCB residual is the maximum [1]; the propagation path is the longest when satellite elevation angle is equal to others [5]. IGSO satellites orbital height is 35,786 km; the orbit accuracy is the highest; its DCB

residual is minimum; its propagation path is about as same as GEO satellite's when satellite elevation angle is equal to others. MEO satellites orbital height is 21,528 km; its orbit accuracy and DCB accuracy are somewhere in between GEO and IGSO; its range accuracy as same as IGSO satellites.

Based on the above considerations, the relationships between the BeiDou different satellites orbital are complex. Elevation angle and SNR don't well describe the observations quality of different orbital satellite. This paper presents a mixed weight algorithm suitable for BDS DCB solver, as follows:

$$p = S \cdot \frac{\sin^2(e)}{\rho^2} \quad (4.6)$$

where e is satellite elevation angle, ρ is distance between the satellite and the receiver (km), the scaling factor s is an empirical values that defined by Eq. (4.7).

SNR is generally 45Db-Hz, consistent with the change of satellite elevation angle. Signal quality is good when the signal to noise ratio is large [4, 5]. Ionospheric residuals combination P_4 is the difference of two frequencies pseudorange. Signal qualities of different frequencies are not identical. Elevation angle model cannot fully describe the observations quality because the DSNR is always changing (see Fig. 4.1). Observations value should be considered have large errors and reduced the weight when DSNR is large, though the satellite elevation angle is big. This paper analyzed the DSNR between B1 and B2 frequency of three orbiting satellites. The result showed that IGSO satellites DSNR are no more than 4. Through trial and error analysis, the results were better when the difference of

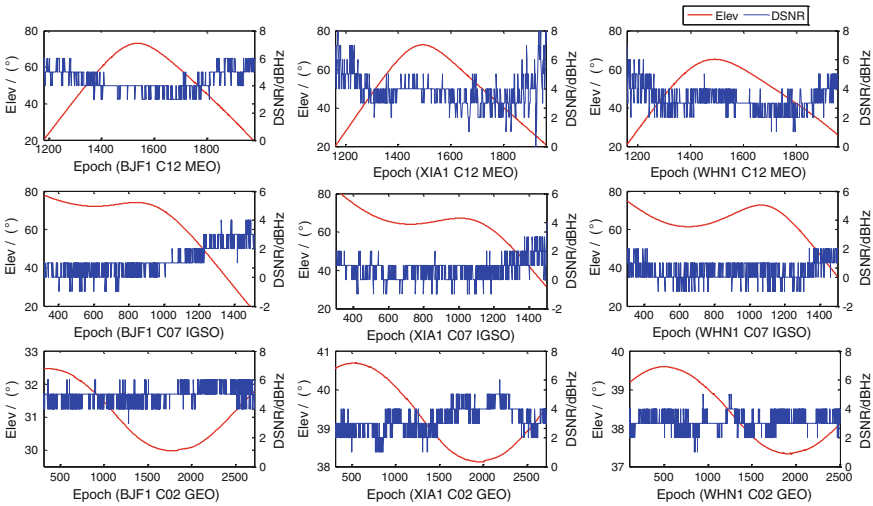


Fig. 4.1 DSNR changes with elevation of MEO, IGSO and GEO satellites of B1, XIA1 and WHN1 station in 125 Day, 2013

SNR1 and SNR2 took 4 as threshold. When DSNR is greater than 4, take 4 as a molecular appropriate to reduce observation power ratio.

$$S = \begin{cases} 1, & |SNR2 - SNR1| \leq 4 \\ \frac{4}{|SNR2 - SNR1|}, & \text{others} \end{cases} \quad (4.7)$$

4.4 Experiment and Result Analysis

4.4.1 BDS SNR and Elevation Angle Analysis

This paper selects the observation of BJJ1, XIA1, WHN1 stations in IGMAS 125 days of 2013, the sampling interval is 30 s, the cutoff elevation is 20°. Data pre-processing marks the position of cycle slip and excludes the gross errors in observations measurements. Experimental analyses the relationship between elevation angle and DSNR of three orbiting satellites. The results of PRN 2 GEO satellites, PRN 7 IGSO satellites and PRN 12 MEO satellites are showed in Fig. 4.1. $DSNR = SNR2 - SNR1$ is the difference of SNR2 and SNR1. Elev is satellite elevation angle.

It can be seen from Fig. 4.1:

- (1) When the elevation angles are equal, IGSO satellites DSNR are smaller than GEO and MEO satellites and not more than 4; it shows a low frequency change; it changes in a small range. MEO satellites DSNR are large; it shows a high frequency variation, it changes in a big range. GEO satellites are somewhere in between. There are great differences between the different stations.
- (2) GEO satellites DSNR change rapidly along with the change of elevation angle, followed with the IGSO satellite. MEO satellites DSNR change very frequent and deeply when elevation angle is less than 40°. IGSO satellites DSNR sometimes are smaller than 0, namely SNR2 is less than SNR1. GEO and MEO satellite SNR2 both are bigger than SNR1 every time. Here we can obtain that it is impossible to judge the observations quality of different frequency by simple comparison with SNR size.
- (3) There is a strong correlation between DSNR and elevation angle, namely the bigger the elevation angle, the smaller the DSNR value. When elevation angle is greater than 60°, DSNR remains stable. But DSNR change frequently, elevation model does not fully describe the changes. Sometimes there is a great difference of DSNR between consecutive epochs. Although the elevation angle is equal, quality of different frequency pseudorange observations cannot be exactly the same. Elevation model cannot full description the quality of P_4 pseudorange. DSNR can better reflect the local variations of the P_4 pseudorange.

4.4.2 Weight Algorithm Analysis

This paper selects the observation measurements of stations in IGMAS 125–132 days of 2013; the sampling interval is 30 s; the cutoff elevation is 10° . Data pre-processing marks the position of cycle slip and excludes the gross errors in observations measurements. Experiments use the following three weight algorithm program to estimate DCB:

Scheme 1: use sine function based elevation angle as weight algorithm.

$$p = \sin^2 e;$$

Scheme 2: use sine function based elevation angle and distance between the satellite and the receiver combination as the weight algorithm. $p = \sin^2 e / \rho^2$;

Scheme 3: use the elevation angle, distance and SNR scaling factor as the weight algorithm that is defined by Eqs. (4.6) and (4.7).

Table 4.1 shows the DCB weighted average value of the three kinds of schemes in consecutive 8 days. Figure 4.2 shows the standard deviations of different schemes.

Considering Table 4.1 and Fig. 4.2, we can draw the following conclusions:

- (1) We can see from DCB standard deviation of 125–132 days in 2013, except C04, C05 satellites, all satellites DCB standard deviation are less than 0.4 ns. BeiDou satellites DCB are stability. GEO satellites DCB residuals are larger than others, especially on the C04, C05 that are significantly bigger than the other satellites. A number of days estimation results also show that C04 and C05 satellite are biggest in single-day change, which may be due to that it is the farthest distance and low elevation angle that results in poor signal quality. Paper [3] has come to similar conclusions.

Table 4.1 BDS satellites DCB of three schemes

PRN	DCB (schemes 1)	DCB (schemes 2)	DCB (schemes 3)
1	15.73	15.69	15.7
2	6.67	6.67	6.68
3	4.82	4.92	4.94
4	4.88	4.79	4.77
5	-0.89	-0.94	-0.88
6	0.17	0.18	0.2
7	5.4	5.43	5.43
8	2.16	2.17	2.17
9	-6.56	-6.54	-6.54
10	-6.92	-6.91	-6.95
11	-7.7	-7.65	-7.65
12	-6.29	-6.21	-6.21
13	-7.45	-7.55	-7.57
14	-4.02	-4.05	-4.09

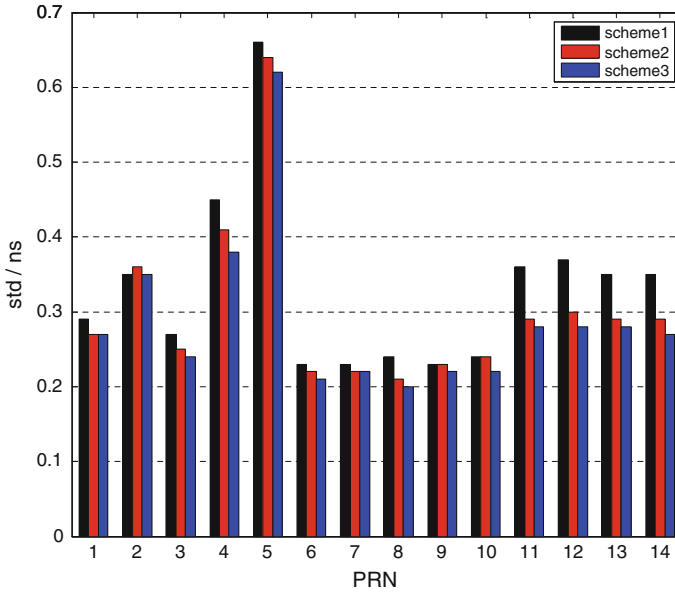


Fig. 4.2 BDS satellites DCB standard deviations

- (2) Different stochastic model has a great effect on the DCB solutions. The weight algorithm taking into account the distance reduces the result standard deviations by increase the MEO satellites weighted, improving the MEO satellites DCB accuracy particularly.
- (3) The new weight algorithm considering the DSNR and distance can better describe the observation measurements quality of different frequency. It improves the results accuracy by increase appropriately the weight of IGSO and MEO satellites compared to the other programs. It improves the GEO satellite 1–15 %, IGSO satellite 4.5–16 %, MEO satellites 20–22 %.

Figure 4.3 shows the DCB change sequence in consecutive 8 days of three estimate schemes of BDS C02 GEO satellites, C06 IGSO satellites and C12 MEO satellites DCB.

From Fig. 4.3, we can conclude that:

- (1) IGSO satellites DCB are the most stable, not more than 0.1 ns in single day change; GEO satellites DCB are generally within 0.4 ns; MEO satellites DCB stabilities are somewhere in between and no more than 0.3 ns.
- (2) Different ways to set weight algorithm results in a certain influence that the differences of a single day estimate results may reach in 0.2 ns. The new weight algorithm is more reasonable to describe the quality of the observed values and enhances the stability of single day DCB.

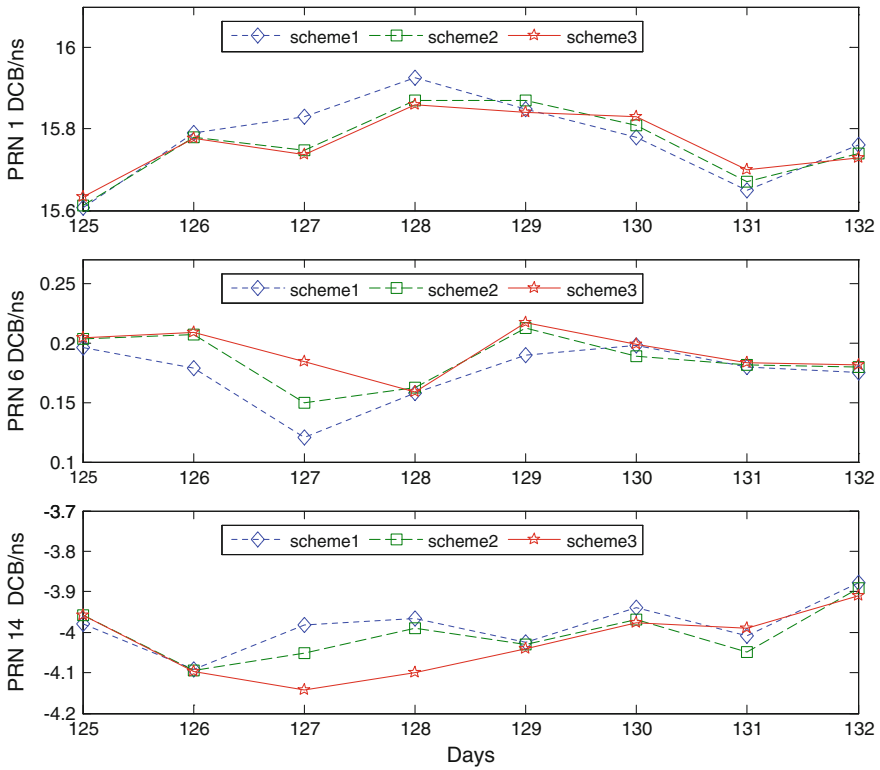


Fig. 4.3 DCB variations of partial BDS satellites in 8 days

4.5 Conclusions

This paper described the DCB estimation function model, analyzed the differences of different orbit satellites' signal quality, proposed a new weight algorithm based on distance between the satellite and the receiver and SNR difference of B1 and B2. Researches show that:

- (1) IGSO satellites DSNR are small and relatively stable, better than the MEO and GEO satellites significantly.
- (2) The new weight algorithm considered the distance and DSNR can describe the observations quality of different frequency well, improve estimate accuracy and enhance the stability. It has greatly improved the accuracy of the GEO satellites and IGSO satellites, especially for MEO satellites. I believe that with the increase of MEO satellites in BDS, the new weight algorithm will be continue to improve the accuracy of calculation.

Acknowledgments This work was partially sponsored by National Natural Science Foundation of China (Grant No. 41274016, Grant No. 41174006 and Grant No. 40974010).

References

1. Li ZS, Yuan Y, Li H et al (2012) Two-step method for the determination of the differential code biases of COMPASS satellites. *J Geodesy* 86(11):1059–1076
2. Jin R, Jin S, Feng G (2012) M_DCB: Matlab code for estimating GNSS satellite and receiver differential code biases. *GPS Solut* 16(4):541–548
3. Fan JC, Wu XL, Li YX et al (2013) COMPASS satellites DCB parameter accuracy assessment based on tri-frequency data. *Chin Space Sci Technol* 33(4):62–70
4. Wu XL, Ping JS, Liu L et al (2011) Hardware delay resolution for regional satellites navigation system. *Geomatics Inf Sci Wuhan Univ* 36(10):1218–1221
5. Zhang Q, Zhao Q, Zhang H et al (2014) BDS satellites and receivers DCB resolution. In: *China satellite navigation conference (CSNC) 2014 proceedings, vol III*. Springer, pp 187–197
6. Dai WJ, Ding XL (2008) Zhu JJ (2008) Comparing GPS stochastic models based on observation quality indices. *Geomatics Inf Sci Wuhan Univ* 33(7):1–9
7. Yuanxi Yang (2010) Contribution and challenges of compass/BeiDou satellites navigation system. *Acta Geod Cartogr Sin* 39(1):1–6
8. Zhang XH, Ding LL (2013) Quality analysis of the second generation compass observables and stochastic model refining. *Geomatics Inf Sci Wuhan Univ* 38(7):832–836
9. Xiao GR, Sui LF, Liu CJ et al (2014) A method of determining the weight matrix for BeiDou navigation satellite system single point positioning. *Acta Geod Cartogr Sin* 43(9):902–907
10. Yuan YB, Ou JK (2005) A generalized trigonometric series function model for determining ionospheric delay. *Prog Nat Sci* 15(8):1015–1019
11. Xie YB, Chen JP, Wu JC et al (2014) Ionospheric electron content and hardware delay analysis based on different constraints. *Geomatics Inf Sci Wuhan Univ* 33(7):799–803
12. Cheng PF, Li W, Bei JZ (2012) Precision analysis of BeiDou range measurement signals. *Acta Geod Cartogr Sin* 41(5):690–695
13. Beijing China Satellites Navigation Office (2013) BeiDou satellites navigation system signal in space interface control document
14. Liu WP (2014) Research on precise orbit determination of BeiDou navigation satellite system. Information Engineering University, Zhenzhou

Chapter 5

Estimation Strategy and Accuracy

Analysis of GNSS Real-Time Precise Satellite Clock Error

Liang Chen, Changjiang Geng, Quan Zhou and Wenhai Jiao

Abstract The rapid development of the Beidou navigation satellite system (BDS) accelerates the user demands for high-precision position and navigation. International GNSS Monitoring and Assessment System (iGMAS) built by China independently provides a good platform for GNSS high-precise real-time products. In this paper, the precise satellite clock error resolving model and strategy are studied, the feasibility of real-time precise clock error estimation on the premise of guarantee accuracy is researched, and the precision of real-time precise clock error using iGMAS site resources (built and planning) is verified. The results show: There is convergence process in the Un-difference mode to realize real-time clock estimation and the convergence RMS accuracy comparing to IGS is about 0.2 ns; Epoch single-difference model eliminates the continuous ambiguity, so there is no convergence process, the precision is better than 0.2 ns; On the whole, the precision of Un-difference mode and epoch single-difference model is quite; From the calculation efficiency and high-precision clock error application, the epoch single-difference model is adapt to GNSS real-time precise clock estimation because of less resolved parameters; Using iGMAS site resources (built and planning) recently, the precision of clock estimated by epoch single-difference model is about 0.2 ns.

Keywords iGMAS · Precise clock estimation · Un-difference · Epoch single-difference · Accuracy analysis

L. Chen (✉) · C. Geng · Q. Zhou
Test and Assessment Research Center, China Satellite Navigation Office,
1 Fengyingdong Road, Beijing 100094, China
e-mail: sdkdchenliang@163.com

L. Chen · C. Geng · Q. Zhou
GNSS Engineering Center, China Academy of Aerospace Electronics Technology,
1 Fengyingdong Road, Beijing 100094, China

W. Jiao
China Beijing Institute of Tracking and Telecommunication Technology,
Beijing 100094, China

5.1 Introduction

With the providing services of Beidou navigation satellite system for Asia-Pacific area and the development of the European Galileo navigation satellite system, there are more choices for navigation users. In order to assessment the GNSS open service signal performance, the concept of monitoring and assessment of GNSS Open Services is present by China in 2010, and the international GNSS Monitoring and Assessment System (iGMAS) has begun to be built [1]. By the end of 2014, iGMAS has completed the first-phase construction. By design, iGMAS consists of more than 30 global tracking stations, three data centers, several analysis centers et al. [2].

With the development of GNSS technology, real-time and high-precision products become one of the developing directions. Currently, the precise satellite orbit and clock error products have been provided by IGS, even though the update speed and accuracy of the ultra-orbit products have been meet the real-time users' requirements, the ultra-clock error accuracy is about 3 ns [3] which cannot meet the high-precision demands. In 2002, the real-time products plan was carried out by IGS, which can provide 0.3 ns real-time clock error product now [4].

As one of analysis centers (Test and Assessment Center, TAC) of iGMAS, depending on our high-precision products, the precise satellite clock error resolving model, strategies and accuracy analysis are mainly studied in this paper. On the premise of guarantee accuracy, the feasibility of achieving real-time clock error resolving based on epoch single-difference model is demonstrated. On this basis, using iGMAS site data and IGS data nearby iGMAS planning site, real-time clock error product accuracy provided by iGMAS recently is verified.

5.2 The Estimation Model of Precise Satellite Clock Error

At present, there are two kinds of methods to obtain the precise satellite clock error. One is to estimate the satellite clock error and orbit parameter together in the process of satellite precise orbit determination, the other is to estimate precise clock error in which the precise satellite orbit, station coordinates and the earth rotation parameters have been fixed as products in process of satellite precise orbit determination. Because of an excess of estimated parameters (such as orbit, earth rotation parameters, troposphere, clock, ambiguity, site coordinate and so on) in normal equation, the calculation efficiency is not quick and data processing interval is large normally (300 s), so subtly precise satellite clock cannot be get in the first processing. In the second method, the real-time clock error can be calculated by fixing some parameters as ultra-product such as the satellite orbit, the earth rotation parameter, and troposphere etc., in which processing efficiency is raised due to the number of parameters estimated in normal equation reduced.

The principle and processing strategy of the un-difference model and epoch single-difference model based on the second method are studied, and its accuracy (take GPS for example) is compared and analyzed in next paper.

5.2.1 The Un-difference Model

In the dual-frequency data processing, the ionosphere-free combination is usually used to eliminate the effect of ionosphere. The observation equation is as follows [5]:

$$\begin{cases} v_{k,PC}^s = \rho_k^s + c(dt_k - dT^s) + D_{k,trop}^s + \varepsilon_{k,PC}^s - P_{k,LC}^s \\ v_{k,LC}^s = \rho_k^s + c(dt_k - dT^s) + \lambda_{LC} \cdot N_{k,LC}^s + D_{k,trop}^s + \varepsilon_{k,LC}^s - \lambda_{LC} \cdot \Phi_{k,LC}^s \end{cases} \quad (5.1)$$

Among them, λ , Φ , N is no ionosphere combination wavelength, phase observations and ambiguity, respectively; P is GNSS code observations; ρ is the station-satellite geometrical distance; dt and dT is receiver and satellite clock error; c is the speed of light; D_{trop} is troposphere delay; ε is the observation noise etc.

Using the observation error equation established by the formula (1), real-time precise satellite clock error is estimated by Square Root Information Filter algorithm. In this processing, satellite orbit and station coordinates are fixed; Ambiguity is treated as a constant when there are no cycles or repaired and when cycle occurs, ambiguity is estimated as one new parameter; antenna phase center offset model is the newest IGS08 model; Receiver clock error and satellite clock error parameters are both treated as white noise. Detailed estimation strategy is shown in Table 5.1.

5.2.1.1 The Epoch Single-Difference Model

There are a lot of ambiguity parameters in un-differenced model, which affect the real-time clock error estimation efficiency. The ambiguity parameters are eliminated in epoch single-differenced model, which is more suitable for the real-time clock error estimation. The observation equation is as follows:

$$\begin{cases} v_{k,PC}^s(i, i+1) = \Delta\rho_k^s(i, i+1) + c(\Delta dt_k(i, i+1) - \Delta dT^s(i, i+1)) + \Delta D_{k,trop}^s(i, i+1) \\ \quad + \varepsilon_{k,PC}^s(i, i+1) - \Delta P_k^s(i, i+1) \\ v_{k,LC}^s(i, i+1) = \Delta\rho_k^s(i, i+1) + c(\Delta dt_k(i, i+1) - \Delta dT^s(i, i+1)) + \Delta D_{k,trop}^s(i, i+1) \\ \quad + \varepsilon_{k,LC}^s(i, i+1) - \lambda_{LC} \cdot \Phi_k^s(i, i+1) \end{cases} \quad (5.2)$$

Among them, $\Delta\rho_k^s(i, i + 1)$ is the station-satellite geometrical distance changes between adjacent epoch; station and satellite coordinates are fixed as precise result from precise orbit determination in this process. So the parameters estimated in epoch single-difference model include $\Delta dt_k(i, i + 1)$, $\Delta dT^s(i, i + 1)$, $\Delta D_{k,trop}^s(i, i + 1)$, and its meaning the changes of receiver clock error, satellite clock error and the troposphere changes on relative to upper epoch, respectively.

From above equation, we can get that the parameters to be estimated in pseudo-range and phase observation are the same. In order to minimize the impact on the precision of clock error, the weight of pseudo-range observation can be reduced. The change value of satellite clock error is estimated actually in this model, so the initial satellite clock error need to be introduced in this model, which is get from broadcast satellite clock error in this paper. The specific strategy is as Table 5.1.

Table 5.1 The estimation strategy of the GNSS clock error

	Parameter	Model
Observation information	Observation	LC, PC
	Observation intervals	30 s
	Priori information	P1:3.0 m; L1:0.02 cycle
	Elevation mask angle	7°
	Observation weight	p = 1, elev > 30° p = 2sin(elev), elev ≤ 30°
Error correction	Phase winding	correct
	PCV	IGS08 model correction
	Tide correction	Model correction
	Relativistic effects Earth rotation	Model correction
Parameters estimation	Reference clock	ALIC
	Orbit	Fixed as precise orbit products
	Site coordinates	Fixed as precise site coordinates
	EOP	Fixed as IERS products
	Troposphere	Saastanmoine Model + white noise
	Clock error of Satellite	Estimation, Broadcast clock error + white noise, priori-constraint 30,000 m
	Clock error of receiver	Estimation, Pseudorange estimation + white noise, priori-constraint 9000 m
Ambiguity	Un-difference estimation/Epoch single-difference elimination	

5.3 The Comparison Method of Clock Error

Compared with IGS final precise clock error, quadratic difference method of clock error assessment method is adopted in this paper. Firstly, one satellite is selected as reference satellite (PRN1 in this paper), in order to eliminate the reference clock differences selected in estimation strategy between clock error estimated in this paper and IGS, other satellites clock errors result got from this paper and IGS final products make a difference with this reference satellite. Use this result, the other difference is done between same satellites of those two products to get quadratic difference, then quadratic differential time series RMS is statistic as follows [6]:

$$RMS = \sqrt{\frac{\sum_{i=1}^n (\Delta_i - \bar{\Delta})^2}{n}} \quad (5.3)$$

Among them, Δ_i is the quadratic difference of i epoch, $\bar{\Delta}$ is the average. Quadratic differential result contains one system deviation, which is due to reference clock error. For users of double-difference data process, initial satellite clock bias have been eliminated in difference between satellites, will not affect the positioning result; In un-difference data processing, the weight of pseudorange observation is very low, those bias can be thought as been absorbed by the carrier phase ambiguity, so ignore the influence of pseudorange deviation [4]. Quadratic difference RMS can reflect the degree of coincidence between clock results estimated in this paper with IGS final clock products effectively.

5.4 Analysis of Test Results

In order to analyze the characteristics, compare precise clock error accuracy and estimation efficiency of the two models, global 40 IGS tracking stations (see Fig. 5.1) observation data from October 17 to 23, 2014 (Day of year 290–296) is analyzed by software developed by ourselves in this article. In order to imitate the real time mode, the GPS orbit is fixed as GPS ultra-orbit products provided by our iGMAS analysis center.

5.4.1 The Comparison of the Precise Clock Error Accuracy

Figure 5.2 shows the precise clock error accuracy with IGS clock on October 17, 2014 (Day of year 290) by un-difference and epoch single-difference model, respectively (take PRN 12 and 15 for example). It can be seen from the results that there is a convergence process due to need to estimate the ambiguity parameters in

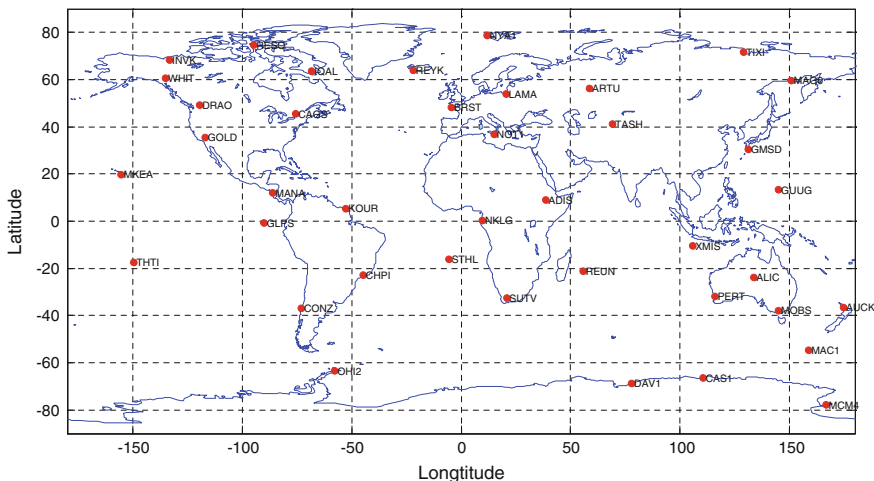


Fig. 5.1 40 IGS tracking stations distribution around the world

un-difference model and the result is stable after convergence (see as Fig. 5.2b). However, because epoch single-difference model has eliminated the ambiguity parameter, there is no convergence process (see as Fig. 5.2b). There is an obvious constant-bias in single-difference model; because the single-difference model estimates the satellite clock differences relative to the upper epoch, the existence of clock error of first epoch causes the bias.

Figure 5.3 is the statistical RMS comparison of un-difference model and epoch single-difference model from DOY 290–296, 2014. The RMS of un-difference

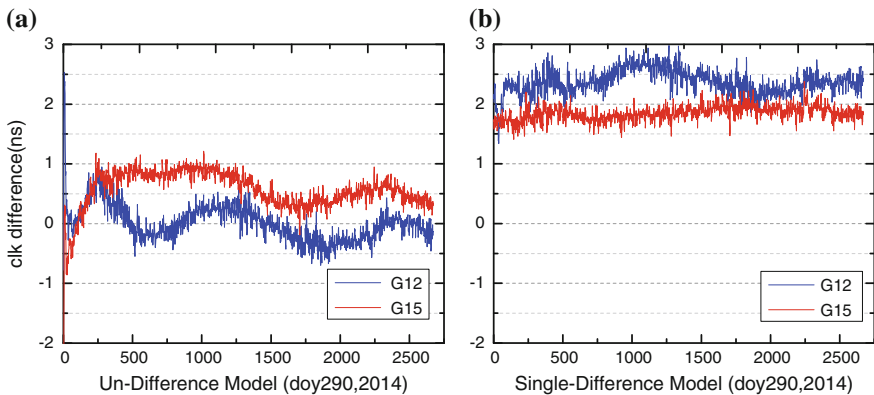


Fig. 5.2 Diagram of precise clock error estimated by un-difference model and epoch single-difference model comparing with IGS

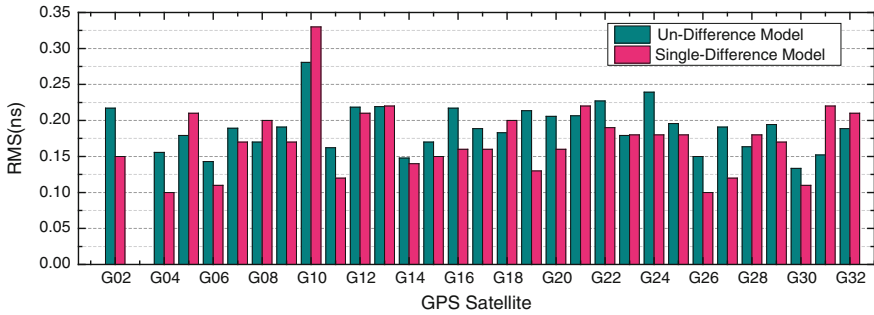


Fig. 5.3 The difference between the model and the epoch single difference model than the RMS statistics

model is the result after convergence and there are no PRN 3 results due to stoppage in experiment. From the results, clock error RMS accuracy estimated by two models comparing with IGS results is about 0.2 ns, most satellites accuracy of epoch single-difference model is superior to 0.2 ns. On the whole, the clock error accuracy estimated by the two models is about the same.

5.4.2 Estimation Efficiency Analysis

Due to irregularity of the satellite clock changing, the real-time clock error estimation system needs higher estimation speed in order to meet the demands of high precision in GNSS high-accuracy real-time applications. In order to verify the clock error data processing efficiency of the two methods, about 40 stations observation data on October 17, 2014 (DOY290) are processed by our clock estimation software in the IBM X3850 server (processor Inter(R) Xeon(R) E7-4820 2.0 GHz, 16 g memory, Windows sever2008 64-bit operating system), and the time comparison results between un-difference model and epoch single-difference model for calculating each epoch time are shown in Fig. 5.4.

From Fig. 5.4, calculation consumption time of un-difference model is more than epoch single-difference model due to large of parameters in normal equation, most is in 0.8–0.9 s, however, the epoch single-difference model is about 0.1–0.2 s. There are some larger consumption times epochs in the two models, and un-difference model's are more than epoch single-difference's, which may be caused by the poor data quality, the reason need to further study.

Combining the accuracy and estimation efficiency of the two models of clock error resolving, and considering the update requirements of real-time clock error at the same time, epoch single-difference model is more suitable for GNSS real-time clock error calculation.

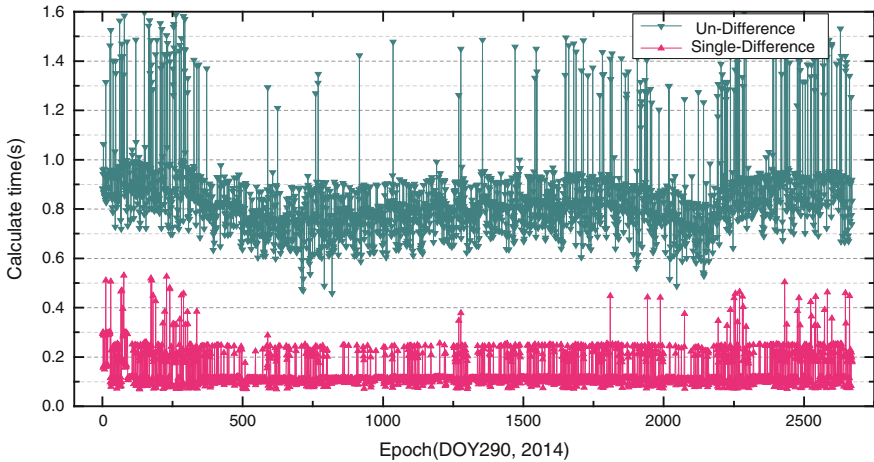


Fig. 5.4 Each epoch calculation consumption time of un-difference model and epoch single-difference model

5.4.3 Validation of iGMAS Real-Time Clock Error Estimation

By the end of 2014, iGMAS has completed the first phase of construction. The image below is the part of iGMAS’s sites having been built, looking forward to building and planning. Due to iGMAS tracking stations are in the construction stage, so in the next part, using iGMAS site data and IGS data nearby iGMAS planned site, real-time clock error product accuracy provided by iGMAS recently is verified by epoch single-difference model (Figs. 5.5 and 5.6).

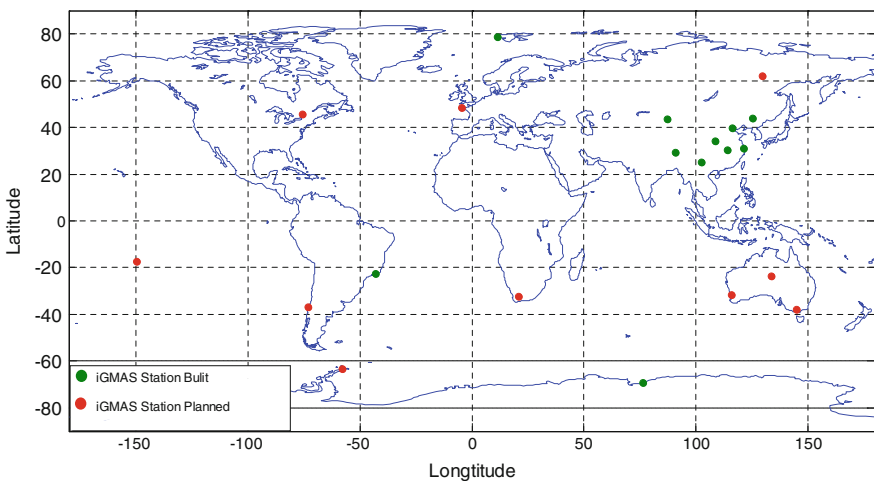


Fig. 5.5 iGMAS the existing, building and planning site

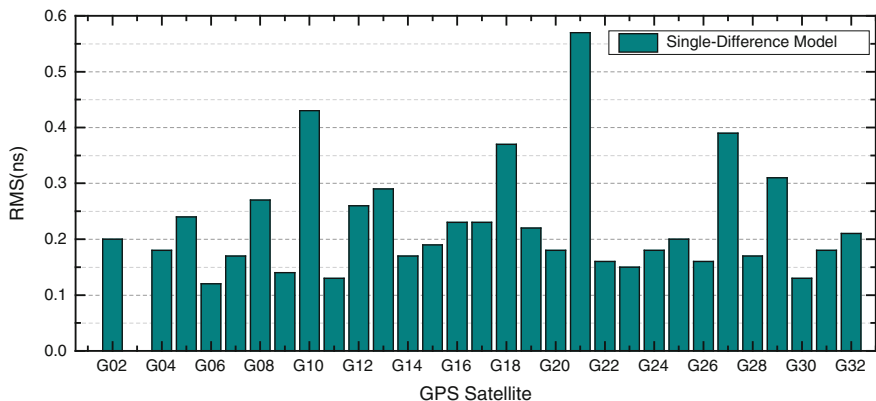


Fig. 5.6 Precise clock error accuracy estimated by epoch single-difference model using iGMAS test site

From the results, the clock error accuracy estimated by epoch single-difference model is about 0.2 ns using about 20 iGMAS test site. What the RMS of G10 and G21 is larger than other satellites is caused by less effective observation for the two satellites from station. So, reasonable station distribution and enough observation data are the premise conditions to obtain high-precision clock error.

5.5 Conclusions

From above discussion we will get conclusions as follow:

- (1) due to ambiguity exists in un-difference model estimation clock error, there is convergence process, after which the RMS accuracy comparing with IGS final clock products is about 0.2 ns; While, because of epoch single-difference model decoding continuous epoch ambiguity, there is no convergence process and RMS is better than 0.2 ns comparing with IGS; On the whole, the clock error accuracy estimated by the two models is about the same.
- (2) Combining the accuracy and estimation efficiency of the two models of clock error resolving, and considering the update requirements of real-time clock error at the same time, epoch single-difference model is more suitable for GNSS real-time clock error calculation because of less parameter and no exist convergence.
- (3) Using iGMAS site resources recently, the clock error accuracy estimated by epoch single-difference model is about 0.2 ns.

Because there are lots of troposphere parameters in epoch single-difference model, more research need to do in order to improve the estimation efficiency.

Acknowledgments Thanks to GNSS data and high-precision products provided by international GNSS Monitoring and Assessment System (iGMAS), and the support from China Satellite Navigation Conference 2012 youth outstanding papers found project (CSNC2012-QY-4).

References

1. Jiao WH, Ding Q, Li JW et al (2011) Monitoring and assessment of GNSS open services. *Sci Sin Phys Mech Astron* 41:521–527. doi:[10.1360/132011-359](https://doi.org/10.1360/132011-359) (in Chinese)
2. Jiao W (2014) International GNSS monitoring and assessment system and its new progress, In: The fifth China satellite navigation conference, Nanjing, China
3. International GNSS Service (IGS): <http://igsceb.jpl.nasa.gov>
4. Song W (2011) Research on real-time clock offset determination and real-time precise point positioning. Wuhan University
5. Hugentobler R, Rolf D, Fridez P (2004) Bernese GPS software version 5.0 draft. Berne: Astronomical Institute, University of Berne
6. Lou Y (2008) Research on real-time precise GPS orbit and clock offset determination. Wuhan University

Chapter 6

Integrating BDS and GPS to Accelerate Convergence and Initialization Time of Precise Point Positioning

Zongpeng Pan, Hongzhou Chai, Zehui Liu, Kefan Yang,
Yang Chong and Yangyin Xu

Abstract Due to the precision of BDS orbits and satellites clock, BDS PPP need long time to convergence to centimeter accuracy than GPS PPP. Combined BDS/GPS PPP can benefit from more visible satellites and enhanced satellite geometry distribution, which can accelerate the convergence speed of PPP. In this contribution, Combined BDS/GPS PPP and its integer ambiguity resolution were investigated. Using data set of 8 MGEX stations, the visible satellites and PDOP between GPS constellation and BDS/GPS constellations were analyzed. The results show combined BDS/GPS constellations can increase visible satellites and reduce PDOP. The improvement rate of visible satellites and PDOP is 89 and 31 %. Then, the positioning accuracy and convergence speed of single system PPP and combined BDS/GPS PPP were compared. For PPP hourly solutions, the positioning accuracy and convergence time of combined GPS/BDS PPP is better than single system PPP solutions. For PPP daily solutions, however, the positioning accuracy is not significantly improved. Finally, the integer ambiguity resolution of PPP was conducted. Only the GPS ambiguities were fixed and leaved BDS ambiguities as float values. The experiment results with MGEX BDS/GPS data indicate that combined GPS/BDS PPP with GPS integer ambiguity resolution can further improve positioning accuracy of PPP hourly solutions. Comparing with float solutions, the positioning accuracy is improved by 59.1 % in N component, 87.0 % in E component and 39.1 % in U component.

Keywords Precise point positioning · BDS · GPS · Convergence time · Ambiguity resolution

Z. Pan (✉) · H. Chai · K. Yang · Y. Chong · Y. Xu
Information Engineering University, Institute of Surveying and Mapping,
Zhengzhou 450001, China
e-mail: panzongpeng@yeah.net

Z. Liu
61175 Troops, Beijing, China

6.1 Introduction

The regional constellation of BeiDou Navigation Satellite System (BDS) has been officially operation since the end of 2012. It can provide service of navigation, timing and positioning in the Asia-Pacific Region. At the same time, IGS MGEX analysis centers and iGMAS analysis centers focus on BDS orbit and satellites clock determination [1–3]. The MGEX analysis centers such as WHU, GFZ and ESA can provide BDS precise orbit and satellites clock which lead to more researches focus on BDS precise positioning. The positioning accuracy of BDS PPP can reach a few centimeters [1–4], but more convergence time is needed and the positioning accuracy is less than GPS PPP [1, 2, 4, 5]. The reason is that BDS orbit and clock products have a lower accuracy than GPS products, due to the small network of monitor stations and the accuracy antenna model and solar radiation pressure model are unknown.

To improve performance of a GNSS mainly depends on increasing the number of navigation satellites, optimizing the spatial geometric configuration and improving of the accuracy of the observation signals [5]. At the same time, the fusion of multiple GNSSs can significantly increase the number of observed satellites and optimize the spatial geometry which can shorten convergence time of positioning. Li et al. [3] presented that the combined BDS/GPS PPP which based on un-differenced dual-frequency observations can shorten convergence time but it did not improve positioning results [4]. Zhao et al. [1] presented that when BDS and GPS data were combined to perform static PPP, the results were slightly worse compared to GPS only solutions. This could be caused by the multipath of some BDS GEO satellites and the process noise of intersystem biases for BDS. For kinematic PPP, however, the BDS/GPS combinations significantly improved the accuracy of GPS only PPP solutions. However, the above study focused on daily solutions [1]. The combined GPS/BDS positioning accuracy with short-time observations (such as 30–60 min) is unknown and PPP ambiguities are estimated as float values. Considering the PPP ambiguity resolution (PPP-AR) can improve positioning accuracy of short time.

With the development of PPP-AR in recent years, PPP ambiguity-fixed solution can improve positioning accuracy and shorten PPP convergence time [6–10]. But the convergence period to ambiguity-fixed solution is over several tens of minutes. The reason for this long period is that the slowly change geometry of satellites and other un-modeled error in observations lead to a long period before ambiguity resolution can be attempted [10, 11]. Cai and Gao [12] presented that combined GPS/GLONASS PPP can improve positioning reliable and reduce convergence time [12]. Jokinen et al. [13] and Li et al. [14] showed that combined GPS/GLONASS PPP can improve accuracy of GPS float ambiguity and reduce the time to first-fixed solution [13, 14]. Above studies focused on GPS/GLONASS PPP, however, combined GPS/BDS PPP was little studied. The impact of adding BDS observations to improve GPS PPP ambiguity resolution is worthy of investigation.

This contribution focuses on Integrating BDS and GPS to accelerate convergence and initialization time of PPP. Firstly the PDOP value and visible satellites are analysis between single system and combined BDS/GPS in the Asia-Pacific Region. Secondly, the PPP convergence time of different system combination is investigated and then the positioning accuracy of BDS, GPS and combined BDS/GPS PPP in hourly solutions is compared. Finally, GPS PPP ambiguity resolution was conduct and leaving BDS PPP ambiguities as real values, for the reason that BDS orbit and clock products have a lower accuracy than GPS products and the accuracy antenna model of BDS can't acquire. The impact of GPS PPP ambiguity resolution for combined BDS/GPS PPP hourly solutions is also investigated

6.2 Combined BDS/GPS PPP Model

6.2.1 Function Model

For combined BDS/GPS PPP, in order to eliminate the ionosphere effect, the ionosphere-free observations are used. Considering hardware delays in observations, for a satellite j of any GNSS s (GPS or BDS) observed by receiver r , the ionosphere-free pseudo range and carrier phase observations can be expressed as follows

$$\begin{aligned} P_{IF,r}^{s,j} &= \rho_r^{s,j} + c(dt_r^s - dt^{s,j}) + M_r^{s,j}d_{trop,r} + c(b_{IFP,r}^s - b_{IFP}^{s,j}) + \varepsilon(P_{IF,r}^{s,j}) \\ \Phi_{IF,r}^{s,j} &= \rho_r^{s,j} + c(dt_r^s - dt^{s,j}) + M_r^{s,j}d_{trop,r} + c(b_{IF\phi,r}^s - b_{IF\phi}^{s,j}) + \lambda_{IF}^s N_{IF,r}^{s,j} + \varepsilon(\Phi_{IF,r}^{s,j}) \end{aligned} \quad (6.1)$$

where $\Phi_{IF,r}^{s,j}$ and $P_{IF,r}^{s,j}$ are ionosphere-free carrier phase and pseudo range observations, the subscript IF for ionosphere-free combination, r for station number, superscript s for GNSS, j for a given satellite, $\rho_r^{s,j}$ as geometric distance between satellite and receiver, c for the speed of light in vacuum, dt_r^s as the receiver clock errors, $dt^{s,j}$ as the satellites clock errors, $b_{IFP,r}^s$, $b_{IF\phi,r}^s$, $b_{IFP}^{s,j}$ and $b_{IF\phi}^{s,j}$ as receiver-dependent or satellite-dependent pseudo range and phase hardware delay, $d_{trop,r}$ for tropospheric zenith delay (ZTD), $M_r^{s,j}$ for the mapping function, $N_{IF,r}^{s,j}$ as the ionosphere-free ambiguity, $\varepsilon(\Phi_{IF,r}^{s,j})$, $\varepsilon(P_{IF,r}^{s,j})$ for the carrier phase and pseudo range measurement noise and other errors, λ_{IF}^s is the wavelength of ionosphere-free combination.

Generally, the satellites clock errors can be corrected using IGS precise clock products $\tilde{dt}^{s,j} = dt^{s,j} + b_{IFP}^{s,j}$ (which include satellite-dependent hardware delay) and the receiver-dependent hardware delay is grouped into receiver clock, then the Eq. (6.1) can be expressed as

$$\begin{aligned}
P_{IF,r}^{s,j} &= \rho_r^{s,j} + c\tilde{d}t_r^s + M_r^{s,j}d_{trop,r} + \varepsilon\left(P_{IF,r}^{s,j}\right) \\
\Phi_{IF,r}^{s,j} &= \rho_r^{s,j} + c\tilde{d}t_r^s + M_r^{s,j}d_{trop,r} + \lambda_{IF}^s B_{IF,r}^{s,j} + \varepsilon\left(\Phi_{IF,r}^{s,j}\right)
\end{aligned} \tag{6.2}$$

where $\lambda_{IF} B_{IF,r}^{s,j} = c\left[\left(b_{IF\phi,r}^s - b_{IFP,r}^s\right) - \left(b_{IF\phi}^{s,j} - b_{IFP}^{s,j}\right)\right] + \lambda_{IF}^s N_{IF,r}^{s,j}$, as float ambiguity which include receiver and satellite-dependent hardware delay, $\tilde{d}t_r^s = dt_r^s + b_{IFP,r}^s$ as the comprehensive receiver clock errors and receiver-dependent hardware delay.

It should be noted that BDS precise orbit and satellites clock products provided by MGEX are based on the ITRF reference frame, which is the same as used in single GPS data processing. It doesn't need to transform coordinate reference frame between two systems. Due to different time system between BDS and GPS, there will have two receiver clock offsets in combined GPS/BDS PPP model, as follows

$$dt_r^G = t_r - t_{sys}^G \quad dt_r^C = t_r - t_{sys}^C \tag{6.3}$$

where dt_r^G , dt_r^C as GPS and BDS receiver clock offset, t_r is time of receiver clock, t_{sys}^G , t_{sys}^C as GPS and BDS system time. Instead of estimating BDS receiver clock offsets, it is preferable to introduce a system time difference parameter as it can reflect the difference between GPS and BDS system times [5, 12]. The BDS receiver clock offset can be expressed as

$$\begin{aligned}
dt_r^C &= t_r - t_{sys}^G - t_{sys}^C + t_{sys}^G \\
&= dt_r^G + dt_{sys}
\end{aligned} \tag{6.4}$$

where $dt_{sys} = t_{sys}^G - t_{sys}^C$ denotes system time difference between GPS and BDS. Considering receiver-dependent hardware delay, the BDS receiver clock offset can be expressed as

$$\begin{aligned}
\tilde{d}t_r^C &= \tilde{d}t_r^G + \tilde{d}t_{sys} \\
\tilde{d}t_{sys} &= dt_{sys} + b_{IFP,r}^C - b_{IFP,r}^G
\end{aligned} \tag{6.5}$$

Introduce Eq. (6.5) into Eq. (6.2) and linearize the equation, the Combined BDS/GPS PPP Model can be expressed as

$$\begin{aligned}
\Delta P_{IF,r}^{G,j} &= \boldsymbol{\mu}_r^{G,j} \Delta \mathbf{r} + c\tilde{d}t_r^G + M_r^{G,j}d_{trop,r} + \varepsilon\left(P_{IF,r}^{G,j}\right) \\
\Delta \Phi_{IF,r}^{G,j} &= \boldsymbol{\mu}_r^{G,j} \Delta \mathbf{r} + c\tilde{d}t_r^G + M_r^{G,j}d_{trop,r} + \lambda_{IF}^G B_{IF,r}^{G,j} + \varepsilon\left(\Phi_{IF,r}^{G,j}\right) \\
\Delta P_{IF,r}^{C,j} &= \boldsymbol{\mu}_r^{C,j} \Delta \mathbf{r} + c\left(\tilde{d}t_r^G + \tilde{d}t_{sys}\right) + M_r^{C,j}d_{trop,r} + \varepsilon\left(P_{IF,r}^{C,j}\right) \\
\Delta \Phi_{IF,r}^{C,j} &= \boldsymbol{\mu}_r^{C,j} \Delta \mathbf{r} + c\left(\tilde{d}t_r^G + \tilde{d}t_{sys}\right) + M_r^{C,j}d_{trop,r} + \lambda_{IF}^C B_{IF,r}^{C,j} + \varepsilon\left(\Phi_{IF,r}^{C,j}\right)
\end{aligned} \tag{6.6}$$

where $\mu_r^{G,j}$, $\mu_r^{C,j}$ denotes GPS and BDS unit direction vector, $\Delta \mathbf{r}$ denotes the three-dimensional coordinate correction. The parameters to be estimated in the equation are $\mathbf{X} = [\Delta \mathbf{r}, \tilde{d}t_r^G, \tilde{d}t_{sys}, d_{trop,r}, B_{IF,r}^{G,j}, B_{IF,r}^{C,j}]$. The unknown vector \mathbf{X} includes three coordinate parameters, a receiver clock offset, a system time difference parameter, a wet zenith tropospheric delay and real-value ambiguity parameters and the Extended Kalman Filter (EKF) can be utilized in parameters estimation.

6.2.2 Stochastic Model and Parameter Estimation Method

Both the fusion of function model and stochastic model in combined GPS/BDS PPP are important. When BDS and GPS data are combined to perform PPP, the initial weight ratio between GPS and BDS observations are set as 1:4, due to the GPS orbit and clock products have better quality compared with BDS ones. Whereas, the weight ratio between pseudo range and phase observations are set as 1:10,000 and the elevation-dependent weighting is also suggested in single GNSS observations.

Since the EKF is applied for combined PPP parameters estimation, appropriate stochastic models for parameters need to be provided. The parameter of system time difference can be modeled as random walk process [5, 12], due to its stable in short time period. The other parameters can be processed the same way as single system PPP and the error corrections such as windup, solid tide et al. must be considered. A detailed and clear discussion of parameters estimation and error corrections [15, 16] can refer to Kouba and Héroux [16].

6.2.3 Zero Difference Integer Ambiguity Resolution

In PPP data processing, the ionosphere-free ambiguities are usually estimated as real-values, due to the existence of the receiver and satellite-dependent phase hardware which also calls uncalibrated phase delays (UPD). As we known, double-difference ambiguities can be fixed, because the UPD are canceled. If the UPD can be canceled in PPP ambiguities, the PPP ambiguity fixing can be attempted. In general, the ionosphere-free ambiguities are decomposed into wide-lane (WL) and narrow-lane (NL) ones, as follows

$$\lambda_{IF} B_{IF,r}^j = \frac{cf_1}{f_1^2 - f_2^2} B_{1,r}^j - \frac{cf_2}{f_1^2 - f_2^2} B_{2,r}^j = \frac{f_2}{f_1 + f_2} \lambda_w B_{w,r}^j + \lambda_n B_{1,r}^j \quad (6.7)$$

where

$$\begin{aligned} B_{w,r}^j &= N_{w,r}^j + b_{w,r} - b_w^j \\ B_{1,r}^j &= N_{1,r}^j + b_{n,r} - b_n^j \end{aligned} \quad (6.8)$$

$B_{w,r}^j$ and $B_{1,r}^j$ as wide-lane and narrow-lane ambiguity which include receiver and satellite-dependent UPD. $N_{w,r}^j$ and $N_{1,r}^j$ denote the original WL and NL integer ambiguity. $b_{w,r}$, $b_{n,r}$, b_w^j and b_n^j denote receiver and satellite-dependent WL and NL UPD. Form above equation, we can see that the separation of UPD and integer ambiguity is key point in ambiguity fixing.

It is difficult to directly separate UPD and integer ambiguity, due to the linear relationship between UPD and integer ambiguity. However, the UPD have an integer part and a fractional part and the integer part will be grouped into integer ambiguity which doesn't lose its integer property.

$$\begin{aligned} B_{w,r}^j &= \tilde{N}_{w,r}^j + f_{w,r} - f_w^j \\ B_{1,r}^j &= \tilde{N}_{1,r}^j + f_{n,r} - f_n^j \end{aligned} \quad (6.9)$$

$\tilde{N}_{w,r}^j$ and $\tilde{N}_{1,r}^j$ denotes the sum of integer ambiguity and the integer part of WL and NL UPD, $f_{w,r}$, $f_{n,r}$, f_w^j and f_n^j denotes fractional part of receiver and satellite-dependent WL and NL UPD, which also call UPD for convenience. If fractional part of UPD can be resolved by server-end and provide to user, then the user can perform integer ambiguity fixing with single receiver.

Generally, the satellite-dependent WL UPD are stable over several days and can be estimated every day and applied to real-time PPP-AR with long update intervals [6, 7]. The fractional part of NL UPD contains not only UPD but also the bias in the estimated ambiguity, which are contaminated by inaccurate modeling of the observations. This results in the fluctuation of the NL UPD. Fortunately, fractional part of NL UPD is rather stable over a certain time span and can be estimated with short-term intervals, such as every 10–15 min. The approach proposed by Li and Zhang [14] can be applied to WL and NL UPD estimation [17] and integer ambiguity resolution in single receiver is presented in following section.

Since WL ambiguities have long wavelength, reaching 0.86 m, WL ambiguities can be easily and firstly fixed. The WL ambiguities can be calculated by taking the time average of the M-W combinations [18, 19] in order to reduce the effect of range noise and multipath.

$$\begin{aligned} \langle B_{w,r}^j \rangle &= \left\langle \left(\frac{f_1}{f_1 - f_2} \Phi_{1,r}^j - \frac{f_2}{f_1 - f_2} \Phi_{2,r}^j \right) - \left(\frac{f_1}{f_1 + f_2} P_{1,r}^j + \frac{f_2}{f_1 + f_2} P_{2,r}^j \right) \right\rangle \\ &= \langle \tilde{N}_{w,r}^j \rangle + f_{w,r} - f_w^j \end{aligned} \quad (6.10)$$

where $\langle * \rangle$ denotes function of taking the time average. After correcting satellite-dependent UPD, the corrected ZD ambiguities should have very similar fractional parts and we take the mean fractional parts of all the corrected ambiguities as receiver UPD. If satellite and receiver-dependent UPD are removed, WL ambiguities can be fixed by rounding to the nearest integer value. In order to ensure accuracy of fixed WL ambiguities, the fixing decision is made according to the

probability P_0 , which is calculated with the following formula [20] and the minimum probability is set as 0.999.

$$P_0 = 1 - \sum_{i=1}^{\infty} \left[\operatorname{erfc} \left(\frac{i - |b - n|}{\sqrt{2}\sigma} \right) - \operatorname{erfc} \left(\frac{i + |b - n|}{\sqrt{2}\sigma} \right) \right] \quad (6.11)$$

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt$$

where, b is real-valued ambiguity, σ is its STD, n is the nearest integer of b . If WL ambiguities are successfully fixed and introduce into ionosphere free ambiguities, then the corresponding NL ambiguities can be obtained by

$$\begin{aligned} B_{1,r}^j &= \frac{f_1 + f_2}{f_1} B_{IF,r}^j - \frac{f_2}{f_1 - f_2} \tilde{N}_{w,r}^j \\ &= \tilde{N}_{1,r}^j + f_{n,r} - f_n^j \end{aligned} \quad (6.12)$$

After correcting the satellite and receiver-dependent NL UPD, the integer property of NL ambiguities can be recovered. Due to the correlation between the PPP ambiguities, the LAMBDA method [21] is applied to solve the NL ambiguities. The criterion for the ratio test is set as 2.0.

If NL ambiguities are also successfully fixed, the ionosphere-free ambiguities can be recovered by

$$\tilde{B}_{IF,r}^j = \frac{f_2}{f_1 + f_2} \left(\tilde{N}_{1,r}^j + f_{n,r} - f_n^j \right) + \frac{f_1 f_2}{f_1^2 - f_2^2} \tilde{N}_{w,r}^j \quad (6.13)$$

Then the ambiguity-fixed solutions can be obtained by highly weighting the $\tilde{B}_{IF,r}^j$ in EKF. From Eq. (6.13), we find that the NL UPD is directly contributing to the ambiguity-fixed solutions. Hence, the accuracy of NL UPD is important to the positioning accuracy.

6.3 Experiments and Results Analysis

6.3.1 Data Collection

The experiments use observations of multi-GNSSs from 8 MGEX stations in March 11, 2014, DOY 70 and the data sampling rate is 30 s. The station distribution is showed in Fig. 6.1. The corresponding precise orbit and satellites clock products and the ‘ground truth’ are provided by GFZ, one of MGEX analysis centers. During data processing, only the dual-frequency GPS/BDS observations are used. At the

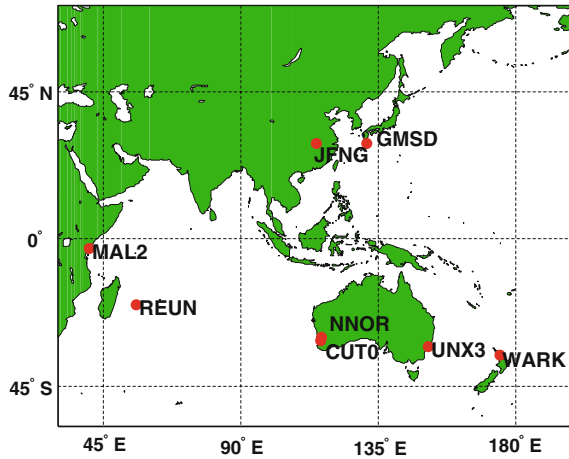


Fig. 6.1 Station distribution

same time, dual-frequency GPS observations of global distribution IGS stations are applied to estimate WL and NL UPD which proposed by Li and Zhang [17].

The PDOP value and visible satellites of GPS and combined GPS/BDS on 8 MGEX stations is analyzed. In Fig. 6.2, the left figure shows the average PDOP value of each station. The right figure shows the average visible satellites of each station. The elevation angle cutoff is set as 10°.

From Fig. 6.2, we can find that when the elevation angle cutoff is set as 10°, the average PDOP value is 2.03 and the number of average visible satellites is 9 for single GPS constellation. It also can be seen that the average PDOP value reduces to 1.41 and the number of average visible satellites increases to 16 when combined GPS/BDS constellations. The improvement rate of visible satellites and PDOP is 89 and 31 %. The above analysis shows combined BDS/GPS constellations can increase visible satellites and reduce PDOP.

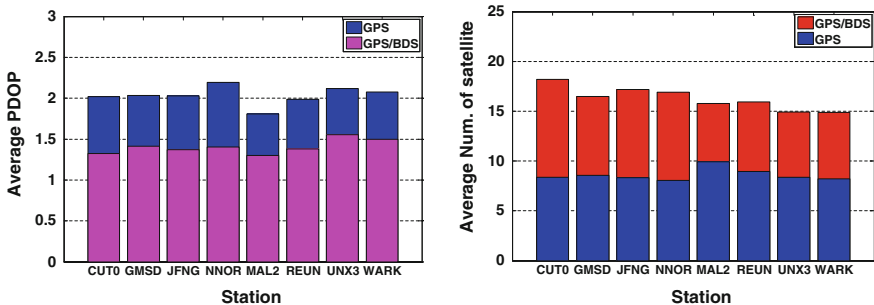


Fig. 6.2 The PDOP value and observed satellites of MGEX stations

6.3.2 Comparison of Convergence Time of Different Combined System PPP

In order to assess the performance of the combined GPS/BDS PPP, the convergence time, the positioning accuracy of hourly solutions and daily solutions are compared among BDS, GPS and combined GPS/BDS. When BDS and GPS data are combined to perform PPP, the initial weight ratio between GPS and BDS observations are set as 1:4.

Figures 6.3 and 6.4 show the convergence time series of BDS, GPS and combined GPS/BDS PPP on station GMSD and NNOR. The positioning results are compared with the ‘ground truth’ which provide by GFZ network solutions. The large fluctuations of time series are caused by simulated data breaks every 2 h on all observed satellites.

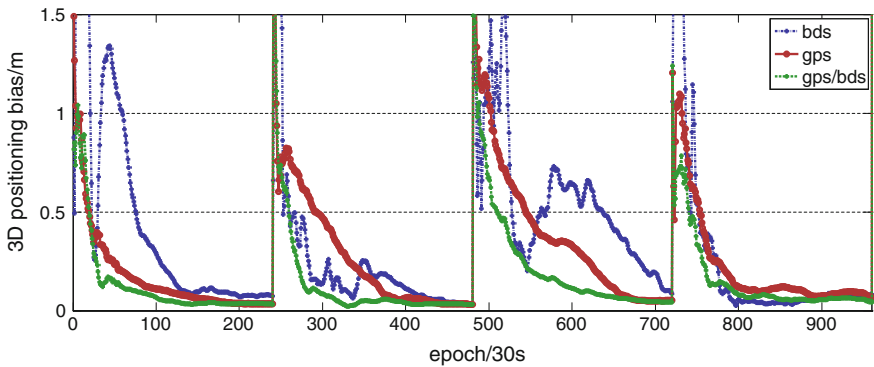


Fig. 6.3 Convergence time series of PPP with different system on GMSD station

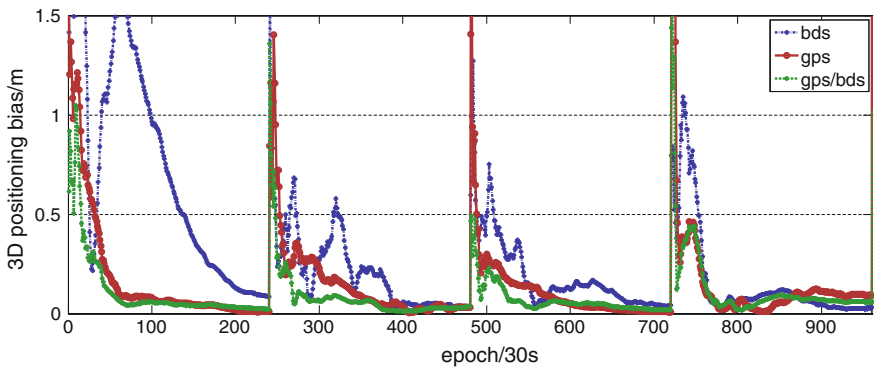


Fig. 6.4 Convergence time series of PPP with different system on NNOR station

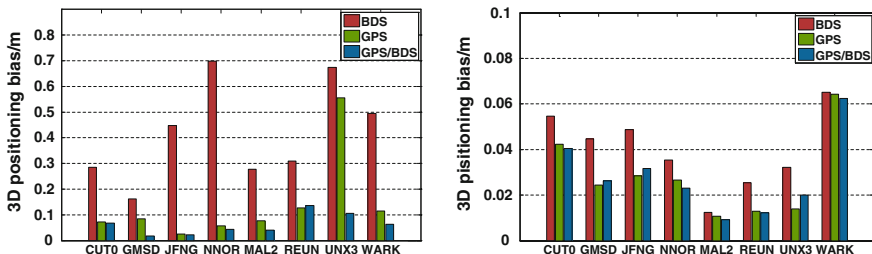


Fig. 6.5 PPP hourly solutions (*left*) and daily solutions (*right*) of 8 MGEX stations with different system

As we can see from Figs. 6.3 and 6.4, the combined GPS/BDS PPP can accelerate convergence compares with GPS PPP and obviously faster than BDS PPP. It also can see that the convergence time of BDS PPP is longer than GPS PPP solutions. The main reason for this phenomenon is that the GPS orbit and satellites clock products have better quality compared with BDS ones.

Figure 6.5 shows the accuracy of PPP hourly solutions and daily solutions of 8 MGEX stations can also indicate this phenomenon. As can be seen from left figure, the positioning accuracy of GPS PPP and combined GPS/BDS PPP hourly solutions are generally better than 10 cm in 3D component and the accuracy of GPS/BDS PPP is better than GPS PPP. It also can be seen, the 3D positioning biases of BDS PPP hourly solutions are reaching a few decimeters. Therefore, the convergence time of BDS PPP is longer than combined GPS/BDS ones. As shown in right figure, however, the daily solutions of BDS PPP are comparative to GPS and combined GPS/BDS PPP. The positioning accuracy of BDS PPP daily solutions is within a few centimeters and slightly worse than GPS one. Whereas, the positioning accuracy of daily solutions of combined GPS/BDS PPP and GPS PPP is at the same level.

In order to assess the performance of convergence and positioning accuracy of combined GPS/BDS PPP, Fig. 6.6 shows the positioning results of PPP in different period of daily data set on station CUTO, JFNG and MAL2. It can be seen, when the data set is less than 2 h, the performance of convergence and positioning

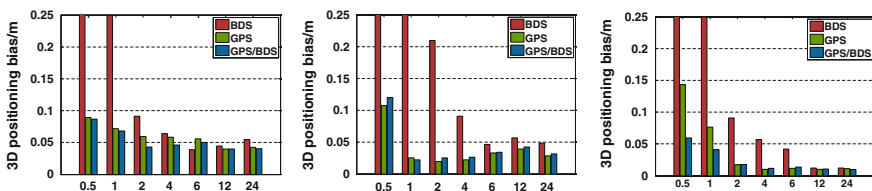


Fig. 6.6 The positioning results of PPP in different period of daily data set on station CUTO, JFNG and MAL2

accuracy of combined GPS/BDS PPP are better than single system PPP solutions. When the data set is more than 2 h, the positioning accuracy of combined GPS/BDS PPP is at the same level and slightly better than BDS PPP solutions.

6.3.3 Comparison Between PPP Ambiguity-Fixed Solutions and Ambiguity-Float Solutions

The fusion of GPS and BDS to PPP can significantly increase the number of observed satellites, optimize the spatial geometry which can shorten convergence time and improve positioning accuracy with short-time observations. However, the improved performance of positioning accuracy of combined GPS/BDS PPP is not obvious. In order to further improve positioning accuracy with short-time observations, GPS PPP ambiguity fixing is attempted, leaving BDS PPP ambiguities as real values, for the reason that BDS orbit and clock products have a lower accuracy than GPS products and the accuracy antenna model of BDS can't acquire. When attempting PPP ambiguity fixing, the WL ambiguity can be calculated by taking the time average of the M-W combinations and fixed to the nearest integer value. WL ambiguity fixing decision is made if the probability P_0 is large than 0.999. The NL ambiguity fixing is attempted by the LAMBDA method and the criterion for the ratio test is set as 2.0. The procedure of PPP ambiguity fixing can reference to Sect. 1.2.3.

Figure 6.7 shows the positioning accuracy of GPS/BDS PPP ambiguities fixed and float solutions in three component and 3D component with hourly observations.

As can be seen from left sub-figure, when PPP ambiguity is fixed, the positioning accuracy in NEU component is improved in different degree. It can be seen more clearly in right sub-figure that the positioning accuracy is within 5 cm of combined GPS/BDS PPP with GPS ambiguity fixing and the performance is better than GPS/BDS ambiguities float solutions.

Table 6.1 statistics the positioning RMS of ambiguities fixed and floats solutions in three component and 3D component with hourly observations and its improvement rate.

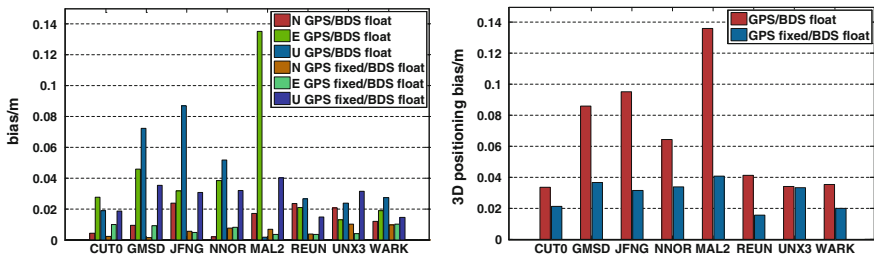


Fig. 6.7 The hourly solutions of combined GPS/BDS PPP with ambiguities fixed or as float values

Table 6.1 The RMS of PPP ambiguities fixed and floats solutions with hourly observations and its improvement rate

	GPS/BDS float (m)	GPS fixed/BDS float (m)	Improvement (%)
N	0.0161	0.0066	59.1
E	0.0553	0.0072	87.0
U	0.0471	0.0287	39.1
3D	0.0744	0.0303	59.3

The result shows, the RMS of ambiguities fixed solutions are better than 1 cm in horizontal components and 3 cm in the vertical. Comparing with float solutions, the positioning accuracy is improved by 59.1 % in N component, 87.0 % in E component and 39.1 % in U component. The improvement in E component is obvious.

6.4 Conclusion

Due to BDS orbit and satellites clock products have a lower accuracy than GPS products, BDS PPP need long time to convergence to centimeter-level and the positioning accuracy is less than GPS PPP solutions. Combined BDS/GPS PPP can benefit from more visible satellites and enhanced satellite geometry distribution which would accelerate the convergence speed of PPP. This contribution focuses on integrating BDS and GPS to accelerate convergence and initialization time of PPP. The experiment results with data set of 8 MGEX stations show

- (1) When the elevation angle cutoff is set as 10° , the average PDOP value is 2.03 and the number of average visible satellites is 9 for GPS only. However, the average PDOP value reduces to 1.41 and the number of average visible satellites increases to 16 when combined GPS/BDS constellations. The improvement rate of visible satellites and PDOP is 89 % and 31 %. Combined BDS/GPS constellations can increase visible satellites and reduce PDOP.
- (2) Combined GPS/BDS PPP can accelerate convergence compares with GPS PPP and obviously faster than BDS PPP. For PPP hourly solutions, the positioning accuracy of GPS PPP and combined GPS/BDS PPP hourly solutions are generally better than 10 cm in 3D component and BDS PPP solutions can only reach a few decimeters. For PPP daily solutions, the positioning accuracy of BDS PPP are within a few centimeters and slightly worse than GPS PPP solutions. The positioning accuracy of combined GPS/BDS PPP and GPS PPP is at the same level.
- (3) Further improvement of positioning accuracy with short-time observations can obtain when GPS PPP ambiguities are fixed. The RMS of ambiguities fixed solutions is better than 1 cm in horizontal components and 3 cm in the vertical. Comparing with RMS of float solutions, the positioning accuracy is improved by 59.1 % in N component, 87.0 % in E component and 39.1 % in U component. The improvement in E component is obvious.

In summary, the fusion of GPS and BDS to PPP can significantly increase the number of observed satellites, optimize the spatial geometry and shorten convergence time of positioning. At the same time, GPS PPP ambiguity fixing can further improve positioning accuracy with short-time observations.

Acknowledgments Thanks to the IGS MGEX and GFZ for providing Multi-GNSS observations and ephemeris products. This work was supported by National Natural Science Foundation of China (41274045).

References

1. Zhao Q, Guo J, Li M, Qu L, Hu Z, Shi C et al (2013) Initial results of precise orbit and clock determination for COMPASS navigation satellite system. *J Geodesy* 87(5):475–486
2. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Teunissen P, Nakamura S et al (2013) Initial assessment of the COMPASS/BeiDou-2 regional navigation satellite system. *GPS Solut* 17(2):211–222
3. Li X, Ge M, Zhang H, Nischan T, Wickert J et al (2013) The GFZ real-time GNSS precise positioning service system and its adaption for COMPASS. *Adv Space Res* 51(6):1008–1018
4. Li W, Teunissen P, Zhang B, Verhagen S (2013) Precise point positioning using GPS and compass observations. In: China satellite navigation conference (CSNC) 2013 proceedings. Springer, Berlin, vol 244, pp 367–378
5. Li M, Qu L, Zhao Q, Guo J, Su X, Li X (2014) Precise point positioning with the BeiDou navigation Satellite system. *Sensors* 1:927–943
6. Ge M, Gendt G, Rothacher M, Shi C, Liu J (2008) Resolution of GPS carrier-phase ambiguities in Precise point positioning (PPP) with daily observations. *J Geodesy* 82 (7):389–399
7. Laurichesse D, Mercier F, Berthias JP, Broca P, Cerri L (2009) Integer ambiguity resolution on undifferenced GPS phase measurements and its application to PPP and satellite precise orbit determination. *Navigation* 56(2):135–149
8. Collins P et al (2010) Undifferenced GPS ambiguity resolution using the decoupled clock model and ambiguity datum fixing. *Navigation* 57(2):123–135
9. Geng J, Teferle FN, Shi C, Meng X, Dodson AH, Liu J (2009) Ambiguity resolution in precise point positioning with hourly data. *GPS Solut* 13(4):263–270
10. Zhang X, Li P, Guo F (2013) Ambiguity resolution in precise point positioning with hourly data for global single receiver. *Adv Space Res* 51(1):153–161
11. Geng J, Meng X, Dodson A, Ge M, Teferle F (2010) Rapid re-convergences to ambiguity-fixed solutions in precise point positioning. *J Geodesy* 84(12):705–714
12. Cai C, Gao Y (2013) Modeling and assessment of combined GPS/GLONASS precise point positioning. *GPS Solut* 17(2):223–236
13. Jokinen A et al (2012) Improving fixed-ambiguity precise point positioning (PPP) convergence time and accuracy by using GLONASS. In: ION GNSS 2012, Nashville, TN
14. Li P, Zhang X (2013) Integrating GPS and GLONASS to accelerate convergence and initialization times of precise point positioning. *GPS Solut* 18(3):461–471
15. Zumberge JF et al (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. *J Geophys Res* 102(B3):5005–5017
16. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock product. *GPS Solut* 5(2):12–28
17. Li X, Zhang X (2012) Improving the estimation of uncalibrated fractional phase offsets for PPP ambiguity resolution. *J Navig* 65:513–529

18. Melbourne WG (1985) The case for ranging in GPS-based geodetic systems. In: Proceedings first international symposium on precise positioning with the global positioning system, Rockville, pp 373–386, 15–19 April 1985
19. Wübbena G (1985) Software developments for geodetic positioning with GPS using TI-4100 code and carrier measurements. In: Proceedings of first international symposium on precise positioning with the global positioning system, Rockville, pp 403–412, 15–19 April 1985
20. Dong DN, Bock Y (1989) Global positioning system network analysis with phase ambiguity resolution applied to crustal deformation studies in California. *J Geophys Res: Solid Earth* 94 (B4):3949–3966
21. Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *J Geod* 70(1–2):65–82

Chapter 7

Study on BeiDou Navigation Satellite Precise Orbit Determination Based on the Extended Kalman Filtering

Yan Wang, Chuanding Zhang and Lijie Song

Abstract In order to the study of BeiDou navigation satellite system (BDS) precise orbit determination, a suit of satellite precise orbit determination software named Orbit Study is developed under VC6.0 platform. Object-oriented programming ideas are used. The implementation of each module is carried out by the corresponding c++ class. This software currently processes undifferenced phase and pseudorange measurements. The Extended Kalman Filter (EKF) is used to estimate parameters. It can satisfy the need of post or real time orbit determination. Firstly, the basic modules of Orbit Study are introduced. The data management mode and the optimization algorithm of improving the computational efficiency are also been introduced in the first part. Secondly, the observation model, dynamic model and EKF orbit determination process are introduced. Thirdly, BDS orbit determination experiment is processed based on the observation of Compass Experimental Test Service network (CETS). Two types for comparison for BDS orbit determination accuracy are used, which are inner and outer consistency test. Although the results show that the accuracy of Orbit Study is 10 cm worse than PANDA in radial direction, it has orbit determination ability of BDS preliminary. At last, the reasons of this result are analyzed and the next plans are proposed.

Keywords Beidou navigation satellite system · Precise orbit determination · Extended kalman filter

Y. Wang (✉) · C. Zhang · L. Song
Institute of Geospatial Information, Information Engineering University, No. 62, Kexue
Road, 450001 Zhengzhou, China
e-mail: wang1yan.hi@163.com

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015
Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_7

7.1 Introduction

BeiDou navigation satellite system (BDS) is a global navigation system, which is independently developed, deployed, and operated by China. The construction of BDS implements the three-step strategy. The second step that operation of the network has completed and it has already provided navigation and positioning services for China and its surrounding. The third step of network operation work is under construction. The final BDS is consisted of 5 GEO, 3 IGSO, and 27MEO satellites [1, 2]. The key to enhancing satellite navigation system's PNT performance lies primarily in two aspects: on the one hand, increasing the number of navigation satellites and optimizing spatial geometric configuration, that is, improving satellites' geometric observation condition at the ground terminal; on the other hand, improving the accuracy of satellite ephemeris, clock offset and relevant model parameters, which means improving accuracy of the system's spatial signals [3, 4]. The development of GPS tells us that satellite precise orbit determination ability is the key part in the process of construction of the whole system. A series of precise orbit determination experiment are processed by many scholars [4–6].

The International GNSS Service (IGS) plays an important role in the accuracy improvement of the GPS orbit. A few analysis centers of IGS have had distinctive precision positioning and orbit determination software [7]. If China wants to improve the orbit accuracy of BDS, a few analysis centers are also need to be established for the study of BDS satellite precise orbit determination for a long time. In China, there are several research institutes. For example, Shanghai Astronomical Observatory Chinese Academy of Sciences, Wuhan University, Institute of Geodesy and Geophysics Chinese Academy of Sciences, China Xi'an Satellite Control Center, Xi'an Research Institute of Surveying and Mapping, Nanjing University, et al. The PANDA of Wuhan University has implemented a series of studies on BDS satellite orbit determination and Compass Experimental Test Service network has established [8, 9]. The SHORDE of Shanghai Astronomical Observatory Chinese Academy of Sciences has used to BDS daily operation control system. The SPODS of Xi'an Research Institute of Surveying and Mapping already has the ability of GNSS positioning and orbit determination [10].

To carry out BDS precise orbit determination research, we must have our own analysis software. So a set of orbit determination software named Orbit Study is developed, which has the ability of precise positioning and orbit determination. This software currently processes undifferenced phase and pseudorange measurements. The Extended Kalman Filter (EKF) is used to estimate parameters. It can satisfy the need of post or real time orbit determination. Firstly, the basic modules of Orbit Study are introduced. The data management mode and the optimization algorithm of improving the computational efficiency are also been introduced in the first part. Secondly, the measurement model, dynamic model and EKF orbit determination process are introduced. Thirdly, BDS orbit determination experiment is processed based on the observation of Compass Experimental Test Service network (CETS). The accuracy of the orbit is assessed by two ways: On the one

hand, overlap arc comparison is undertaken. On the other hand, compare with PANDA's result. The result shows that the results of Orbit Study haven't get the level of PANDA. But it has already had orbit determination ability of BDS preliminary. At last, the reasons of this result are analyzed and the next plans are proposed.

7.2 Software Overview

Orbit Study is developed under VC6.0 platform. Object-oriented programming ideas are used. It mainly includes the following several modules: The first part is data preparation and parameter Settings. The second part is data preprocessing. The third part is orbit integrator, which is carried by Runge-Kutta and Adams-Moulton. The fourth part is EKF filter. The fifth part is observation equation establishing and error correction. The last part is quality control. Every functional module is implemented by the corresponding data processing class. This structure makes the data processing clearly and not easy to have errors.

7.2.1 Data Preparation and Parameter Settings

The job of data preparation is that downloading the corresponding data from IGS data processing center, which includes observation data (RINXE format), broadcast ephemeris, precise ephemeris, clock correction and updating the table files. The table files include EOP parameter, station coordinates and phase center offset data, et al.

The function of the parameter setting module is mainly to set the parameter of data processing, which include the estimation time interval, the path of the data, computation option, the condition of cycle slip detection and gross error rejection and the parameters' prior information and process noise, et al. This part can modify the configuration file directly and can also be written from the software interface.

7.2.2 Data Management Mode

Figure 7.1 shows the construction of data management. The struct of total observation data is shown in Fig. 7.1. The total observation data consists of $n*m$ structure of objects named CStationOBS, where n is the number of station and m is the number of epoch. Each CStationOBS object data consists of t OBSVEC structure of objects, where t is the observation number of a station in one epoch.

The process that the RINEX format observation data transform to the construction of Orbit Study is as follows:

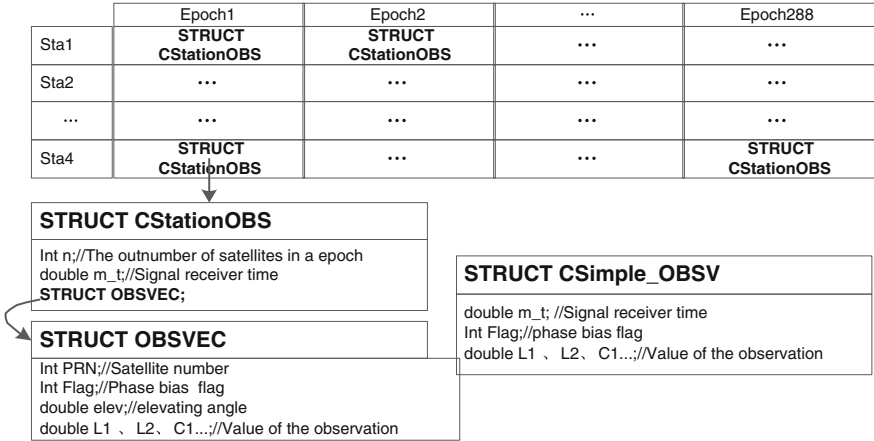


Fig. 7.1 Construction of observation data in Orbit Study

1. The observation data of a station is processed. An IGS station named ASPA is chosen to illustrate, whose sampling interval is 30 s. The header of observation data of ASPA is read and put into to a structure object named OBS_Header. After that, the body of observation data of ASPA is read by the epoch and put the data into the CStationOBS.
2. Cycle slip detection and gross errors rejection are processed by the satellite number order. The observation data of one satellite is extracted from CStationOBS and put into CSimple_OBSV. So it is convenient for a satellite to cycle slip detection and gross error rejection by the order of epoch. The method of cycle slip detection is TurboEdit [11]. The symbol ‘Flag’ in STUCT OBSVEC is the ambiguity flag, where 0 symbolize this observation cannot be used, 1 symbolize this observation has a new ambiguity, 2 symbolize this observation has the same ambiguity with the last observation.
3. The ambiguity flag after step 2 will be put into the CStationOBS in the step 1. In this step, the Flag of each observation will be the number of ambiguity instead. If the observation cannot be used, the ‘Flag’ will be 0.
4. The next station will be processed until the last station. The clean observation data will be wrote into a binary file. To the parameter estimation, the total observation can be read directly from the binary file.

7.2.3 Orbit Integrator

The implement of orbit integrator in Orbit Study is as follows: Starting with 8 orders Runge-Kutta. When the order of right function of Adams integrator is met, Adams integrator is used. In order to improve the accuracy of integrator, the

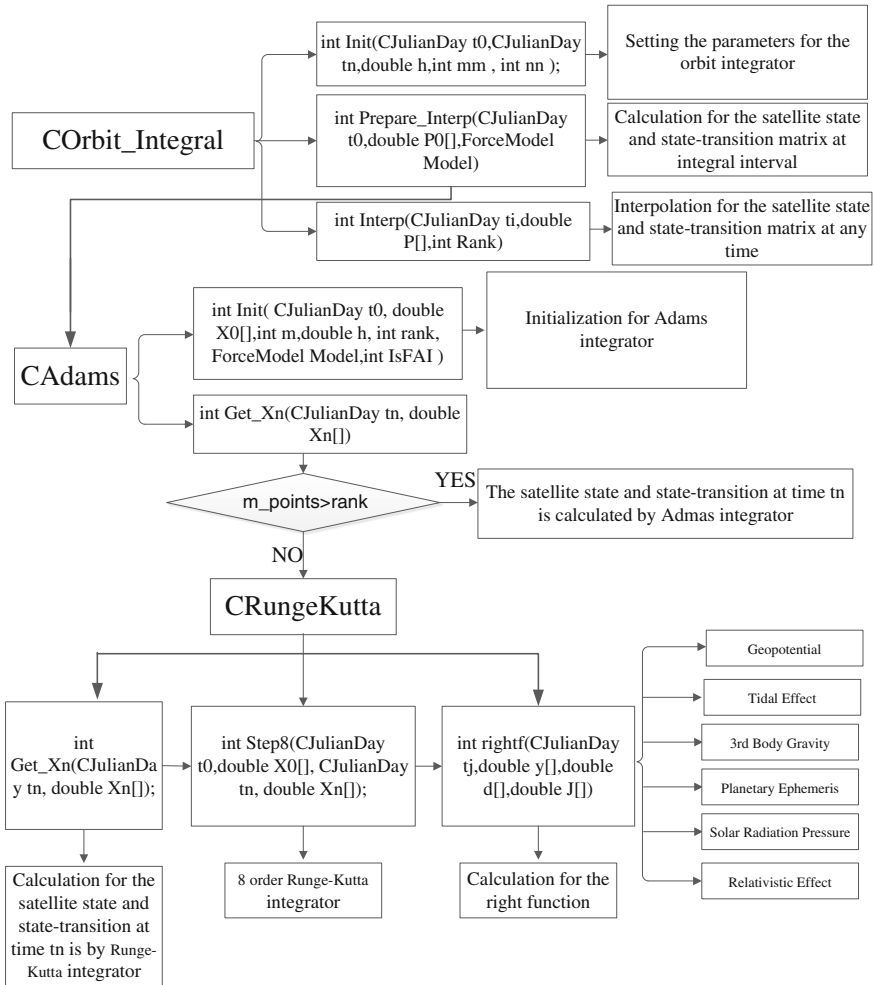


Fig. 7.2 Class of orbit integrator

Adams-Moulton algorithm is used. There are three orbit integrator classes in Orbit Study, which are CORbit_Integral, CAdams and CRungeKutta. Figure 7.2 shows the usage of the integrator.

7.2.4 EKF Parameter Estimation R

For parameter estimation, Orbit Study employs Extended Kalman Filter expressed in Eq. (7.1) as the observation update and Eq. (7.2) as the temporal update.

In addition to the conventional forward filter, the backward filter and the smoother are supported by Orbit Study.

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_k(-)\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_k(-)\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \\ \hat{\mathbf{x}}_k(+) &= \hat{\mathbf{x}}_k(-) + \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_k(-))) \\ \mathbf{P}_k(+) &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k(-) \end{aligned} \quad (7.1)$$

$$\begin{aligned} \hat{\mathbf{x}}_k(-) &= \hat{\mathbf{x}}_k(+) + \int_{t_k}^{t_{k+1}} \mathbf{f}(\hat{\mathbf{x}}_k(+), \tau) d\tau \\ \mathbf{P}_{k+1}(-) &= \mathbf{\Phi}(t_{k+1}, t_k)\mathbf{P}_k(+)\mathbf{\Phi}(t_{k+1}, t_k)^T + \mathbf{Q}_k \end{aligned} \quad (7.2)$$

For the BDS orbit determination, the estimated state vector at t_k consisting of the satellite position and velocity, satellite clock, tropospheric parameters, SRP parameters, and phase bias. Table 7.1 shows the estimated parameters per an epoch if 18 stations' measurements are used to determine the 14 satellites' orbit.

The problem that large-scale matrix multiplication is always met in the EKF parameter estimation. In Eq. (7.1), the calculation of gain matrices \mathbf{K}_k is the most time consuming part. The calculation of the two matrices that $\mathbf{H}_k\mathbf{P}_k(-)\mathbf{H}_k^T$ and $\mathbf{P}_k(-)\mathbf{H}_k^T\mathbf{Q}_k$ is the most time consuming part. The special algorithms are designed in Orbit Study:

1. Because of $\mathbf{P}_k(-)$ is a symmetric matrix, the lower triangular elements are only stored in Orbit Study. The elements of this matrix are stored into a "one dimensional array". To the matrix multiplication, the corresponding elements in the "one dimensional array" are determined by the line and column number in the storage location of $\mathbf{P}_k(-)$. This algorithm save memory space and improve the computational efficiency.

Table 7.1 Estimated parameters per an epoch

Parameters	Number of states	Sats/Stas	Total
Satellite position/velocity	6	14	84
Satellite SRP parameters	9	14	126
Satellite clock	1	14	14
Station receiver clock	1	18	18
Station Tropos. ZTD	1	18	18
Station Tropos. gradient	5	18	90
Carrier phase bias	1	18*14	252
Total parameters per Epoch	1490		

2. As the Table 7.1 shows, the number of the total parameters per epoch is 1490. The $P_k(-)$ in Eq. (7.1) is a matrix that 1490 multiply 1490. When to the temporal update, the corresponding coefficient of many ambiguity parameters is 0. The corresponding elements in $P_k(-)$ are not changed after the temporal update. In order to improve the computational efficiency, Orbit Study takes a special algorithm. Extract the corresponding elements of the ambiguity parameters in this epoch and put into a new matrix $P_k(-)_{ep}$. The dimension of $P_k(-)_{ep}$ is lower than $P_k(-)$. After the temporal update, the elements of $P_k(-)_{ep}$ will be put back to $P_k(-)$. This algorithm improves the computational efficiency and not affects the accuracy.
3. To matrix H_k multiply matrix $P_k(-)$ in Eq. (7.1), it will cost many time if H_k multiply $P_k(-)$ directly. Because H_k is matrix that have n lines, t columns and many elements in it is 0, where n is the observation number in this epoch and t is the number of total parameters. If H_k multiply $P_k(-)$ directly, it is inefficient. So a special algorithm is used in Orbit Study. The non-zero elements in H_k are only stored, which in a “one dimensional array”. These non-zero elements of one observation are constant. They are the corresponding coefficients of parameters of dynamic, troposphere, clocks and phase bias. A index array named $H_k\text{-Index}$ is used to record the location of the elements of H_k . So the computational efficiency is improved.

7.3 Algorithms and Models

7.3.1 Measurement Model

The basic observation for BDS orbit determination is the undifferenced ionosphere-free carrier phase and pseudorange measurements, shown in Eq. (7.3):

$$\begin{aligned}
 \Phi_{LC} &= C_1\lambda_1\Phi_{B1} + C_2\lambda_2\Phi_{B2} = \rho + c(dt - dT) + T + N_{LC} \\
 &\quad - \Delta_{pcvs} - \Delta_{pcvr} + \Delta_{rel} + \Delta_{phw} + \varepsilon_{\Phi_{LC}} \\
 P_{LC} &= C_1P_{B1} + C_2P_{B2} = \rho + c(dt - dT) + T \\
 &\quad - \Delta_{pcvs} - \Delta_{pcvr} + \Delta_{rel} + \Delta_{phw} + \varepsilon_{P_{LC}} \\
 C_1 &= f_1^2 / (f_1^2 - f_2^2) \\
 C_2 &= f_2^2 / (f_1^2 - f_2^2)
 \end{aligned} \tag{7.3}$$

$\Phi_{B1}, \Phi_{B2}, P_{B1}, P_{B2}$	B1/B2 carrier phase measurement (cycle) and pseudorange measurement (m)
ρ	Satellite-station geometric distance (m)
dt, dT	receiver/satellite clock bias (sec)

Table 7.2 Reference frames and interconnection

ECEF	IGS08
ECI	J2000
Precession/Nutation	IAU 2000A/B
ERP parameters	IERS Bulletin B, EOP C04, IGS ERP

N_{LC}	carrier phase bias of ion-free LC (m)
$\Delta_{pcvs}, \Delta_{pcvr}$	satellite/receiver antenna PCV (m)
Δ_{rel}	relativity effect correction (m)
Δ_{phw}	phase-windup effect correction (m)
$\varepsilon_{\Phi_{LC}}, \varepsilon_{P_{LC}}$	measurement noise of ion-free LC and PC (m)
f_1, f_2	B1/B2 carrier frequency (Hz), B1 = 1561.098 MHz, B2 = 1207.140 MHz

The SOFA software package is use to the transformation between inertial coordinate (ECI) and earth-fixed (ECEF) [12]. Table 7.2 shows the details. Table 7.3 shows the parameters and priori variance-covariance of parameters.

Table 7.3 Reference frames and interconnection

Item	Models	Priori variance-covariance
Observation	Undifferenced ionosphere-free carrier phase and pseudorange measurements	Phase: 0.02 cycle Pseudorange 1.0 m
Estimation time span/interval	3 days/300 s	
Dynamic parameters	Satellite position/velocity BRNESE 9 SPR Parameters Prior value computed from broadcast ephemeris orbit smoothing	$10 \times 10 \times 10$ m, 10^{-4} , 10^{-4} , 10^{-4} m/s 9 SPR: 0.1 times of prior value
Station Tropos. ZTD	Saastamoinen Process noise	0.3 m + 0.001 m/s
Station Tropos. Gradient	2 parameters Process noise	0.02 m + 0.0005 m/s
Carrier phase bias	Folat Prior value computed from pseudorange	10 m + 0 m/s
Satellite clock	White noise	100 m
Receive clock	White noise Prior value computed by pseudorange and broadcast ephemeris	1000 m
Station coordinates	Fixed	

Table 7.4 Perturbative force models

Geopotential	EGM96 up to 10 order
Tidal effect	Solid earth, pole tide, ocean loading (IERS 2010)
3rd body gravity	Moon and Sun as point of mass
Planetary Ephemeris	JPL DE405
Eclipse model	Cylidric or penumbra/umbra by earth and moon shadow
Solar radiation pressure model	BERNESE ECOM model
Relativistic effect	IERS 2010

7.3.2 Dynamic Model

Table 7.4 shows the dynamic models in Orbit Study.

7.4 Evaluation of Accuracy

To evaluate the accuracy of the BDS satellite orbit determination of Orbit Study, using world-wide 18 CETS station observation data. Figure 7.3 shows the geometry of the station network used for the evaluation. The estimation time span is 2012/11/05-2012/11/13 (DOY 310-318). During this time, the working satellites are C01,

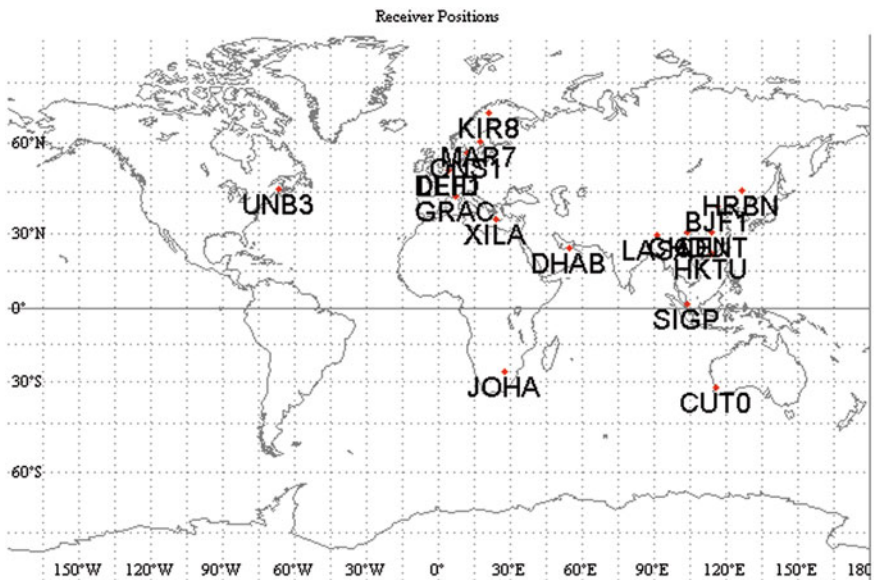


Fig. 7.3 Station Network used for evaluation

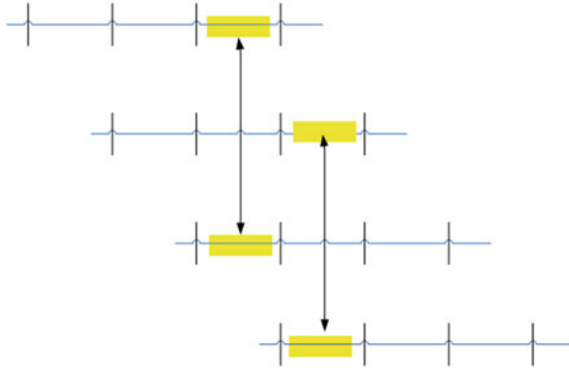


Fig. 7.4 Differences of overlapping arc

C04, C05, C06, C07, C08, C09, C10, C11, and C12. To compare, the orbit determination results of PANDA are also be collected. Precise orbits of BDS satellites are determined with every three successive days as an arc segment. The two arc segments that PANDA computes are collected, which are 2012/11/05-2012/11/08 and 2012/11/06-2012/11/09.

The differences of overlapping arc are used to evaluate the orbit accuracy, which is shown in Fig. 7.4.

Figure 7.5 and Table 7.5 show overlapping accuracy in the radial direction for each of the satellites. This result is given by six arc segments (DOY 310–318). The results show that the average accuracy in radial is about 20 cm.

Figure 7.6 and Table 7.6 show the estimated satellite orbit accuracy with respect to the result of PANDA. The two successive arc segments are compared, which are DOY 310-312 and DOY 311-313.

To make it more clearly, the overlapping accuracy of DOY 310-313 and DOY 311-314 is compared by Orbit Study and PANDA respectively. In Fig. 7.7, the top

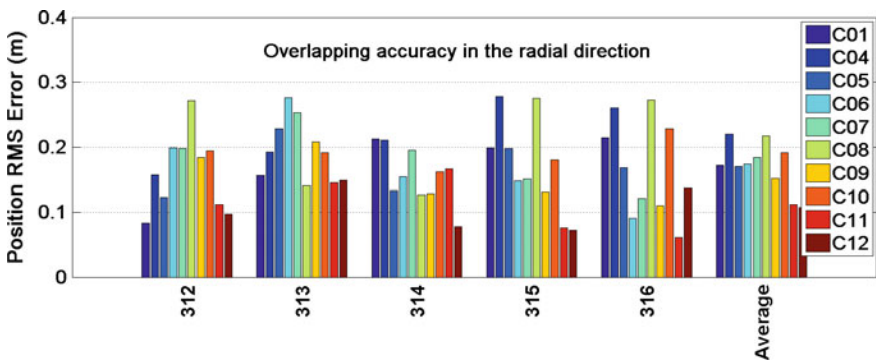


Fig. 7.5 Overlapping accuracy in the radial direction

Table 7.5 Overlapping accuracy in the radial direction (m)

	2013/312	2013/313	2013/314	2013/315	2013/316	Average
C01	0.083	0.156	0.212	0.199	0.214	0.172
C04	0.157	0.192	0.211	0.278	0.260	0.220
C05	0.123	0.228	0.133	0.198	0.168	0.170
C06	0.199	0.276	0.154	0.148	0.091	0.174
C07	0.198	0.253	0.195	0.151	0.121	0.184
C08	0.271	0.141	0.127	0.275	0.272	0.217
C09	0.184	0.208	0.129	0.131	0.110	0.152
C10	0.194	0.191	0.162	0.180	0.228	0.191
C11	0.112	0.145	0.166	0.076	0.061	0.112
C12	0.097	0.149	0.078	0.072	0.137	0.107

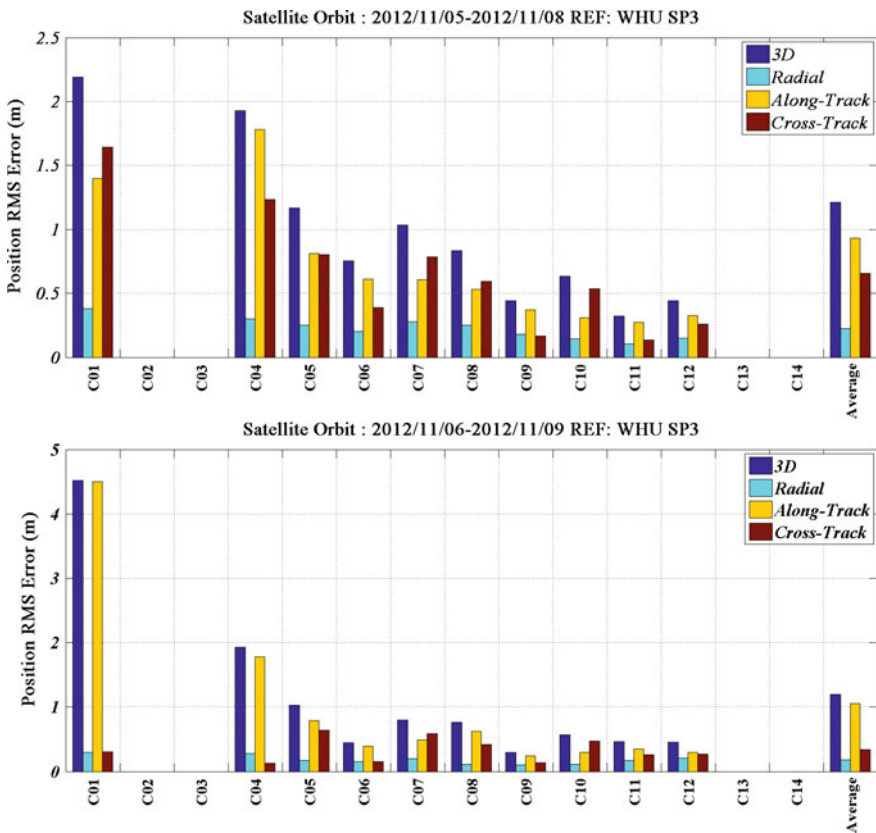


Fig. 7.6 Estimated satellite orbit accuracy with respect to PANDA

Table 7.6 Overlapping accuracy in the radial direction (m)

		3D	Radial	Along	Cross
GEO	310–313	1.565	0.245	1.357	0.592
	311–314	2.494	0.249	2.357	0.358
IGSO	310–313	0.593	0.158	0.384	0.395
	311–314	0.574	0.134	0.410	0.355
MEO	310–313	0.361	0.081	0.292	0.197
	311–314	0.458	0.187	0.321	0.266

one shows the overlapping accuracy of Orbit Study. The bottom one shows the overlapping accuracy of PANDA (Table 7.7).

From orbit determination results, we can see that the BDS orbit determination accuracy of Orbit Study is 10 cm worse than PANDA in the radial direction. There are two reasons for this result:

On the one hand, the strategy for precise orbit determination of PANDA is as follows: firstly, resolving coordinate, clock offset and zenith tropospheric delay

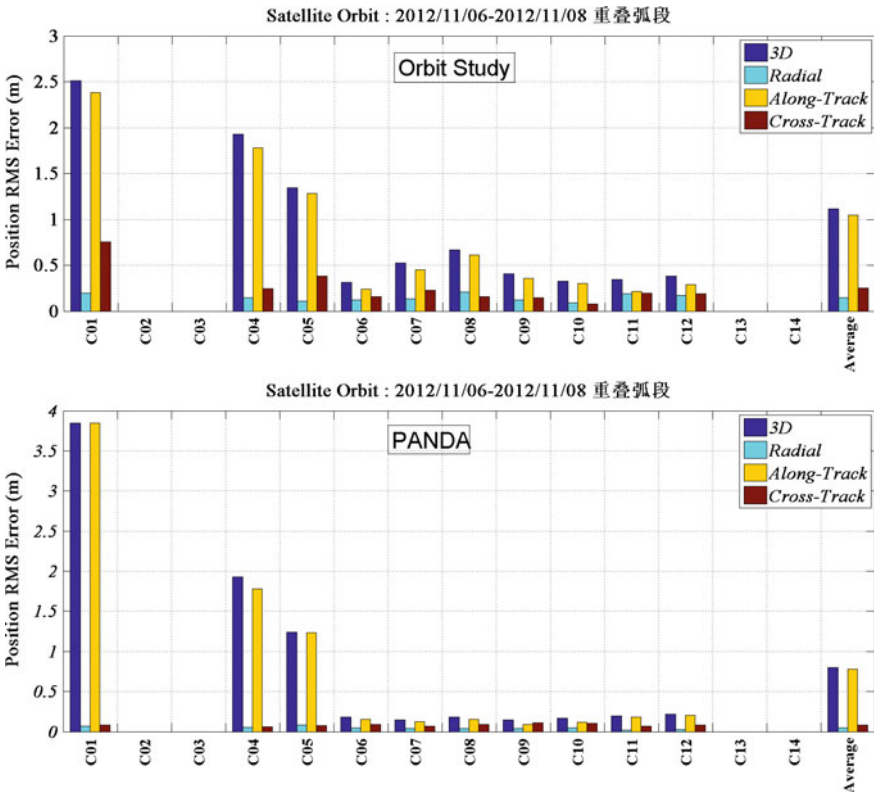


Fig. 7.7 Overlapping accuracy of Orbit Study and PANDA

Table 7.7 Overlapping accuracy of Orbit Study and PANDA (m)

		3D	Radial	Along	Cross
GEO	Orbit study	1.929	0.151	1.817	0.462
	PANDA	2.338	0.069	2.286	0.074
IGSO	Orbit study	0.446	0.132	0.394	0.152
	PANDA	0.164	0.042	0.126	0.092
MEO	Orbit study	0.364	0.178	0.251	0.192
	PANDA	0.206	0.020	0.190	0.076

(ZTD) parameters for ground stations. The GPS measurements are only used in this step; then fixing ground receivers' clock offset and ZTD parameters and resolving the initial position of each of the six Beidou satellites, as well as satellite clock offset and nine light pressure parameters. This strategy is different from Orbit Study. The BDS measurements are only used for resolving the BDS dynamic parameters. The GPS measurements are not used in Orbit Study.

On the other hand, because of this software is preliminary developed, there are many details should be researched. For example, the module that ambiguity resolution is no used in Orbit Study.

7.5 Conclusions and Plans

A set of orbit determination software named Orbit Study is described, which is developed under VC60.0 platform. EKF parameter estimation is used in this software. Many special algorithms for improving the computational efficiency are used in this software. The key technology is described in part 2. The measurement model and dynamic model are described in part 3. The accuracy evaluation result using Orbit Study for BDS orbit determination is described in part 4. Two types are used for accuracy evaluation, which are overlapping accuracy and comparing with PANDA. The estimation result shows the accuracy of 20 cm as the satellite position in radial direction.

The reason why the accuracy of Orbit Study is lower than PANDA is analyzed. The next plans are as follows: Firstly, the error correction for measurements is to be refined. Secondly, the module that ambiguity resolution is joined. At last, the strategy that PANDA for BDS orbit determination is planned to join.

References

1. <http://www.BeiDou.gov.cn/>
2. Yang Y (2010) Progress, contribution and challenges of compass/BeiDou satellite navigation system. *Acta Geodaetica et Cartographica Sinica* 39(1):1–6
3. Catalan C, Hernandez C, Mozo A et al (2008) Improve integrity concept for future GNSS evolutions. In: Proceedings of 21st international technical meeting of the satellite division of the institute of navigation (ION GNSS 2008), Savannah, GA, September 2008, pp 2547–2557

4. Lee YC (2011) A position domain relative RAIM method. *IEEE Trans Aerosp Electron Syst* 47(1):85, 97
5. Shi C, Zhao Q, Li M et al (2012) Precise orbit determination of Beidou satellites with precise positioning. *Scientia Sinica Terrae* 55:1079–1086
6. Steigenberger P, Hugentobler U, Hauschild A et al (2013) Orbit and clock analysis of compass GEO and IGSO satellites. *J Geod.* doi:10.1007/s00190-013-0625-4
7. <http://www.igs.org/>
8. Zhao Q (2004) Research on precision orbit determination theory and software of both GPS navigation constellation and LEO satellites. PhD. Wuhan University,
9. Shi C, Zhao Q, Lou Y et al (2009) PANDA: comprehensive processing software for satellite navigation systems and its research progress. *Spacecraft Eng* 18(4):64–70
10. Ruan R, Jia X, Wu X et al (2014) SPODS software and its result of precise orbit determination for GNSS satellites. In: China satellite navigation conference (CSNC) 2014 proceedings, vol III. doi:10.1007/978-3-642-54740-9_27
11. Blewitt G (1990) An automatic editing algorithm for GPS data. *Geophys Res Lett* 17(3): 199–202
12. <http://www.iausofa.org/>

Chapter 8

Analysis of Multi-frequency BDS/GPS RTK Positioning

Yijun Tian, Dongqing Zhao, Shuangna Zhang, Zhiyong Huang
and Hang Dong

Abstract In this contribution, quality control which includes number of visible satellites, PDOP and multipath of BDS+GPS integrated system in varying cut-off elevations is analyzed by a lot of real datasets, and then the performance of BDS+GPS multi-frequency RTK positioning over short baseline is evaluated. The following conclusions can be drawn: more satellites are visible in the combined BDS+GPS system, which ensures the spatial geometry intensity even in case of high cut-off elevations; With the cut-off elevation increases, the impact of the multipath effects on BDS and GPS is significantly weakened, so the AR success rate is mainly affected by the spatial geometry intensity in case of high cut-off elevations; The triple-frequency BDS-only or combined system can perform as well as that of dual-frequency, but it is not much beneficial to positioning performance; The AR success rate and positioning accuracy of BDS+GPS combined system in case of high of cut-off elevations still remain high, which will significantly improve the availability and reliability of the current RTK positioning under some challenging conditions.

Keywords BDS/GPS · Quality control · RTK · Availability · Reliability

Supported by the Open Research Fund of The Academy of Satellite Application under grant No. 2014_CXJJ-DH_14.

Y. Tian (✉) · D. Zhao · Z. Huang
Information Engineering University, Zhengzhou 450052, China
e-mail: 2531015503@qq.com

S. Zhang
Space Star Technology Co., Ltd, Beijing 10086, China

H. Dong
Troop 61206, Beijing, China

8.1 Introduction

With the built of multiple satellite navigation systems, the navigation applications of a single GPS system have been expanded to the coexistence of multiple satellite systems, and multi-system positioning in continuity, availability, reliability, accuracy and efficiency have more advantages. After the BeiDou satellite navigation system officially provide regional services in the Asia-Pacific region, the number of visible satellites increase to 20 [1], and therefore assessment of the BDS/GPS RTK positioning urgently need to be performed.

Cai [2] tested the BDS data quality and performance for single point positioning (SPP) under different observing conditions. Odolinski [3] compared the performance of the combined BDS-GPS RTK positioning with that of BDS- and GPS-only. He [4] demonstrated the performance of the BDS/GPS single-epoch positioning in both static and kinematic modes. Teunissen [5] gave a formal and empirical analysis of the single-epoch RTK positioning capabilities of BeiDou +GPS. All above are research on single or dual-frequency combined BDS+GPS system, little about triple-frequency is mentioned.

In order to test the data quality and performance of the combined BDS-GPS multi-frequency high-precision RTK positioning, experiments have been conducted under different cut-off elevations. To assess the BDS data quality, the PDOP, code multipath and availability were analyzed and compared to GPS data. The results show that adding BDS observations can significantly improve the positioning accuracy of GPS-only RTK in environments with limited satellite visibility.

8.2 The GPS+BDS GNSS Model

The GPS- or BDS-only double differenced code and phase observation equation is defined as follows:

$$\phi_{ur,*}^{st} = \rho_{ur,*}^{st} + T_{ur,*}^{st} - I_{ur,*}^{st} + \lambda_{*} N_{ur,*}^{st} + \varepsilon_{ur,*}^{st} \quad (8.1)$$

$$P_{ur,*}^{st} = \rho_{ur,*}^{st} + T_{ur,*}^{st} + I_{ur,*}^{st} + e_{ur,*}^{st} \quad (8.2)$$

In which $\phi_{ur,*}^{st}$ and $P_{ur,*}^{st}$ denotes the system specific DD phase and code observation vectors, respectively, with $*$ = {G, B} (G = GPS, B = BDS), $T_{ur,*}^{st}$ and $I_{ur,*}^{st}$ denote the troposphere and ionosphere delay error, $N_{ur,*}^{st}$ denote the DD integer ambiguity vector.

We assume that n_G GPS-satellites are tracked on two frequencies and n_B BDS on triple frequency, the combined short-baseline GPS+BDS model is then given as

$$\begin{aligned}
 E \begin{bmatrix} \phi_G \\ \phi_B \\ P_G \\ P_B \end{bmatrix} &= \begin{bmatrix} A_G & \Lambda_G & 0 \\ A_B & 0 & \Lambda_B \\ A_G & 0 & 0 \\ A_B & 0 & 0 \end{bmatrix} \begin{bmatrix} dX \\ N_G \\ N_B \end{bmatrix} \\
 D \begin{bmatrix} \phi_G \\ \phi_B \\ P_G \\ P_B \end{bmatrix} &= \begin{bmatrix} Q_{\phi_G \phi_G} & & & \\ & Q_{\phi_B \phi_B} & & \\ & & Q_{P_G P_G} & \\ & & & Q_{P_B P_B} \end{bmatrix}
 \end{aligned} \tag{8.3}$$

where $E[\cdot]$ and $D[\cdot]$ denote the expectation and dispersion operator, respectively. The entries of the design matrix is given as

$$\begin{aligned}
 A_* &= \begin{bmatrix} l_*^2 - l_*^1 & m_*^2 - m_*^1 & k_*^2 - k_*^1 \\ \vdots & \vdots & \vdots \\ l_*^{n_*} - l_*^1 & m_*^{n_*} - m_*^1 & k_*^{n_*} - k_*^1 \\ l_*^2 - l_*^1 & m_*^2 - m_*^1 & k_*^2 - k_*^1 \\ \vdots & \vdots & \vdots \\ l_*^{n_*} - l_*^1 & m_*^{n_*} - m_*^1 & k_*^{n_*} - k_*^1 \end{bmatrix} \\
 \Lambda_* &= \begin{bmatrix} \lambda_{*,1} & & & & \\ & \ddots & & & \\ & & \lambda_{*,1} & & \\ & & & \lambda_{*,2} & \\ & & & & \ddots \\ & & & & & \lambda_{*,2} \end{bmatrix}
 \end{aligned}$$

In the above definition, the diagonal matrix Λ_* contains the wavelengths of the observed frequencies and A_* contains the un-differenced receiver-satellite unit direction vectors for GPS and BDS, respectively.

8.3 Quality Analysis of GPS+BDS Measurements

8.3.1 Analysis of Visibility and PDOP Value

The number of visible satellites and PDOP is used to analysis the intensity of space geometry, and the real broadcast ephemeris was used to calculate the satellite position when evaluating the number of visible satellite and PDOP value.

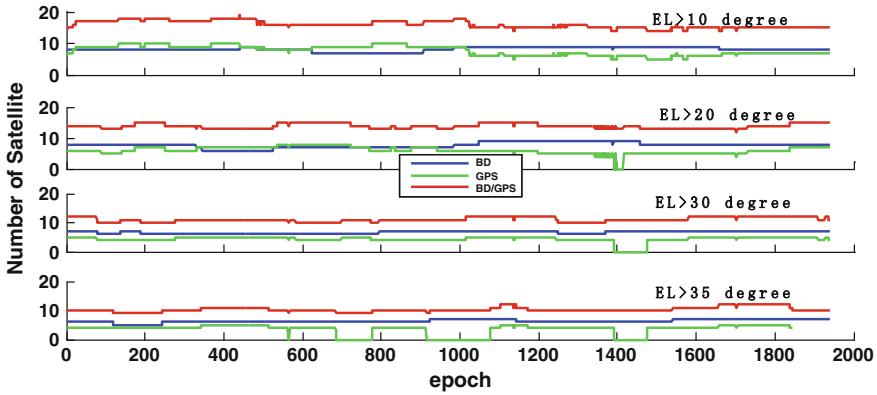


Fig. 8.1 Number of visible satellites and PDOP value

Table 8.1 Average values of visible satellites and PDOP

	15		20		25		30		35	
	Nsat	PDOP	Nsat	PDOP	Nsat	PDOP	Nsat	PDOP	Nsat	PDOP
GPS	6.9	2.6	6.1	3.5	5.2	6.0	4.5	7.7	3.5	7.6
BD	8.0	4.1	7.8	4.3	7.5	4.7	6.6	7.5	6.3	8.3
BG	15.0	1.6	13.9	1.8	12.6	2.2	11.1	3.2	10.3	3.8

In Fig. 8.1, the visible satellites of BDS are between 7 and 9, GPS between 4 and 9 while BDS/GPS combined system between 13 and 17. With the increase of cut-off elevations, visible satellites of BDS- and GPS-only system significantly reduce, while BDS/GPS remains high.

It can be seen from Table 8.1 that: The average visible satellites of BDS are apparently more than GPS, however, the PDOP value is mostly smaller than GPS, which means geometric configuration of BDS still need to strengthen. The availability of BDS/GPS is greatly superior to BDS-only and GPS-only, especially in case of high cut-off elevations.

8.3.2 Multipath Combination

Considering short baseline, we apply system-specific double-differencing (DD) and thus most systematic errors are eliminated. The key elements which affect positioning accuracy are multipath and thermal noise plus noise of code measurements, so we need to construct the multipath combination (MPC) to assess multipath and code noise level of a receiver. The MPCs for the receiver and satellite can be respectively written as:

$$MP1_*^s = P_{*,1}^s - \frac{f_1^2 + f_2^2}{f_1^2 - f_2^2} \phi_{*,1}^s + \frac{2f_2^2}{f_1^2 - f_2^2} \phi_{*,2}^s \quad (8.4)$$

$$MP2_*^s = P_{*,2}^s - \frac{2f_1^2}{f_1^2 - f_2^2} \phi_{*,1}^s + \frac{f_1^2 + f_2^2}{f_1^2 - f_2^2} \phi_{*,2}^s \quad (8.5)$$

$$MP3_*^s = P_{*,3}^s - \frac{2f_1^2}{f_1^2 - f_3^2} \phi_{*,1}^s + \frac{f_1^2 + f_3^2}{f_1^2 - f_3^2} \phi_{*,3}^s \quad (8.6)$$

where P_*^s and ϕ_*^s are the code and carrier phase measurements, respectively, and $*$ = $\{G, B\}$ ($G = GPS, B = BDS$). In the MPCs, the systematic errors and noise of carrier phase measurements are assumed negligible compared with that of code observations. The first-order ionosphere effect and geometric range between the satellite and the receiver are eliminated, and the MPCs only consist of a constant ambiguity term and thermal noise plus noise of code measurements.

Figure 8.2 shows the time series of B1, B2 and B3 for the BDS satellite and L1, L2 for the GPS satellites. It can be seen that with the variation of elevations angles, the MPCs change slowly, especially for the GPS satellites. For example, the time series of G26 have typical multipath effects with stronger variations of a few meters at two ends of the pass and smaller variations in between.

Figure 8.3 illustrates the STD of the multipath combinations for each BDS and GPS satellite. It can be seen that the BDS code noise on the B1 frequency which is slightly smaller than B2 frequency is greatly larger than B3 frequency whereas the GPS code observations on the L1 frequency have a significantly larger noise level than L2 frequency. This indicates that the B1/B2 code are noisier than B3 code while L1 noisier than L2. It should be mentioned that BDS can perform as good as GPS in real time kinematic (RTK) positioning.

8.4 Positioning Results

A single-baseline RTK analysis was performed to verify the quality claims of the previous section and to study the actual BDS+GPS performance. The standard broadcast ephemerides were used to provide for the BDS and GPS satellite orbits and clocks.

We start with the success-rate analysis. The ILS success rate is defined as,

$$P_{ILS} = \frac{\text{Number of correctly fixed epochs}}{\text{total number of epochs}}$$

The results of single-epoch success-rate for seven different cut-off elevations are shown in Table 8.2. In the success-rate behavior of 5 modes, the earlier refer to presence of low-elevation multipath. As the cut-off elevation angle increase to 20°,

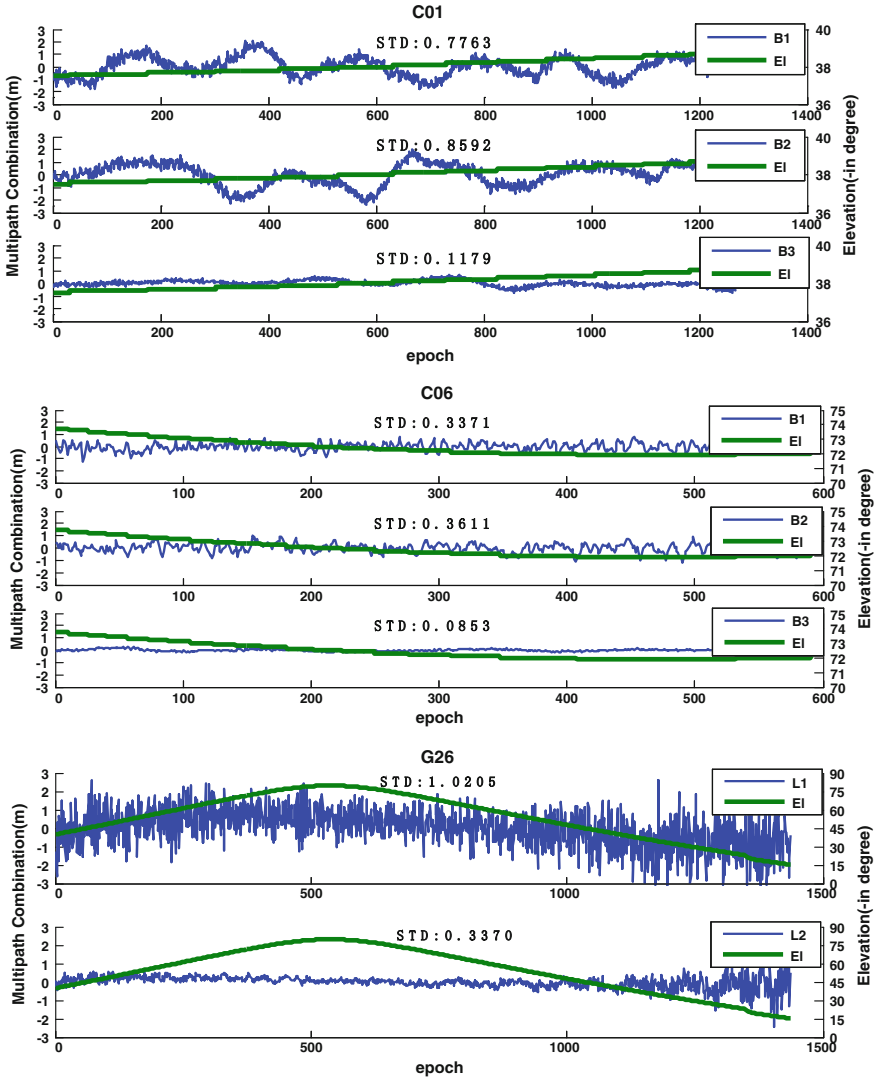


Fig. 8.2 MPCs against elevation for the BDS or GPS satellite

multipath disappears and the success-rate increases. So here we can see an apparent advantage which successful ambiguity resolution at higher cut-off elevation brings. While the cut-off elevation is greater than 20°, the success-rates of BDS-only and GPS-only all get smaller as the cut-off elevation gets larger, and the decrease amplitude of three frequencies is slightly less than double frequencies. This occurrence is mainly related to the number of tracked satellites. As the cut-off elevation gets larger, less satellites of GPS or BDS are tracked, which influence the

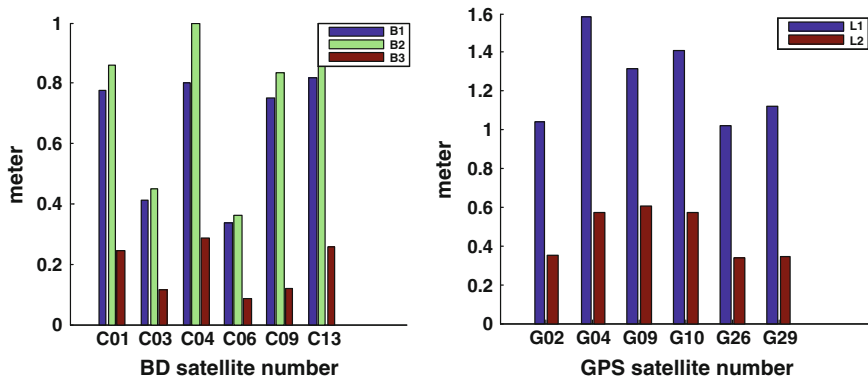


Fig. 8.3 STD of MPCs for each BDS and GPS satellite

Table 8.2 Single-epoch ILS success-rate for 10 – 40°

System	Success rate (%)						
	10	15	20	25	30	35	40
GPS	88.9	97.2	97.6	95.2	95.2	80.7	56.2
BDS2	97.5	98.6	98.9	99.6	97.8	97.2	36.7
BDS3	97.6	98.8	99.0	99.7	99.6	99.4	62.4
BG2	70.9	91.1	98.1	99.9	100	100	100
BG3	70.2	86.9	97.7	99.9	100	100	100

intensity of space geometry. Just as the double-frequency, combined systems achieve higher success rate (up to 25°), even being successful up to 40° cut-off elevation, also the triple-frequency. Comparing the relative effect of an added frequency versus an added system, we can conclude that the effect of the former is apparently lower than the latter.

In Fig. 8.4, results are shown for different systems and different frequencies. We can see that the positioning performance of BDS is as good as GPS, which implies that BDS can realize RTK positioning alone.

See Fig. 8.5, positioning accuracy of single system clearly becomes lower as the cut-off elevations gets larger due to the fact that the total number of their visible satellite drops, especially GPS. While maintaining a higher success-rate, combined system has of course the best positioning performance of all, and it still enables high precision positioning even in case of 35° cut-off elevation. With the addition of BDS to GPS, the improved spatial geometry not only improved ambiguity resolution, but it also avoids that in the many more correctly fixed solutions. However, with the addition B3 to BDS double-frequency, no such improved receiver-satellite geometry is present. Hence, the presence of B3 will now not be benefit to positioning performance.

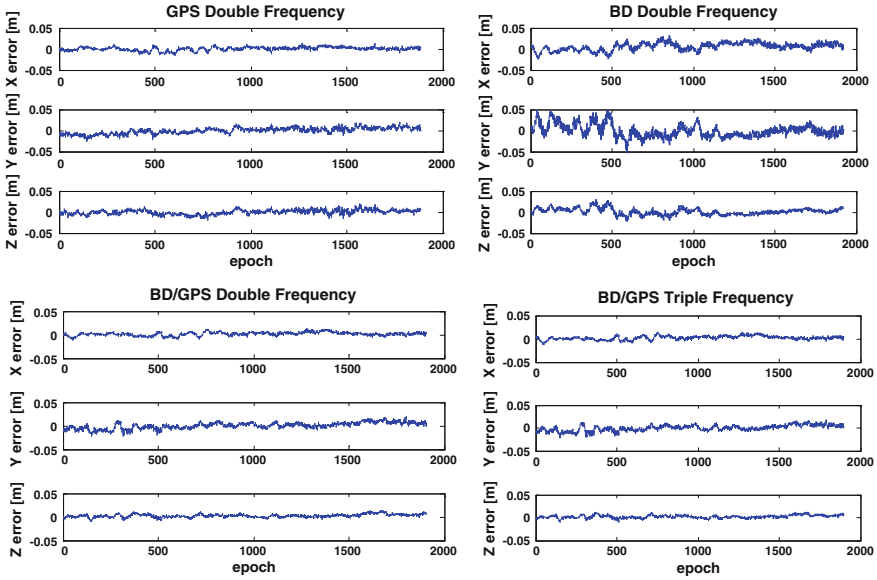


Fig. 8.4 Time series of BDS, GPS and BDS+GPS

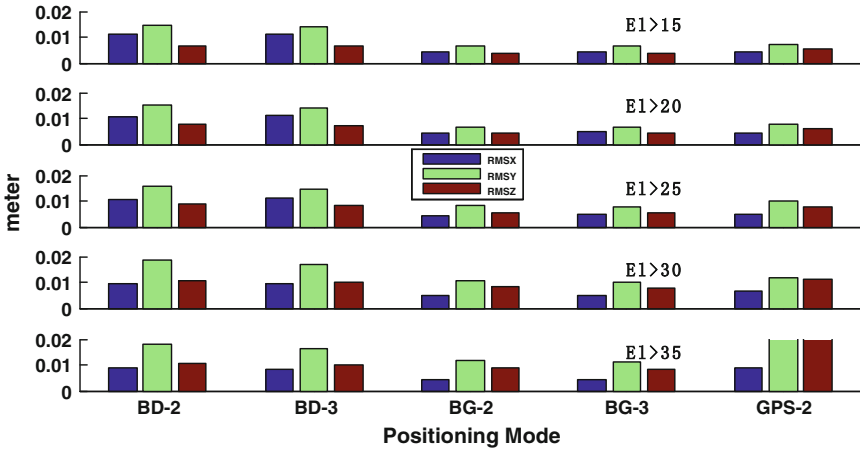


Fig. 8.5 RMS of baseline differences between the ground truth and kinematic solutions

8.5 Summary and Conclusions

In this contribution, quality control and instantaneous RTK positioning performance of a combined BDS+GPS system were analyzed for varying cut-off elevations. It was shown that more satellites are visible in the combined system, which

will significantly reduce the PDOP value, and the spatial geometry intensity remains high even in case of high cut-off elevations. The given analytical MPC provide good understanding of how the multipath affects the AR success rate. The actual ambiguity resolution performance for varying cut-off elevations was also analyzed. Ambiguity resolution and positioning are namely driven by different factors of the GNSS model, especially when increasing the cut-off elevations. The triple-frequency BDS-only or combined system can perform as well as that of double-frequency, but it is not much beneficial to positioning performance. The combined system has of course better performance than the single system. It was also shown that with the combined system much larger cut-off elevations can be used, which means such measurement set-up will significantly increase the GNSS applicability in constrained environments, such as in urban canyons or when low-elevation multipath is present.

References

1. Yang Y et al (2014) Preliminary assessment of the navigation and positioning performance of BeiDou regional navigation satellite system. *Sci China Earth Sci* 57(1):144–152
2. Cai C (2013) An analysis on combined GPS/COMPASS data quality and its effect on single point positioning accuracy under different observing conditions. *Adv Space Res*
3. Odolinski R et al (2013) An analysis of combined COMPASS/BeiDou-2 and GPS single- and multiple-frequency RTK positioning. In: *Proceedings of the institute of navigation PNT 2013*
4. He H et al (2014) Performance assessment of single- and dual-frequency BeiDou/GPS single-epoch kinematic positioning. *GPS Solut* 18(3):393–403
5. Teunissen PJG et al (2014) Instantaneous BeiDou+GPS RTK positioning with high cut-off elevation angles. *J Geodesy* 88(4):335–350

Chapter 9

Upper Atmospheric Density Retrieval from Accelerometer on Board GRACE Mission

Runjing Chen and Bibo Peng

Abstract In order to improve the prediction accuracy of LEO, it is essential to build up an accurate atmospheric model for density prediction. However, most existing atmospheric models belong to the type of semi-empirical model, thus the data sets are not of homogeneous quality and have limited geographical and temporal coverage. Since the space-borne accelerometer could measure the total non-conservative accelerations acting on LEO directly, the atmospheric drag component could be isolated with the help of the solar and earth albedo radiation pressure models, then the atmospheric density can be calculated, which provides necessary data for making evaluation and improvement of the existing atmospheric models. This paper describes the method to retrieve the upper atmospheric density from accelerometer in detail, 3 months of observations spanning from May 2013 to July 2013 are selected to do the experiment, we use the dynamical orbit determination strategy to calibrate the accelerometers, and then retrieve the air density at the altitude of GRACE Mission. The results show that prediction models cannot exhibit the density variation in high frequency, and the in situ measurements are very useful in density analysis, in addition, it is validated that atmospheric density has a positive correlation with the solar activity intensity.

Keywords Space-borne accelerometer · Dynamical orbit determination · Accurate non-conservative force model · Upper air density · Air density model

R. Chen (✉)

Test and Assessment Research Center, China Satellite Navigation Office, GNSS Center,
China Academy of Aerospace Electronic Technology, 1 Fengyingdong Road,
Beijing 10094, China
e-mail: chenrunjing@163.com

B. Peng

Institute of Geodesy and Geophysics, Chinese Academy of Sciences, 340 Xudong Road,
Wuhan 430077, China

© Springer-Verlag Berlin Heidelberg 2015

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015*

Proceedings: Volume III, Lecture Notes in Electrical Engineering 342,

DOI 10.1007/978-3-662-46632-2_9

9.1 Introduction

Compared to high-orbit satellite, LEOs suffer from a significant increase of the air mass density in the surrounding circumstance. As the atmospheric drag is proportional to the air mass density, it plays the most essential role in all the non-conservative forces acting on LEOs. Unfortunately, it is very difficult to predict the air mass density in this region accurately beforehand due to its complex physical interaction with a multiple of space weather factors such as solar activity, geomagnetic field variation, earth rotation, etc.

In situ measurement of atmospheric density in high temporal and spatial resolution acts as an essential means to assess and improve existing atmospheric density models. SuperStar accelerometer on board GRACE Mission could accurately measure all the non-gravitational acceleration acting on the satellite, which provide a new way to retrieve the air density. The SuperStar accelerometer consists of a proof mass inside its box, which is immune to non-gravitational force. For satellite body itself, it suffers from not only the gravitational force but also the non-gravitational force. Due to the suffering force differences between the satellite surface and the proof mass, the relative motion takes place. By producing the necessary voltage at the electrodes, the proof mass could be kept at the center of the body. The voltage can be converted to the non-gravitational force acceleration in proportion, and then we finally get the value of non-gravitational force [1]. As the atmospheric drag act as the most important perturbation non-gravitational force, and the solar and earth radiation pressure can be calculated precisely by accurate physical models, the drag could be isolated from the total non-gravitational acceleration. Based on the aerodynamics, the drag coefficient can be obtained, atmospheric density will finally be retrieved when the satellite's geometry and its attitude are known.

To study the atmospheric density is of great significance. It can be applied in scientific investigations as well as in many types of satellite orbit calculations. It contributes to prediction of the LEOs orbit and its lifetime, especially for the re-entry trajectory. A priori knowledge of the satellite position and attitude in future time helps the scientists to analyze potential risk for satellites and send maneuver signals if necessary. In addition, there are a large number of orbiting objects in space which can be classified into space debris, they should be carefully tracked, as some of them have very large area-to-mass ratio, it might cause catastrophic consequences in a collision with an operational satellite [2].

9.2 Accelerometer Calibration

Due to the complexity and particularity of the space environment, it is very difficult to calibrate accelerometer in the laboratory, and the output values from accelerometer are not the true value for non-gravitational acceleration, therefore calibration

must be done. In general, the real non-gravitational acceleration acting on GRACE can be expressed as:

$$\vec{f}_{true} = \vec{B} + S\vec{f}_{obs} \quad (9.1)$$

where \vec{f}_{obs} is the raw measurement value of ACC in the satellite reference frame [3]; \vec{B} is the bias vector in the measured acceleration, S is a 3x3 orthogonal matrix with its diagonal elements as the scale vector for \vec{f}_{obs} . The off-diagonal elements of S are all set to be zero, and \vec{f}_{true} is the true value for accelerometer.

To estimate the bias and scale factors, we consider them as orbit dynamical unknown parameters and fit the orbit to the GPS phase observations in the process of POD to work them out. The dynamic orbit determination technique is applied to obtain best fit orbit. A software developed by IGG¹ is used to implement this task. Gravitational forces are precisely calculated following the *IERS 2010* [4], which includes a series of corrections to the in situ gravitational model coefficients. As for non-gravitational forces, they are replaced by the observations from SuperSTAR accelerometer where a priori calibration parameters [5] has been applied to correct the raw data.

The direct output of observed non-gravitational forces from accelerometer are represented in the SRF (Science Reference Frame), which is defined as follows. It has its origin at the center of mass of the satellite, the X axis points to the its twin sister satellite which nominally aligns with the velocity direction, the Z axis is perpendicular to the X direction pointing toward the radial direction, and the Y axis completes the right-handed coordination system with the Z and X axes. In order to transform the ACC measurement to the inertial frame, we need to use the quaternions from the SCA1B file to form rotation matrix which is released together in the Level1B file packages. For missing data in SCA1B file, we use the following method to make interpolation of the quaternions at those epochs we need [6].

Given are two arbitrary normalized quaternions referring to epochs t_a and t_b ($t_b > t_a$), linear interpolation to an epoch between t_a and t_b , we should firstly determine the sign ambiguity. If the scalar product of the vector parts of the two quaternions is negative, all elements in q_b , have to change their sign:

$$q_b = -q_b \quad \text{if} \quad (q_{a1}q_{b1} + q_{a2}q_{b2} + q_{a3}q_{b3} < 0) \quad (9.2)$$

The quaternion describing the differential rotation between the rotations q_b and q_a can be written by

$$q_{ab} = q_a^* q_b \quad (9.3)$$

¹Institute of Geodesy and Geophysics, Chinese Academy of Sciences.

where

$$\begin{aligned}
 q_{ab4} &= q_{a4}q_{b4} + q_{a1}q_{b1} + q_{a2}q_{b2} + q_{a3}q_{b3} \\
 q_{ab1} &= q_{a4}q_{b1} - q_{a1}q_{b4} + q_{a3}q_{b2} - q_{a2}q_{b3} \\
 q_{ab2} &= q_{a4}q_{b2} - q_{a2}q_{b4} + q_{a1}q_{b3} - q_{a3}q_{b1} \\
 q_{ab3} &= q_{a4}q_{b3} - q_{a3}q_{b4} + q_{a2}q_{b1} - q_{a1}q_{b2}
 \end{aligned} \tag{9.4}$$

The rotation angle corresponding to the rotation described by q_{ab} is

$$\Phi_{ab} = 2 \arccos(q_{ab4}) \tag{9.5}$$

This angle may be linearly interpolated

$$q_{at} = \frac{t - t_a}{t_b - t_a} \Phi_{ab} \tag{9.6}$$

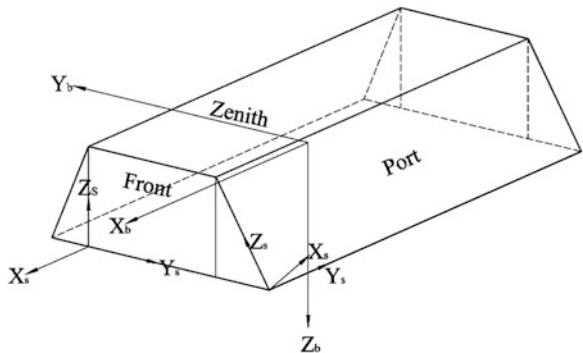
The quaternion q_t for interpolating t can finally be written as

$$q_t = q_a q_{at} \tag{9.7}$$

9.3 Satellite Configuration

The simplified surface model as well as the optical parameters for each plane are given in Ref. [3]. Figure 9.1 is the sketch map for the model, which is a combination of six panels. The Front and Rear refer to the head and the tail plane of GRACE Mission in the velocity direction respectively, Starboard and Port are the planes at both sides, Zenith and Nadir correspond to the top and the bottom plane. X_b, Y_b, Z_b in the figure refer to the three axes of satellite frame, while X_s, Y_s and Z_s are the panel frame with X_s pointing to the normal direction of each panel.

Fig. 9.1 Simplified GRACE surface model



9.4 Solar and Earth Radiation Pressure Acceleration

The GRACE Mission suffers from the radiation pressure due to the incident radiation from solar and earth, which is the net momentum change per unit time. Acceleration by solar radiation can be calculated as follows [11, 12]:

$$\vec{f}_{solar} = -f\gamma \frac{S_{\oplus}}{c} \left[\frac{A_u}{r} \right]^2 \sum_{k=1}^6 \frac{A_k}{m} \cdot (\vec{n}_k \cdot \vec{s}) \left\{ (1 - \rho_k) \vec{s} + 2 \left[\frac{\delta}{3} + \rho_k \cdot (\vec{n}_k \cdot \vec{s}) \right] \vec{n}_k \right\} \quad (9.8)$$

where f accounts for the satellite eclipse, usually a conical model is applied to determine whether the satellite is exposed to the solar ray, S_{\oplus} is the solar flux constant with value of 1366.1 W/m^2 , c is the speed of light in a vacuum, A_u is the astronomical unit, and r is the distance from the Sun to the satellite, m is the satellite mass which can be obtained from the Level 1B MAS file. A_k , ρ_k and δ_k indicate the area, specular and diffuse reflectivity coefficients of the panel k respectively. Equation (9.8) shows that the solar radiation force is the sum of two components: one acts along the opposite direction of radiation \vec{s} , the other one acts along the opposite of normal direction. γ can be used to adjust the resultant force to account for model error.

Earth radiation pressure comprises two components, one is the solar radiation reflected from the earth surface directly when it reaches the earth, the other component is the radiation reemitted from the earth which has been absorbed earlier. In this case, we divide the earth into many small elements, and consider the earth radiation pressure is the sum force from each of these elements. The equation can be expressed in the following way:

$$\vec{f}_{refl} = -\gamma^{refl} \frac{S_{\oplus}}{c} \left[\frac{A_u}{R} \right]^2 \frac{1}{m} \sum_{i=1}^N \sum_{k=1}^6 v_i \cdot (\vec{n}^i \cdot \vec{R}) \cdot A_i \cdot A_k \cdot (\vec{n}_k \cdot \vec{s}^i) \left\{ (1 - \rho_k^{vis}) \vec{s}^i + 2 \left[\frac{\delta_k^{vis}}{3} + \rho_k^{vis} (\vec{n}_k \cdot \vec{s}^i) \right] \vec{n}_k \right\} \quad (9.9)$$

$$\vec{f}_{emit} = -\gamma^{emit} \frac{S_{\oplus}}{c} \left[\frac{A_u}{R} \right]^2 \frac{1}{m} \sum_{i=1}^N \sum_{k=1}^6 \frac{e_i}{4} A_i \cdot A_k \cdot (\vec{n}_k \cdot \vec{s}^i) \left\{ (1 - \rho_k^{IR}) \vec{s}^i + 2 \left[\frac{\delta_k^{IR}}{3} + \rho_k^{IR} (\vec{n}_k \cdot \vec{s}^i) \right] \vec{n}_k \right\} \quad (9.10)$$

where γ^{refl} and γ^{emit} is the adjustment coefficients for the modeled force, \vec{R} is the unit vector pointing from the earth to the sun, v_i , e_i are the albedo and emissivity of the i th earth surface element, \vec{n}^i corresponds to its normal direction, \vec{s}^i is the unit vector for the i th element pointing from the surface to the satellite, A_i is the area of the element, ρ_k^{vis} , δ_k^{vis} , ρ_k^{IR} , δ_k^{IR} are reflectivity and emissivity coefficients for k th panel of the satellite. Other symbols have the same meaning as the equations above.

9.5 Upper Atmospheric Density Retrieval

The atmospheric drag acts primarily in the along-track component, meanwhile ACC observations in X component have few spike signals, so it is ideal to use measurement in this component to retrieve upper atmospheric density. As the calibrated ACC observations consist of all the non-gravitational accelerations, drag could only be obtained when the solar and earth radiation acceleration have been removed from the total acceleration. Therefore the atmospheric density can be calculated as follows:

$$f_{obs-d} = f_{cal} - f_{solar} - f_{reft} - f_{emit} \quad (9.11)$$

$$\rho = \frac{2m\vec{f}_{obs-d,x}}{C_a A v_r^2} \quad (9.12)$$

Where m is the satellite mass, $\vec{f}_{obs-d,x}$ is the atmospheric drag in the along-track component, C_a is the drag coefficient, A is the projected area in the along-track component for GRACE, v_r is the satellite moving velocity relative to surrounding atmosphere. In fact, the in-track velocity change induced by atmospheric rotation and wind is neglectable compared to satellite's velocity in the same direction, so calculations of the accurate drag coefficient and area-to-mass ratio are key points in density retrieval. For satellite mass reduction due to gas burning, we can get the real-time mass value from MAS Level 1B file.

9.6 Experiment and Result

In order to retrieve the upper atmospheric density following the method proposed in this paper, we use 92 days of the data for GRACE-A and GRACE-B spanning from May 1st to July 31st in 2013 to do the experiment. Dynamical orbit determination strategy is applied to work out calibration factors, all the models used in the software are listed in Table 9.1. For calculation of the solar radiation pressure acceleration, conical shadowing model is employed to account for the shadowing factor, solar position can be obtained from DE405 ephemeris, the simplified GRACE surface model is used to represent the shape of the GRACE Mission. As the reflectivity and emissivity coefficients for each small area element on the earth surface are required in calculation of the earth radiation acceleration, measurements from spectrometer FM1 on board Terra Mission are used which includes worldwide $2.5^\circ \times 2.5^\circ$ monthly-average albedo and longwave flux [13]. In addition, HWM07 model [14] is used to get the wind, satellite mass can be found in the MAS1B file, the projected area for GRACE Mission in the flight direction can be calculated based on attitude data from SCA file, drag coefficient can be obtained using the Sentman model [15].

Table 9.1 Models and Parameters used in POD

Model	Description
Gravitational model	EGM2008 (150 × 150) [7]
Correction to low-degree coefficients	a: correction to \bar{C}_{20} \bar{C}_{30} \bar{C}_{40} using secular rates; b: correction to $\bar{C}_{21}(t)$ and $\bar{S}_{21}(t)$
Solid Earth tides	a: frequency-independent part: correction up to 4° coefficients b: frequency-dependent part: correction up to 2° coefficients
Ocean tides	FES 2004 [8], truncated up to 50 degrees
Solid Earth pole tide	Correction to $\bar{C}_{21}(t)$ and $\bar{S}_{21}(t)$
Ocean pole tide	correction up to 3° coefficients
N body perturbation	JPL DE405 [9]
General relativity	<i>IGRS 2010</i> [4]
Reference Frame	J2000.0
Precession/nutation	IAU2000A
Earth orientation parameters	EOP 08 C04 [10]
Non-gravitational acceleration	Accelerometer observations
POD arc length	24 h
Parameters to be estimated	Description
Initial satellite state	6 parameters (position and velocity)
Accelerometer calibration	3 bias factors, 3 scale factors
Ambiguity	Real value solution
Clock error	Real value solution

We consider the bias and scale factors as unknowns to estimate in POD, Fig. 9.2 is the result. It can be evidently seen that the bias and the scale factors were inversely proportional to each other. GRACE-A had the scale factors varying from 0.898 to 0.971, while the bias factors varied from -1322 to -1220 nm/s²; For GRACE-B, the scale factors varied from 0.906 to 0.942 while the bias factors varied from -634 to -591 nm/s². Overall, GRACE-A was more stable in calibration than GRACE-B. As the twin satellites were only 220 km apart from each other, they should suffer from almost the same non-gravitational acceleration in magnitude. Figure 9.3 displayed the time series of the difference between the calibrated measurements of GRACE-A and GRACE-B for one week, which showed that non-gravitational accelerations were more or less the same for GRACE-A and GRACE-B for the majority of the epochs, and the mean difference was only 4 nm/s² with the standard error of 11.027 nm/s². Big difference points corresponded to those epochs when the satellite flew across earth polar areas.

Atmospheric density retrieved from ACC observations on May 12, 2013 was compared with the output values from NRLMSISE-00 [16] and JB2008 models [17] in Fig. 9.4, good consistence could be found among the three lines, which

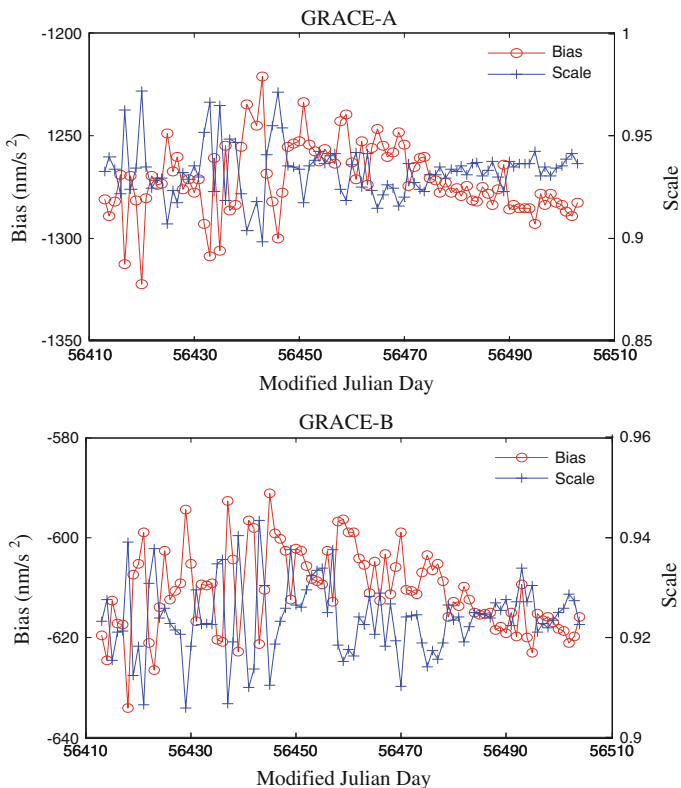


Fig. 9.2 Estimated calibration factors for GRACE-A and GRACE-B

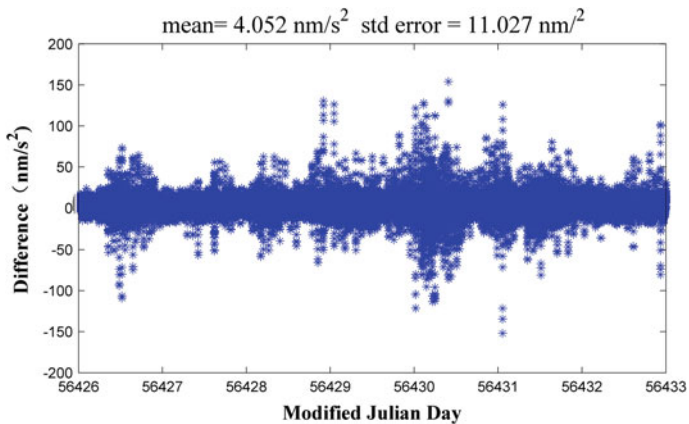
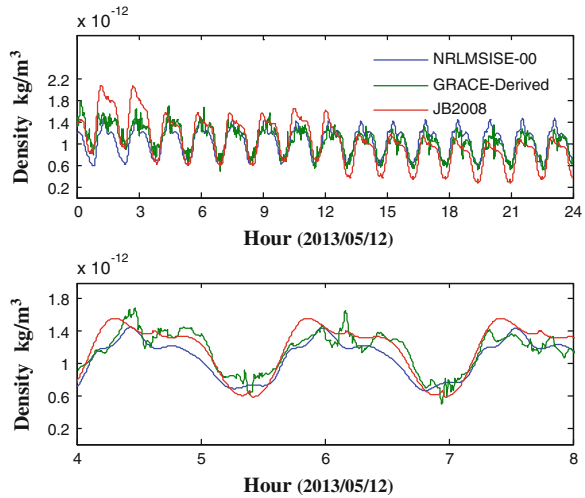


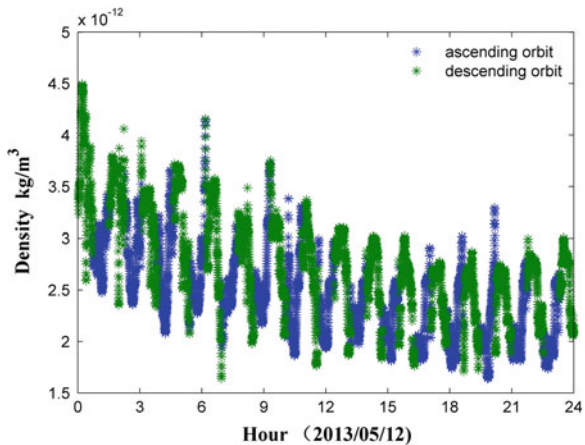
Fig. 9.3 Non-gravitational acceleration difference between GRACE-A and GRACE-B

Fig. 9.4 Atmospheric density derived from accelerometer on May 12, 2013 compared with predicted values by NRLMSISE-00 and JB2008 model



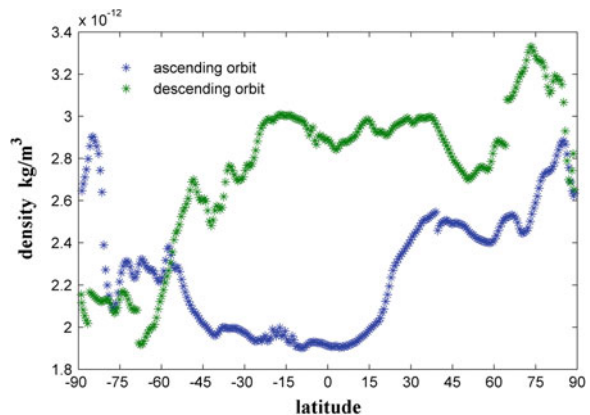
agreed with each other in trend and magnitude order. Only the ACC-derived atmospheric density could capture the variation in high frequency, while empirical models only reflect the stable variations, indicating that improvements were required for existing models. Figure 9.5 gives the atmospheric density difference for ascending and descending orbit arcs of GRACE-A Mission. As only hemisphere of the earth can be illuminated by sun light, the ascending and descending arcs correspond to the day and night side of the earth respectively (except that when orbital plane is perpendicular to the sun light). Due to absorption of the solar radiation energy, atmospheric density in the illuminated side becomes significantly larger than the dark side. It can be evidently seen that the descending orbit arc was in the

Fig. 9.5 Atmospheric density of ascending orbits and descending orbits on May 12, 2013



light side in Fig. 9.5, where the maximum density in this side was 50–70 % larger than the minimum value in the dark side, especially the first orbit arc with the most pronounced density discrepancy. In addition, when satellite moved around the earth, the earth also rotated itself synchronously, so different angles were formed between incident radiation vector and the GRACE orbit plane. Figure 9.6 gave the relationship between the atmospheric density and the latitude in the 7th orbit arc. In this time period, the satellite passed the longitude concentrated between W61° and W55° for the descending arc and concentrated between E129°–E134° for the ascending arc with the UTC ranging between 11:00 and 12:30, the sun's rays was in the vicinity of the meridian at the Greenwich, so solar illumination was more closer to the descending arc, making the atmospheric density around larger than it was for the ascending arc. Following conclusions could be drawn from Fig. 9.6: for light side of the earth, density decreased with increasing latitude in southern hemisphere, while reverse was found between N30° and N75° but decrease rapidly above N75°. Overall, density decrease in the southern hemisphere was more pronounced than the southern hemisphere, so that density in the arctic area was notably larger than the southern pole area; while for the dark side, atmospheric density increased along with the increasing latitude for both northern and southern hemispheres, especially for the northern hemisphere, but a small decrease area was found between N45° and N60°. To study the relationship between density variation and solar activity intensity, Fig. 9.7 gave the variation scope of atmospheric density at perigee of GRACE-A from May 1st to July 31st, 2013 along with corresponding F10.7 value at that day, where strong positive correlation could be evidently seen, big fluctuation of density was found at those days with high values of F10.7, which was one order magnitude larger maximally than quiet days.

Fig. 9.6 The relationship between variations of atmospheric density and latitude on May 12, 2013



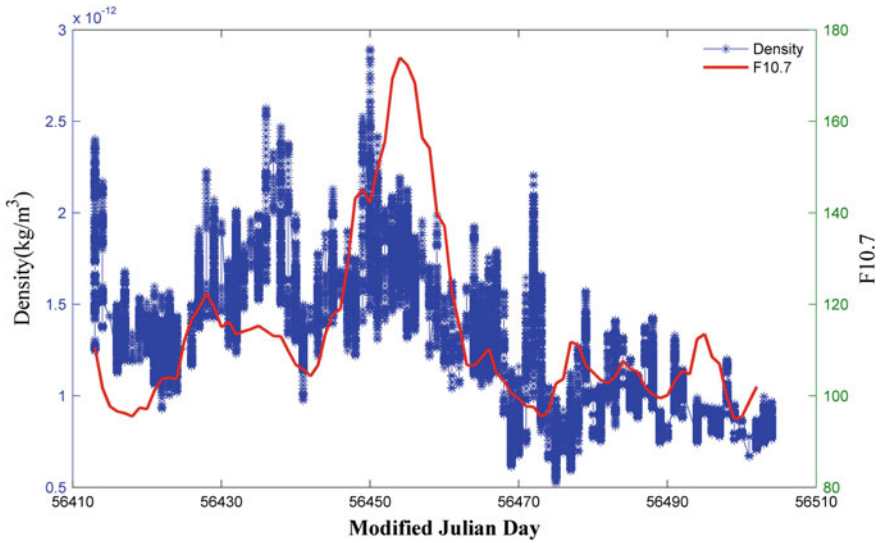


Fig. 9.7 The variation scope of Atmospheric density at perigee of GRACE-A from May 1st to July 31st, 2013 along with the F10.7 value at the corresponding day

9.7 Conclusion

This paper introduces the method to retrieve the atmospheric density by accelerometer on board GRACE Mission, and 92 days of data spanning from May 1 to July 31 in 2013 for GRACE-A and GRACE-B was used to do the experiment, we compare the derived density with the output from NRLMSISE-00 and JB2008 atmospheric models, indicating the superiority of accelerometer in extracting atmospheric density change in the high frequency. Besides, the atmospheric density is significantly larger in the light side of the earth than it is in the dark side, and distinct difference in variation can be found between the two sides. In addition, it is verified that atmospheric density variation is highly correlated with the solar activity intensity.

Acknowledgments This work was supported by Natural Science Foundation of China (40774012).

References

1. Bruinsma S et al (2004) Atmospheric densities derived from CHAMP/STAR accelerometer observations. *Planet Space Sci* 52(4):297–312
2. Doornbos E (2012) *Thermospheric density and wind determination from satellite dynamics*. Springer, Berlin

3. Case K et al (2002) GRACE level 1B data product user handbook. JPL Publication D-22027
4. Petit G, Luzum B (2010) IERS conventions (2010) (No. IERS-TN-36). Bureau International Des Poids Et Mesures Sevres (France)
5. Bettadpur S (2009) Recommendation for a-priori bias and scale parameters for level-1B ACC data (Version 2). GRACE TN-02
6. Gruber T, Rummel R, Abrikosov O, van Hees R (2010) GOCE level 2 product data handbook, No 4.2. GO-MA-HPF-GS-0110
7. Pavlis NK, Holmes SA, Kenyon SC, Factor JK (2008) An earth gravitational model to degree 2160: EGM2008. In: EGU General Assembly, pp 13–18
8. Lyard F, Lefevre F, Letellier T, Francis O (2006) Modelling the global ocean tides: modern insights from FES2004. *Ocean Dyn* 56(5–6):394–415
9. Sitarski G (2002) Warsaw ephemeris of the solar system: DE405/WAW. *Acta Astronomica* 52:471–486
10. Bizouard C, Gambis D (2009) The combined solution C04 for Earth orientation parameters consistent with international terrestrial reference frame 2005. In: *Geodetic reference frames*. Springer Berlin, pp 265–270
11. Moe K, Moe MM (2005) Gas–surface interactions and satellite drag coefficients. *Planet Space Sci* 53(8):793–801
12. Oliver M, Eberhard G (2000) *Satellite orbits: models, methods and applications*. Springer, Berlin
13. Wong T et al (2000) Clouds and the earth’s radiant energy system (CERES) validation plan erbe-like averaging to monthly toa fluxes (SUBSYSTEM 3.0)
14. Drob DP, Emmert JT, Crowley G, Picone JM, Shepherd GG, Skinner W, Vincent RA (2008) An empirical model of the Earth’s horizontal wind fields: HWM07. *J Geophys Res: Space Phys* (1978–2012): 113 (A12)
15. Sentman LH (1961) Free molecule flow theory and its application to the determination of aerodynamic forces (No. LMSC-448514). Lockheed missiles and space Co Inc Sunnyvale, CA
16. Picone J et al (2002) NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *J Geophys Res: Space Phys* (1978–2012) 107(A12): SIA 15-11–SIA 15-16
17. Bowman BR, Tobiska WK, Marcos FA, Huang CY, Lin CS, Burke WJ (2008) A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices. In: *AIAA/AAS astrodynamics specialist conference and exhibit*, Honolulu, Hawaii

Chapter 10

Dynamic GPS Precise Point Positioning for Deformation Monitoring Using Prior Information

Zhiping Liu, Ziqiang Zhao and Qiuzhao Zhang

Abstract The single epoch deformation monitoring method based on the relative positioning principle was analyzed and its limitations were investigated. This paper presents an intensive study on the deformation features of monitoring objects using a deformation transformation matrix from the rectangular space to topocentric Cartesian coordinates and proposes a dynamic PPP deformation estimation method that includes prior information regarding the deformation features of the monitoring objects. The method was applied to a case study using high-rate data from four international GNSS service (IGS) sites located in a magnitude-9.0 earthquake zone in Japan. The results showed an effective improvement on the precision and convergence of the deformation solution and were in agreement with current results, proving that our method is effective and reliable.

Keywords GPS · Dynamic precise point positioning · Prior information · Deformation monitoring · Earthquakes

10.1 Introduction

Currently, GPS data is applied to several deformation monitoring studies (e.g., surface subsidence in urban areas and subsistence of highway subgrades and high slopes). Nowadays, the extraction of GPS deformation information generally involves the single epoch deformation monitoring method based on the relative positioning principle [1]. However, this method cannot be applied to the dynamic

Z. Liu (✉) · Q. Zhang
School of Environment Science and Spatial Informatics,
China University of Mining and Technology, Xuzhou 221116, China
e-mail: zhpliu@cumt.edu.cn

Z. Zhao
Shanghai Geotechnical Investigations and Design Institute Co., Ltd,
Shanghai 200032, China

large deformation monitoring of long baselines [2, 3] since long bases result in less common visible satellites with the consequent failure of the deformation estimation. In addition, to avoid an ambiguous solution and cycle slip detection, a range restriction to the actual deformation must be applied. Unfortunately, the above requirements are difficult to meet for the deformation monitoring of mining pits, long-span bridges, high-rise buildings, and earthquake displacement.

On March 11, 2011 (05:46:23 UTC), a violent earthquake of magnitude 9.00 at a depth of 20 km [4] occurred 130 km east of the Sendai Port in Honshu, Japan. After the earthquake, several organizations, including the Jet Propulsion Laboratory (JPL) and the Geographical Survey Institute of Japan, quickly provided the coseismic displacement and inverted the dislocation distribution using high-frequency observation data from GPS monitoring stations in Japan [5]. Abundant research has focused on the surface deformation features caused by the earthquake in both Japan and surrounding areas [6–8].

This paper focuses on the deformation monitoring method based on precise point positioning (PPP). First, we introduced the dynamic PPP estimation method and used it to analyze the prior deformation features of four types of monitoring objects according to their characteristics. A new dynamic PPP deformation monitoring method using the prior information was then suggested, for which high-frequency data was selected from four international GNSS service (IGS) stations (DAEJ, MIZU, SHAO, and USUD) to perform a deformation calculation and analysis and verify the validity and effectiveness of the method.

10.2 Dynamic GPS Precise Point Positioning

10.2.1 The Zero-Difference Observation Model

The equations for the pseudo-range and phase zero-difference observations were:

$$P_{i,f}^j = \rho_i^j - c(dt_i + B_{i,f_r}) + c(dt^j + B_{f_p}^j) + I_{i,f}^j + T_i^j + re_i^j + \varepsilon_{i,f_p}^j \quad (10.1)$$

$$\Phi_{i,f}^j = \rho_i^j - c(dt_i + B_{i,f_\phi}) + c(dt^j + B_{f_\phi}^j) + \lambda \cdot N_f^j - I_{i,f}^j + T_i^j + re^j + \varepsilon_{i,f_\phi}^j \quad (10.2)$$

where $P_{i,f}^j$ is the pseudorange, $\Phi_{i,f}^j$ is the carrier phase (unit: m), ρ_i^j is the geometrical distance between the satellite and the center of the receiver, (see Eq. (10.3)), c is the speed of light, f is the signal frequency of the receiver, and λ is the wavelength of the corresponding frequency f . For Eq. (10.2), N_f^j represents the ambiguity of the whole cycles, dt_i and dt^j are the clock corrections for the receiver and satellite, respectively, B_i and B^j are the corrections for the differential code bias for the receiver and satellite, $I_{i,f}^j$ is the correction for the ionosphere, T_i^j is the correction for the neutral atmosphere

delay, re_i^j is the correction for the equivalent range delay due to the relativistic effect of the j th satellite, and ε_i^j represents the observation noise.

The expression of ρ_i^j in Eqs. (10.1–10.2) was rewritten as:

$$\begin{cases} \rho_i^j = \|\mathbf{X}^j - \mathbf{X}_i\| \\ \mathbf{X}^j = \mathbf{x}^j + \mathbf{pc}^j + \mathbf{er}^j \\ \mathbf{X}_i = \mathbf{x}_{i,0} + \mathbf{pc}_i + \mathbf{el}_i + \mathbf{ol}_i \end{cases} \quad (10.3)$$

where $\mathbf{X}_i = (X_i, Y_i, Z_i)$ is the corrected coordinate of the observation station, $\mathbf{X}^j = (X^j, Y^j, Z^j)$ is the corrected coordinate of satellite j , $\mathbf{x}^j = (x^j, y^j, z^j)$ is the barycentric coordinate of satellite j , $\mathbf{x}_{i,0} = (x_{i,0}, y_{i,0}, z_{i,0})$ is the coordinate of the observation station i , \mathbf{pc}^j and \mathbf{pc}_i are the corrections for the phase center of the satellite and receiver antennas, respectively, \mathbf{er}^j is the correction for the effect of the earth's rotation, \mathbf{el}_i is the correction for the terrestrial tide, and \mathbf{ol}_i is the correction for the sea tide.

10.2.2 Dynamic PPP Estimation Method

We developed a dynamic PPP model [9] using high-precision IGS products and GPS dual-frequency iono-free combination observations and correcting for the satellite phase center, relativistic effect, earth's rotation, zenith tropospheric delay, and terrestrial and sea tide effects. The coordinate correction parameters for the observation stations, the clock correction of the receiver, the residual delay of the wet atmosphere zenith, and the epoch-by-epoch ambiguity also had to be estimated using successive adjustment methods based on the satellite elevation fixed weight model [1, 10]. The authority downgrading had to be conducted on a cycle slip of the carrier phase and the interpolation of the satellite orbit based on the receiving time of the signal had to be revised during signal transmission, unlike for the interpolation of the clock correction. The IGS precise clock correction products were obtained from dual-frequency iono-free combination observations, so that no hardware delay correction was needed [10].

For the parameter estimation in the dynamic PPP, the correction for the observation station's coordinates and clock, and the residual delay of the wet atmosphere zenith were time-varying parameters, while the ambiguity parameters were constant and treated as time-unvarying parameters in the absence of lock-loss or cycle slips. Setting \mathbf{x}_k as the time-unvarying parameter, \mathbf{y}_k as the time-varying parameter, and $\mathbf{L}_k, \mathbf{P}_k$ as the observation and weight matrix of the k th epoch, the adjustment model was expressed as:

$$\mathbf{V}_k = \mathbf{A}_k \cdot \underset{u_1 \times 1}{\mathbf{y}_k} + \mathbf{B}_k \cdot \underset{u_2 \times 1}{\mathbf{x}_k} - \mathbf{L}_k \quad (10.4)$$

where $\mathbf{y}_k = [dx \ dy \ dz \ c \cdot dt \ dT_{wet}]^T$, $\mathbf{x}_k = [N^I \ \dots \ N^s]^T$, and $\mathbf{A}_k, \mathbf{B}_k$ represent the corresponding design matrices, with dx, dy, dz as the correction parameters for the observation station's coordinates, $c \cdot dt$ as the clock correction of the receiver (equivalent distance), dT_{wet} as the residual delay of the wet atmosphere zenith, and N as the ambiguity of the iono-free combination.

Setting $\hat{\mathbf{x}}_{k-1}, \mathbf{Q}_{\hat{\mathbf{x}}_{k-1}}$ as the time-unvarying parameter estimation and the association factor matrix after adjustment in one cast for the first $k-1$ epochs, we obtained the following fictitious observation equation:

$$\bar{\mathbf{V}}_{\hat{\mathbf{x}}_{k-1}} = \mathbf{x}_k - \hat{\mathbf{x}}_{k-1} \quad (10.5)$$

Thus, considering the observation information of the k th epoch and the estimation information $\hat{\mathbf{x}}_{k-1}, \mathbf{Q}_{\hat{\mathbf{x}}_{k-1}}^{-1}$ of the time-unvarying parameter for the first $k-1$ epochs, according to the generalized adjustment principles, we obtained:

$$\mathbf{V}_k^T \mathbf{P}_k \mathbf{V}_k + \bar{\mathbf{V}}_{\hat{\mathbf{x}}_{k-1}}^T \mathbf{Q}_{\hat{\mathbf{x}}_{k-1}}^{-1} \bar{\mathbf{V}}_{\hat{\mathbf{x}}_{k-1}} = \min \quad (10.6)$$

Using the generalized adjustment, the following equation for the adjustment parameters was obtained:

$$\begin{bmatrix} N_{a,k} & N_{ab,k} \\ N_{ba,k} & \bar{N}_{b,k} \end{bmatrix} \begin{pmatrix} \mathbf{y}_k \\ \mathbf{x}_k \end{pmatrix} = \begin{pmatrix} \mathbf{L}_{a,k} \\ \bar{\mathbf{L}}_{b,k} \end{pmatrix} \quad (10.7)$$

where $N_{a,k} = \mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k$, $N_{b,k} = \mathbf{B}_k^T \mathbf{P}_k \mathbf{B}_k$, $N_{ab,k} = \mathbf{A}_k^T \mathbf{P}_k \mathbf{B}_k$, $N_{ba,k} = \mathbf{B}_k^T \mathbf{P}_k \mathbf{A}_k$ and $\mathbf{L}_{a,k} = \mathbf{A}_k^T \mathbf{P}_k \mathbf{L}_k$, $\mathbf{L}_{b,k} = \mathbf{B}_k^T \mathbf{P}_k \mathbf{L}_k$, $\bar{\mathbf{L}}_{b,k} = \mathbf{L}_{b,k} + \mathbf{Q}_{\hat{\mathbf{x}}_{k-1}}^{-1} \hat{\mathbf{x}}_{k-1}$, $\bar{N}_{b,k} = N_{b,k} + \mathbf{Q}_{\hat{\mathbf{x}}_{k-1}}^{-1}$.

The sequential adjustment estimation [1] was obtained by solving the time-varying parameters $\hat{\mathbf{y}}_k$ by back substitution after solving the time-unvarying parameters $\hat{\mathbf{x}}_k$ or by solving the time-unvarying parameters $\hat{\mathbf{x}}_k$ by back substitution after solving the time-varying parameters $\hat{\mathbf{y}}_k$. In addition, the variation in the ambiguity parameters must be processed (i.e., increased or decreased) because of the change in satellite attitude and receiver lock-loss, among other factors.

10.3 Dynamic PPP Deformation Monitoring Using Prior Information

The deformation monitoring needs accurate initial coordinates of the observation stations, which can be solved by static PPP for follow-up data processing and later obtention of the coordinate correction or deformation time series.

The current GPS single epoch deformation monitoring method estimates first the deformation under the WGS84 coordinate system, followed by its conversion to the topocentric Cartesian coordinate system, which is easier for deformation explanation. However, it does not include the prior information of the deformation features

in the data processing. Thus, to unify the relationship between deformation estimation and explanation, a transformation expression of the deformation or correction under the topocentric Cartesian and rectangular space coordinate systems must be obtained first [1, 11]:

$$\begin{pmatrix} dN \\ dE \\ dU \end{pmatrix} = \begin{bmatrix} -\sin \bar{\varphi} \cos \bar{\lambda} & -\sin \bar{\varphi} \sin \bar{\lambda} & \cos \bar{\varphi} \\ -\sin \bar{\lambda} & \cos \bar{\lambda} & 0 \\ \cos \bar{\varphi} \cos \bar{\lambda} & \cos \bar{\varphi} \sin \bar{\lambda} & \sin \bar{\varphi} \end{bmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} \quad (10.8)$$

where $(dx, dy, dz)^T$ is the deformation or correction of the rectangular space coordinate system, $(dN, dE, dU)^T$ is the deformation or correction of the topocentric Cartesian coordinate system, and $\bar{\varphi}, \bar{\lambda}$ are the latitude and longitude of the observation stations, respectively, which can be calculated using the approximate coordinates of the observation stations.

In deformation monitoring, such as for mining pits and engineering slopes, it is usually easy to obtain the azimuth tendency after the deformation from the variation of the coordinates. Then, the following equation, including the known deformation azimuth angle, is obtained according to Eq. (10.8):

$$\sin \alpha \cdot dN - \cos \alpha \cdot dE = V_\alpha \quad (10.9)$$

where α represents the azimuth angle of the station deformation, $V_\alpha \in Norm\{0, \sigma_\alpha^2\}$.

For high and large-span bridges (large linear structures), the deformation features manifest as distinct horizontal hunting and elevation vibration, with almost no deformation in the direction of the bridge path. The prior information is then given by the following equation:

$$\cos \beta \cdot dN + \sin \beta \cdot dE = V_\beta \quad (10.10)$$

where β represents the azimuth angle along the bridge path, $V_\beta \in Norm\{0, \sigma_\beta^2\}$.

For the monitoring of surface subsidence in urban areas and subsistence of highway subgrades, the deformation features manifest as distinct surface subsidence, with almost no changes in the horizontal direction (i.e., $dN \approx 0, dE \approx 0$). Then, using Eq. (10.8), the restraint equation can be expressed as:

$$\begin{cases} dN = V_{dN} \\ dE = V_{dE} \end{cases} \quad (10.11)$$

where $V_{dN} \in Norm\{0, \sigma_{dN}^2\}, V_{dE} \in Norm\{0, \sigma_{dE}^2\}$.

For the monitoring of plate or crustal movement and high-rise buildings, the deformation usually manifests as horizontal deformation, while the approximate elevation of the observation stations remains almost unchanged (i.e., $dU \approx 0$). Then, the following restraint equation can be obtained based on Eq. (10.8):

$$dU = V_{dU} \quad (10.12)$$

where $V_{dU} \in Norm\{0, \sigma_{dU}^2\}$.

In addition, we can obtain a more accurate residual delay of the wet atmosphere zenith in the time-varying parameter x_k by a random walk method [12] as:

$$dT_{wet} = V_{wet} \quad (10.13)$$

where $V_{wet} \in Norm\{0, \sigma_{wet}^2\}$.

The restraint equation including Eqs. (10.9–10.12) and (10.13) can be expressed as:

$$C_k \cdot y_k = \bar{V}_{y_k}, \quad Q_{\bar{V}_{y_k}} \quad (10.14)$$

where $Q_{\bar{V}_{y_k}} = diag\{\sigma_\alpha^2, \sigma_\beta^2, \sigma_{dN}^2, \sigma_{dE}^2, \sigma_{dU}^2, \sigma_{wet}^2\}$ and C_k is obtained from Eqs. (10.8–10.12).

In conclusion, combining Eqs. (10.4), (10.5) and (10.14), the dynamic PPP deformation monitoring method with prior information for the time-varying parameters can be obtained and expressed by the following equation based on a generalized adjustment:

$$\begin{bmatrix} \bar{N}_{a,k} & N_{ab,k} \\ N_{ba,k} & \bar{N}_{b,k} \end{bmatrix} \begin{pmatrix} y_k \\ x_k \end{pmatrix} = \begin{pmatrix} L_{a,k} \\ \bar{L}_{b,k} \end{pmatrix} \quad (10.15)$$

where $\bar{N}_{a,k} = N_{a,k} + C_k^T Q_{\bar{V}_{y_k}}^{-1} C_k$.

The sequential adjustment estimation can be obtained by solving the time-varying parameters \hat{y}_k by back substitution after solving the time-unvarying parameters \hat{x}_k as:

$$\begin{cases} \hat{x}_k = Q_{\hat{x}_k} \left(\bar{L}_{b,k} - N_{ba,k} N_{a,k}^{-1} L_{a,k} \right) \\ Q_{\hat{x}_k}^{-1} = \bar{N}_{b,k} - N_{ba,k} \bar{N}_{a,k}^{-1} N_{ab,k} \end{cases} \quad (10.16)$$

$$\begin{cases} \hat{y}_k = \bar{N}_{a,k}^{-1} \left(L_{a,k} - N_{ab,k} \hat{x}_k \right) \\ Q_{\hat{y}_k}^{-1} = \bar{N}_{a,k}^{-1} + \bar{N}_{a,k}^{-1} N_{ab,k} Q_{\hat{x}_k}^{-1} N_{ba,k} \bar{N}_{a,k}^{-1} \end{cases} \quad (10.17)$$

or by solving the time-unvarying parameters \hat{x}_k by back substitution after solving the time-varying parameters \hat{y}_k as:

$$\begin{cases} \hat{y}_k = Q_{\hat{y}_k} \left(L_{a,k} - N_{ab,k} \bar{N}_{b,k}^{-1} \bar{L}_{b,k} \right) \\ Q_{\hat{y}_k}^{-1} = \bar{N}_{a,k} - N_{ab,k} \bar{N}_{b,k}^{-1} N_{ba,k} \end{cases} \quad (10.18)$$

$$\begin{cases} \hat{x}_k = \bar{N}_{b,k}^{-1}(\bar{L}_{b,k} - N_{ba,k}\hat{y}) \\ Q_{\hat{x}_k} = \bar{N}_{b,k}^{-1} + \bar{N}_{b,k}^{-1}N_{ba,k}Q_{\hat{y}_k}N_{ab,k}\bar{N}_{b,k}^{-1} \end{cases} \quad (10.19)$$

Nevertheless, the dynamic PPP deformation monitoring method with prior information proposed in this paper is designed for certain practices and not all requirements from Eqs. (10.9–10.12) can be met in actual situations. In most cases, only one or certain restraint equations are selected. Appropriate constraint conditions combined with engineering practices must then be chosen for each specific practical application.

10.4 Results and Discussion

To validate the method proposed in this paper, high-frequency GPS (00:00–07:30 UTC, from March 11, 2011) and precise IGS ephemeris data were selected for four IGS stations (DAEJ, MIZU, SHAO, and USUD), located at epicentral distances of about 140, 450, 1362, and 2088 km, respectively (Fig. 10.1) from the magnitude-9.0 earthquake that occurred in Japan on March 11, 2011 (05:46:23 UTC). The movement of the stations during the earthquake and the recovery from deformation for short periods were calculated and analyzed.

The deformation for all stations was calculated using the dynamic PPP method (Fig. 10.2). To clearly show the deformation results for all stations, the results for USUD, DAJE, and SHAO were moved -1, +1, and +2 m, respectively.

Previous research has shown that two main types of waves: P and S, are generated during an earthquake. P-waves are relatively less destructive than S-waves, which have a great tangential destructive power, causing horizontal warping and extrusion of the crust. The horizontal displacement caused by an earthquake is greater than the vertical, which mostly returns to its initial state. Thus, the prior information was added to the elevation direction using Eq. (10.12) and the

Fig. 10.1 Map of the earthquake epicenter and IGS stations

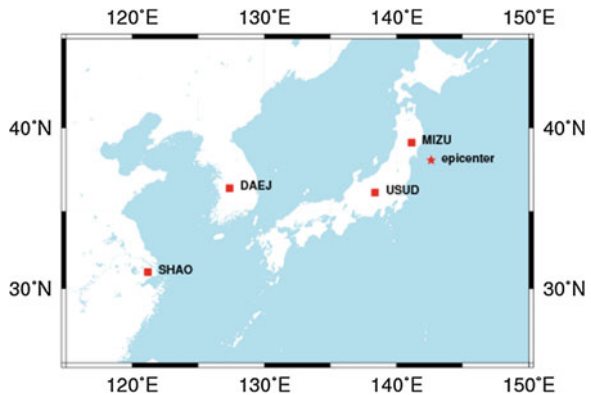
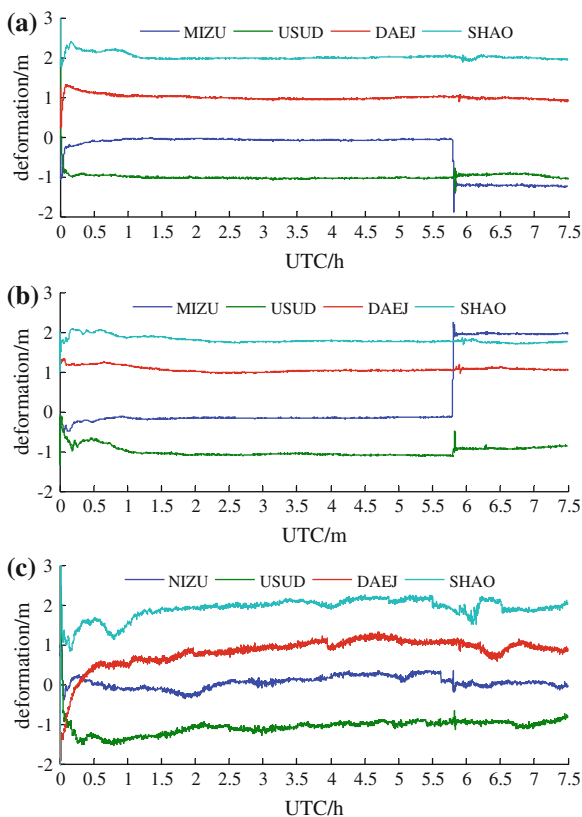


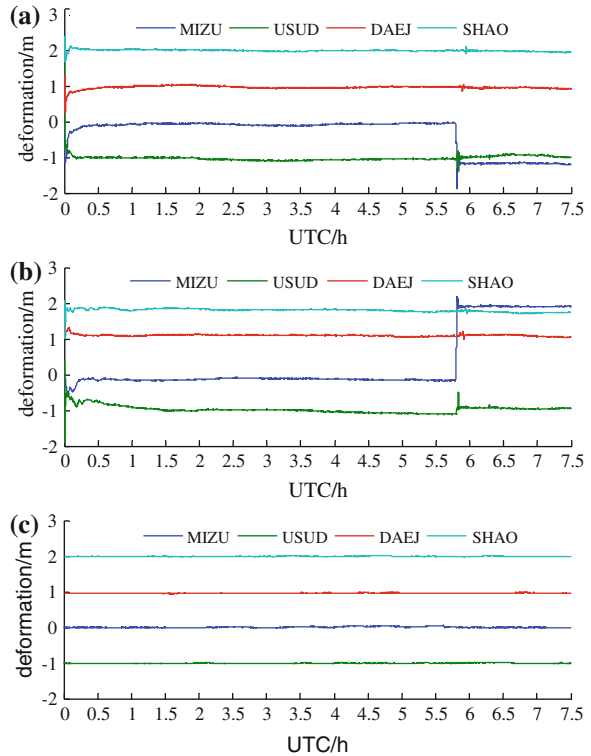
Fig. 10.2 Deformations calculated using the dynamic precise point positioning (PPP) for four international GNSS service (IGS) stations. **a** Deformation in the north direction. **b** Deformation in the east direction. **c** Deformation in the vertical direction



deformation estimation results using the dynamic PPP with prior information are shown in Fig. 10.3.

The north direction deformation time series (Fig. 10.2a) showed that for MIZU, which was closest to the earthquake epicenter, the maximum instant southward displacement of 1.61 m and the permanent deformation of 1.20 m was generated after the earthquake. Oppositely, USUD was deformed instantly during the earthquake, with a maximum instant southward displacement of 0.27 m, which mostly recovered after the earthquake. The stations of DAEJ and SHAO, located further from the epicenter, showed small deformations during the earthquake and suffered a very small impact. The deformation time series in the east direction (Fig. 10.2b) revealed a maximum instant eastward displacement for MIZU of about 2.76 m, with a remaining permanent deformation of up to 2.10 m. The maximum eastern displacement for USUD was 0.57 m, with a permanent translation of 0.15 m, while DAJE and SHAO had very little deformation during the earthquake and nearly completely recovered after it. According to the deformation time series (Fig. 10.2c), the deformation estimation errors for the four stations on the vertical direction were much bigger than for the north and east directions, with some low-frequency

Fig. 10.3 Deformations calculated using the dynamic precise point positioning (PPP) with prior information for four international GNSS service (IGS) stations. **a** Deformation in the north direction. **b** Deformation in the east direction. **c** Deformation in the vertical direction



deformation. The analyses suggested that this was mainly caused by imperfect data processing procedures such as the cycle slip detection or the low geometric strength of observation models.

The deformation time series for the north direction (Fig. 10.2a) showed that station MIZU was the most affected by the earthquake, with a maximum instant southward displacement of 1.65 m and a permanent deformation of 1.17 m after the earthquake. The arrival time of the seismic waves was 05:47:11 as inversed from the deformation series. For USUD, the maximum southward displacement was 0.22 m, with a northward permanent deformation of 0.05 m and an arrival time of the seismic waves of 05:48:34. However, DAJE and SHAO were only displaced for 0.02–0.03 m during the earthquake, fully recovering in a very short period, with almost no permanent deformation. The arrival time of their seismic waves was 05:52:20 and 05:54:58, respectively. The deformation time series in the east direction (Fig. 10.3b) showed a maximum instant eastward displacement for MIZU of 2.87 m, with a permanent deformation of 2.06 m after the earthquake. For USUD, the maximum displacement was 0.62 m, with a permanent deformation of 0.16 m. DAJE and SHAO showed very little deformation during the earthquake, fully recovering in a very short period. The vertical deformation time series

(Fig. 10.3c) revealed a small subsidence within 0.10 m for all stations, representing only a slight effect of the earthquake.

Comparing Figs. 10.2 and 10.3, we see that the maximum deformation difference for MIZU in the north direction was 0.04 m, with a permanent deformation difference of 0.03 m, while for the east direction, the maximum deformation difference was 0.11 m, with a permanent deformation difference of 0.04 m. For USUD, the difference in the maximum deformation in the north direction was 0.05 m, with a permanent difference of 0.05 m, while for the east direction, the maximum difference was 0.05 m, with a permanent difference of 0.03 m. Thus, the absolute difference in the maximum horizontal deformation between these two methods was 0.04–0.11 m, while the absolute difference in the permanent displacement was 0.03–0.05 m. In relation to the total deformation, the relative difference between the two results was 1–5 %. However, for the vertical direction (Fig. 10.2), the results did not agree with the fact that earthquake waves spread successively from the smaller effect and shorter distance from the earthquake epicenter. Oppositely, the results including the prior information (Fig. 10.3) were consistent with previous work [4, 8, 13]. In addition, the dynamic PPP deformation monitoring method using prior information improved the precision and convergence of the deformation solution for the north, east, and vertical directions.

According to the arrival time of the seismic waves and distance from the epicenter, the transmission speed of the seismic waves obtained by inversion was 3.1–4.1 km/s, in agreement with Liu et al. [4]. In addition, since MIZU and USUD were the stations most affected by the earthquake and had the largest deformation, to better represent their horizontal movement during the earthquake, the plane movement track of the deformation time series for these two stations was shown for 05:46:00–05:51:00 UTC. Figure 10.4 clearly shows the plane movement for these two stations, indicating that the crustal movement caused by the earthquake was characterized by deformation resilience. The closeness to the earthquake epicenter resulted in a higher resilience proportion, in agreement with Zhang et al. [13].

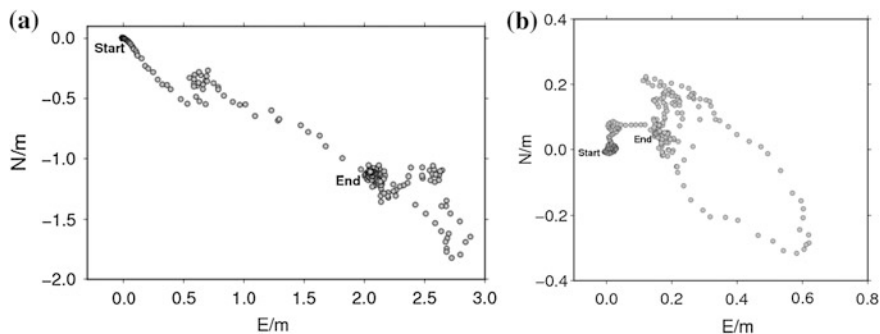


Fig. 10.4 Motion trajectory in the horizontal direction for the international GNSS service (IGS) stations MIZU **a** and USUD **b** for 05:46:00–05:51:00 UTC

10.5 Conclusion

This paper highlights that the current single epoch deformation-monitoring methods are difficult to apply in large dynamic deformation estimations of long baselines. The deformation estimation method based on dynamic precise point positioning was presented. This method involved the construction of the deformation transformation matrix from the rectangular space to the topocentric Cartesian coordinate system, which is easier for deformation explanation, followed by the unification of the deformation estimation and explanation standards to better consider the deformation features during data processing. In addition, we obtained prior deformation features for four types of monitoring objects by combining engineering mechanics, engineering geology, and engineering practices and proposed a dynamic PPP deformation monitoring method using this prior information. The new method was applied to the calculation and analysis of the dynamic deformation of four Japanese stations due to the 2011 magnitude-9.0 earthquake that occurred in Japan. The obtained deformation features were consistent with current results, confirming a new reliable method for the estimation of large dynamic deformation for long baselines.

Acknowledgments This work was supported by the NSFC (No. 41204011), SKLGED (No. 2014-3-2-E), and SZBF (No. 2011-6-B35) projects.

References

1. Liu Z (2014) GNSS slope monitoring and deformation analysis. Surveying and Mapping Press, Beijing
2. Liu Z, He X et al (2011a) GPS single-epoch deformation monitoring method using sequential adjustment and carrier phase combinations. *J Geomat Sci Technol* 28(2):P88–P93
3. Liu Z, He X et al (2011b) Dynamic triple-difference method for single frequency GPS deformation monitoring. *J Tongji Univ Nat Sci* 39(7):P1074–P1078
4. Liu Y, Fan S, Qian F et al (2012) GPS responses to the 2011 MW9.0 earthquake in Japan. *Acta Oceanologica Sinica* (6):P59–P65
5. Shao Z, Wu Y, Jiang Z et al (2011) The analysis of coseismic slip and near-field deformation about Japanese 9.0 earthquake based on the GPS observation. *Chinese J Geophys* 54(9): P2243–P2249
6. Zhao G, Li P (2012) Coseismic and pre-seismic displacements observed by fiducial stations in China caused by the 2011 Japan M9.0 earthquake. *Earthquake* (2):P129–P134
7. Zhang F, Zhan W, Sun D (2012a) Analysis of GPS baseline time series before and after Tohoku-oki M9.0 earthquake in Japan in 2011. *J Seismol Res* (2):P190–P200
8. Shen F, Li J, Guo F (2012) Analysis of the coseismic displacements along Eastern Coast of China caused by the 2011 Tohoku-Oki earthquake using epoch-by-epoch PPP. *Geomat Inf Sci Wuhan Univ* (11):P1345–P1347
9. Kouba J (2009) A guide to using international GNSS service (IGS) products [DB/OL]. <http://igscb.jpl.nasa.gov/igscb/resource/pubs/UsingIGSProductsVer21.pdf>. Accessed 10 Jan 2011

10. Ziqiang Zhao (2013) Research and implementation of precise barometer altitude augmented GPS dynamic precise point positioning. China University of Mining and Technology, Xuzhou
11. Liu Z, Yu Q, Zha J (2015) Fast coordinate transformations for both XYZ-BLH and XYZ-RhA. *Sci Surv Mapp* (under publishing)
12. Ge M, Liu J (1996) The estimation methods for tropospheric delays in global positioning system. *Acta Geodaetica Et Cartographica Sinica* 25(4):P285–P291
13. Zhang X, Guo F, Guo B et al (2012b) Coseismic displacement monitoring and wave picking with high-frequency GPS. *Chinese J Geophys* 55(6):P1912–P1918

Chapter 11

Positioning Accuracy Analysis of Beidou Continuous Operation Tracking Stations

Xiangxin Guo, Qile Zhao, Shenghua Jiang and Min Li

Abstract Based on PANDA software, GPS and Beidou observation data of four Beidou continuous operation tracking network (iGMAS/BETS/CMONOC/MGEX) are calculated in precise point positioning mode, while GPS and Beidou positioning accuracy and observation residuals of each tracking station are analyzed, which can provide reference for the establishment of centimeter-level reference coordinate frame using Beidou system. The results are as follows: single day positioning repeatability with Beidou are 11.06, 7.33, 30.13 mm in E, N, U directions respectively; the STD of differences between GPS and Beidou are 11.36, 7.43, 31.73 mm in E, N, U directions, which indicates that Beidou system can achieve centimeter-level precision positioning at service area; the average of differences between GPS and Beidou are 4.87, 2.65, 29.38 mm in E, N, U directions, which demonstrates an obvious systematic bias of precision positioning between GPS and Beidou; according to tracking network, GPS and Beidou positioning accuracy and residuals are statistically analyzed, which shows that CMONOC tracking network has the best positioning accuracy, while MGEX and BETS tracking network are worse and iGMAS tracking network is much lower.

Keywords Beidou navigation satellite system · Precise point positioning · iGMAS/BETS/CMONOC/MGEX

11.1 Introduction

China is implementing the BeiDou Navigation Satellite System (referred to as “Beidou System”) independently. The initial satellite navigation system was completed in December 27, 2012. It contains 4 medium earth orbit satellites,

X. Guo (✉) · Q. Zhao · M. Li
GNSS Research Center of Wuhan University, 129 Luoyu Road,
Wuhan 430079, Hubei, China
e-mail: guoxx@whu.edu.cn

S. Jiang
Wuhan Geomatic Institute, 209 Wansongyuan Road, Wuhan 430022, Hubei, China

5 inclined geosynchronous orbit satellites and 5 geostationary orbit satellites in orbit and provides coverage in the Asia-Pacific region (55°S–55°N, 70°E–150°E) with positioning, navigation, timing, and short-message communication service capabilities. The system is designed to provide global coverage around 2020 [1, 2]. Many scholars have verified Beidou system's basic navigation and positioning service performance in the Asia-Pacific region [3–6]. Its satellite visibility and DOP can meet the needs of different users. Pseudorange and carrier phase measurement accuracy is equivalent to the US GPS system. MEO and IGSO satellites orbit determination accuracy are better than 10 cm. Precise point positioning accuracy in horizontal direction and vertical direction is better than 1 and 3 cm respectively.

Nowadays there are four Beidou high precise observation network operating steadily in the world. For monitoring and evaluating the four systems (GPS/GLONASS/Beidou/Galileo) operation condition and key performance indicators, China has established an international GNSS Monitoring and Assessment System (iGMAS), including eight domestic and four overseas tracking stations. To further enhance the theory and application of the Beidou system, Wuhan University emplaced 20 Beidou Experimental Tracking Stations (BETS) worldwide. Crustal movement observation network of China (referred to as “state land network”, CMONOC) also has 26 base stations to upgrade to GPS/Beidou dual mode receiver. For all-round development of multimode GNSS technology research, IGS globally deployed 114 MGEX (Multi GNSS Experiment) tracking stations, 30 tracking stations among having the ability to receive Beidou data [7–10].

In this paper, GPS and Beidou observation data taken from the Asia-Pacific region iGMAS/BETS/CMONOC/MGEX tracking stations was used to non-difference pattern Precise Point Positioning (PPP) based on the integrated satellite navigation data processing software (Position And Navigation Data Analysis, PANDA) developed independently by Wuhan University, comparing the four network tracking station coordinates, analyzing Beidou continuous operation tracking stations positioning accuracy, verifying Beidou wide range positioning levels, providing reference for establishing centimeter coordinate reference frame in the use of Beidou system, laying the foundation for regional high-precision positioning, navigation and timing services.

11.2 Data Sources

This paper adopts GPS and Beidou observation of Beidou continuous operation tracking stations in Beidou system service area, whose 50-day observation time ranging from 8 September to 27 October, 2014 (day of year from 251 to 300, 2014). The distribution of Beidou continuous operation tracking stations is shown in Fig. 11.1, divided by iGMAS/BETS/CMONOC/MGEX tracking network, where number of iGMAS tracking station is 6, number of BETS for 7 stations, number of

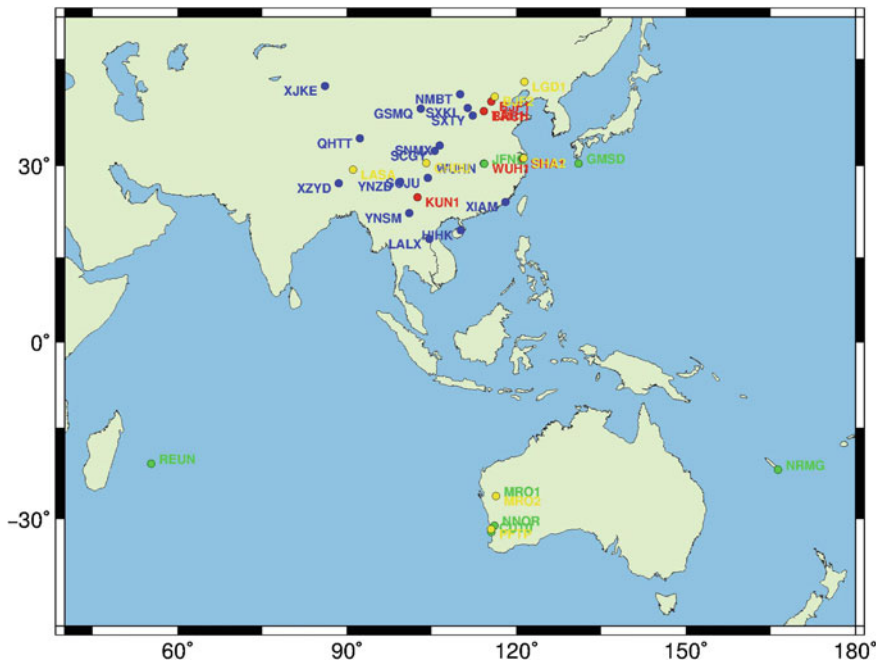


Fig. 11.1 Distribution of Beidou continuous operation tracking stations (*red* represents for iGMAS, *yellow* represents for BETS, *blue* represents for CMONOC, *green* represents for MGEX)

CMONOC tracking station is 16, and number of MGEX tracking station is 7. Each of tracking station information is shown in Table 11.1, which contains the type of receiver stations and antenna types.

11.3 Positioning Solution Strategy

In this paper, 50 days of GPS/Beidou observation data from DOY 251 to 300 in 2014 taken from iGMAS/BETS/CMONOC/MGEX tracking station was calculated in the method of non-difference pattern PPP with the PANDA software, independently developed by Wuhan University. Mathematical model of the geometry information, the signal dissemination of atmospheric information, astronomy and geodynamic information was established. The model error detection, sequential least squares and nonlinear state estimation were adopted in data analysis and processing, eventually obtaining massive ground tracking station network precise coordinates [11–13]. GPS/Beidou PPP solution strategy is shown in Table 11.2, and the final products of Beidou precise orbit and clock error products released by

Table 11.1 Information of Beidou continuous operation tracking stations

Network	Station	Receiver type	Antenna type	Station	Receiver type	Antenna type
iGMAS	BRCH	CETC-54 GMR-4011	LEIAR25.R4 LEIT	TAH1	CETC-54 GMR-4011	LEIAR25.R4 LEIT
	KUN1	UNICORE UB4B0I	NOV750.R4 NOVS	WUH1		
	SHA1			BJF1		
BETS	CHDU	UNICORE UR240	HX-BS483A	SHA2	TRIMBLE NETR9	TRM59900.00 NONE
	LASA			LGD1		
	BJF2			MRO2		
	PFTP					
CMONOC	GSMQ	TRIMBLE NETR9	TRM59900.00 SCIS	SXKL	TRIMBLE NETR9	TRM59800.00 SCIS
	HIHK			SCGY		
	LALX			WUHN		TRM59900.00 SCIS
	NMBT			XIAM		
	QHTT			XJKE		
	SXTY			XZYD		
	SCJU			YNSM		
	SNMX			YNZD		
	MGEX			CUT0		
GMSD		NRMG	TRM57971.00 TZGD			
JFNG		TRM59800.00 NONE	NNOR	SEPT POLARX4	SEPCHOKE_MC NONE	
MRO1		TRM59900.00 NONE				

Wuhan University are employed. Its accuracy can meet the needs of centimeter-level positioning [14, 15]. Note that the antenna phase center variations (PCV) and antenna phase center offset (PCO) of Beidou satellite and receiver has not been calibrated, and it will not be considered.

11.4 Positioning Results Analysis

The above iGMAS/BETS/CMONOC MGEX tracking station daily static solution were calculated to get single day repeatability precision of GPS and Beidou of the tracking station, while standard deviation (STD) of GPS and Beidou difference are shown in Fig. 11.2.

As can be seen from Fig. 11.2, GPS single day repeatability precision is substantially better than 3 mm in the horizontal direction and substantially better than 5 mm in the elevation direction; Beidou single day repeatability precision is substantially better than 1 cm in the horizontal direction and substantially better than 3 cm in the elevation direction; deviation STD difference between GPS and Beidou

Table 11.2 GPS/Beidou PPP solution strategy

	Parameter	Model
Observation	Observation data	No difference non-ionosphere combination (pseudorange and carrier)
	Elevation mask	7
	Antenna phase center correction	GPS: IGS absolute phase center correction; Beidou: Ignored
Error correction	Phase wrapping	GPS: correct; Beidou: Ignored
	Tropospheric model	Saastamoinen + GMF model + Process noise 20 + 2 cm/sqrt (hour)
	Ionospheric correction	Dual-frequency and ionospheric combination observations to eliminate
	Atmosphere loading	Leave out
	Relativistic correction	IERS2010
	Tide correction	Earth tide, sea tide, pole tide (IERS2010)
	Satellite ephemeris	GPS: IGS Precision orbit; Beidou: WHU Precision orbit
	Receiver clock error	process estimate + white noise, priori constraints: 9000 m
	Satellite clock error	GPS: IGS Precision clock; Beidou: WHU Precision clock
	Station coordinate	Parameter estimation

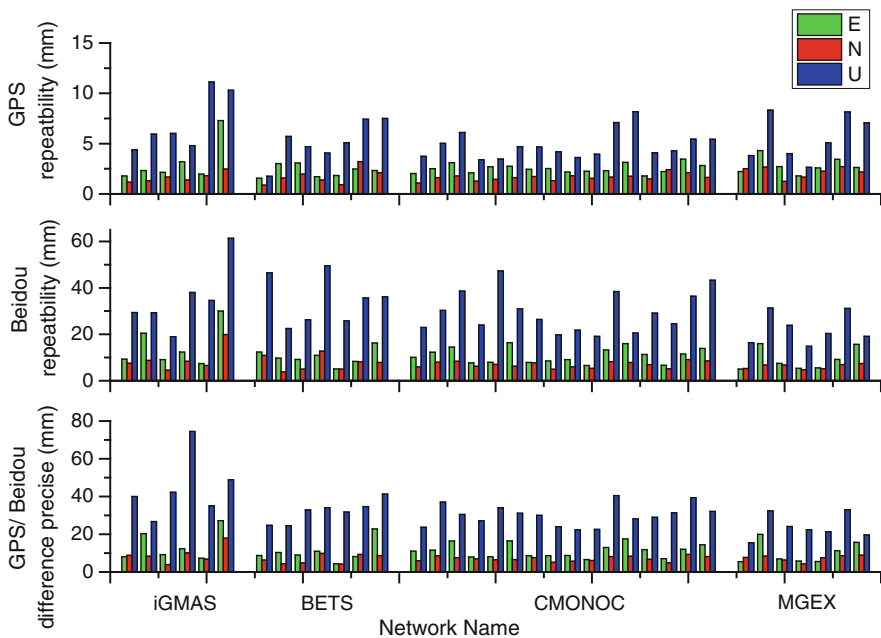


Fig. 11.2 GPS/Beidou repeatability and their difference STD

Table 11.3 GPS/Beidou repeatability

Component	GPS (mm)			Beidou (mm)		
	E	N	U	E	N	U
IGMAS	3.11	1.64	7.09	14.78	9.29	35.27
BETS	2.28	1.72	5.18	10.26	7.65	34.63
CMONOC	2.52	1.64	4.83	10.84	6.97	29.6
MGEX	2.81	2.17	5.58	9.16	6.17	22.44
ALL	2.63	1.76	5.42	11.06	7.33	30.13

is substantially less than 1 cm in the horizontal direction and generally within 3 cm in the elevation direction. The results show that there is a big discrepancy in Beidou horizontal direction and elevation accuracy because at the present stage 14-satellite Beidou system has 5 GEO satellites, which remained relatively immobile ground points and located at the equator. These observations of GEO satellites has weak constraint for elevation direction geometric, whose geometric precision attenuation factor is small, therefore the direction of precision elevation was significantly lower than the horizontal accuracy [16].

According to iGMAS/BETS/CMONOC/MGEX tracking network, GPS/Beidou repeatability and their difference statistics are shown in Tables 11.3 and 11.4.

As can be seen from Table 11.3, GPS positioning precision with Beidou continuous operation tracking stations is 2.63, 1.76 and 5.42 mm in E/N/U direction respectively; Beidou positioning precision is 11.06, 7.33 and 30.13 mm in E/N/U direction respectively, which verifies that Beidou system positioning precise can reach 1 cm in horizontal direction and 3 cm in elevation direction. Comparing positioning precise of four Beidou continuous operation tracking network, precision of CMOMOC network is the best and that of iGMAS network is the worst, while precision of MGEX network is little better than that of BETS network.

As can be seen from Table 11.4, the STD of differences between positioning results with GPS and Beidou is 11.36, 7.43 and 31.73 mm in E/N/U direction respectively. The gap between Beidou and GPS positioning accuracy is mainly due to the poor products of Beidou satellite orbit and clock errors, which currently only reaches 10–20 cm 3-dimensional orbit accuracy [14] with an order of magnitude worse than that of GPS. The average difference between GPS and Beidou in E, N, U three directions is respectively 4.87, 2.65 and 29.38 mm, which indicates the

Table 11.4 statistics of difference between GPS and Beidou

Component	RMS (mm)			STD (mm)			Average (mm)		
	E	N	U	E	N	U	E	N	U
iGMAS	14.91	9.96	49.41	14.06	9.3	44.58	4.96	3.57	21.31
BETS	11.17	7.8	37.44	10.6	6.78	31.97	3.52	3.86	19.49
CMONOC	12.43	7.3	45.04	11.24	7.02	30.18	5.31	2.00	33.43
MGEX	11.22	7.57	39.62	10.07	7.4	24.03	4.95	1.60	31.50
ALL	12.36	7.89	43.24	11.36	7.43	31.73	4.87	2.65	29.38

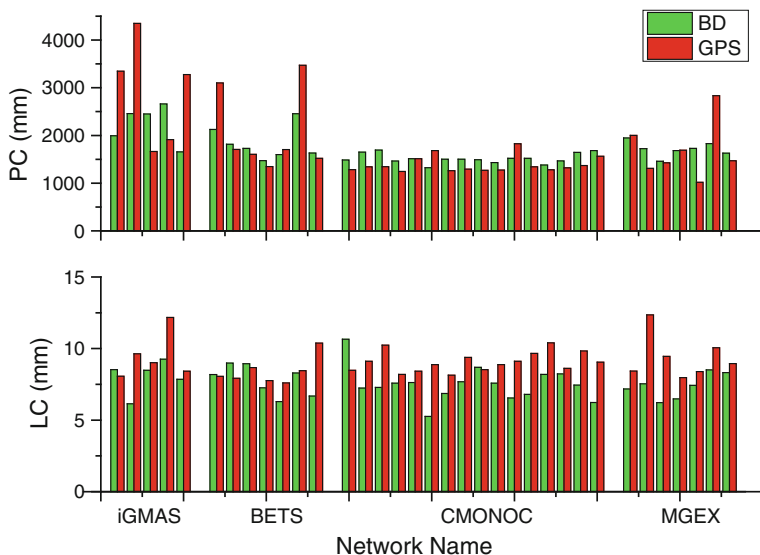


Fig. 11.3 Beidou and GPS residuals

presence of GPS and Beidou significant systematic bias, especially in the elevation direction. System error is influenced with a variety of relevant factors such as the satellite and the receiver phase center corrections, satellite orbit and clock errors and frame of reference [4, 8, 16].

Observation residuals contain a variety of noise sources, including orbit error, PCO and PCV model corrections, multipath effects and noise observations [17–19], and it is one of the most important indicators to assess noise observations and positioning accuracy. Statistically analyzing GPS and Beidou pseudorange ionosphere-free combination of observations (Pseudorange Combination, PC) with carrier-free ionization combined observations (Carrier Phase Combination, LC) residuals respectively, each Beidou and GPS residuals of Beidou continuous operation tracking stations are shown in Fig. 11.3.

As can be seen from Fig. 11.3, PC residuals of GPS and Beidou observation are around 1.8 m, while LC residuals of GPS and Beidou observation are around 8 mm, and residuals of GPS observation are greater than most of residuals of Beidou observation.

According to iGMAS/BETS/CMONOC/MGEX tracking network, GPS and Beidou residuals statistics are shown in Table 11.5. As can be seen from Table 11.5, GPS and Beidou PC residuals are 2.06 and 1.84 m, while LC residuals are 8.98 and 7.76 mm. The reason why GPS residuals are slightly larger than Beidou ones is that the determination of GPS precise orbit and clock errors does not rely on these four tracking networks. However, Beidou orbit determination used iGMAS/BETS/MGEX data, whose orbit and clock errors absorbed some orbit station-related errors. What is more, GPS orbit products is based on IGS, while Beidou orbit

Table 11.5 Statistics of Beidou and GPS residuals

mm	GPS		Beidou	
	PC	LC	PC	LC
iGMAS	2902.68	9.49	2246.56	7.98
BETS	2221.68	8.31	1895.80	8.27
CMONOC	1389.74	9.08	1510.13	7.37
MGEX	1728.88	9.02	1695.43	7.40
ALL	2060.75	8.98	1836.98	7.76

determination and positioning are based on PANDA software, resulting a high degree of consistency [6, 20].

From Table 11.5 it can also be found that CMONOC tracking network has the least residuals, while the iGMAS the largest. The residual of MGEX tracking network is less than that of BETS. The result is consistent with the comparison results of the above-mentioned various tracking network positioning accuracy, indicating that iGMAS tracking station data has relatively poor quality.

11.5 Conclusion

Based on non-differential data processing mode PANDA software of Wuhan University, observations of Beidou continuous operation tracking station (iGMAS/BETS/CMONOC/MGEX) were precise positioning analyzed, while statistical analysis of Beidou and GPS positioning accuracy and observation residuals of each tracking station were done, the results are as follows:

- (1) The STD of differences between GPS and Beidou of four tracking network are respectively 11.36, 7.43 and 31.73 mm in E, N, U three directions, which shows regional post-positioning accuracy of Beidou system has reached 1 cm in horizontal direction and 3 cm in elevation direction. The large discrepancy with GPS positioning accuracy is because of slightly worse precision of Beidou satellite orbit and clock products;
- (2) The average of differences between GPS and Beidou are respectively 4.87, 2.65, 29.38 mm in E, N, U three directions. The systematic bias between GPS and Beidou is caused by a variety of factors such as satellite and receiver phase center corrections, satellite orbit and clock errors, the frame of reference and so on;
- (3) According to positioning accuracy and residuals statistics of each tracking network, the same conclusion is that positioning accuracy of CMONOC tracking network is the best, which is better than those of MGEX and BETS tracking network is worse, while positioning accuracy of iGMAS tracking network is slightly poor.

- (4) The fact that Beidou regional system can achieve centimeter-level precision positioning can provide a reference for the establishment of centimeter-level coordinate reference frame using Beidou system.

Acknowledgments Thanks for each Beidou continuous operation tracking network (iGMAS/BETS/CMONOC/MGEX) providing observation data; thanks for IGS providing precise GPS orbit and clock poor product; thanks for Wuhan University providing Beidou precise orbit and clock product and PANDA software. This work was supported by the iGMAS project of China, the National Nature Science Foundation of China (No. 41325015, No. 41274049, No. 41204029) and the National “863 Program” of China (Grant No. 2014AA123101).

References

1. China Satellite Navigation Office. BeiDou navigation satellite system in public service performance specifications (Version 1.0) [EB/OL]
2. Yang Y (2010) Progress, contribution and challenges of compass/Beidou satellite navigation system. *Acta Geodaetica Cartogr Sin* 39(1):1–6
3. Yang Y, Li J, Xu J, Tang J, Guo H, He H (2011) Contribution of the compass satellite navigation system to global PNT users. *Chin Sci Bull* 56:2813–2819
4. Yang YX, Li JL, Wang AB et al (2014) Preliminary assessment of the navigation and positioning performance of BeiDou regional navigation satellite system. *Sci China: Earth Sci* 2014(57):144–152
5. Lou Y, Liu Y, Shi C et al (2014) Precise orbit determination of BeiDou constellation based on BETS and MGEX network. *Scientific reports*, vol 4
6. Li M, Qu L, Zhao Q et al (2014) Precise point positioning with the BeiDou navigation satellite system. *Sensors* 14(1):927–943
7. Jiao W, Ding Q, Li J et al (2011) Monitoring and assessment of GNSS open services. *J Navig* 64(S1):S19–S29
8. Shi C, Zhao Q, Li M, Tang W, Hu Z, Lou Y, Zhang H, Niu X, Liu J (2012) Precise orbit determination of BeiDou Satellites with precise positioning. *Sci China: Earth Sci* 55:1079–1086
9. Montenbruck O, Steigenberger P, Khachikyan R et al (2013) IGS-MGEX: preparing the ground for multi-constellation GNSS science. In: 4th international colloquium scientific and fundamental aspects of the Galileo programme, ESA
10. Gan W, Zhang R, Zhang Y et al (2007) Development of the crustal movement observation network in China and its applications. *Recent Dev World Seismol* (7):43–52
11. Shi C, Zhao Q, Lou Y et al (2009) PANDA: comprehensive processing software for satellite navigation systems and its research progress. *Spacecraft Eng* 4:64–70
12. Zhao Q (2003) Research on precision orbit determination theory and software of both GPS navigation constellation and LEO satellites. Wuhan University
13. Li M (2011) Research on Multi-GNSS precise orbit determination. Wuhan University
14. Zhao Q, Guo J, Li M et al (2013) Initial results of precise orbit and clock determination for COMPASS navigation satellite system. *J Geodesy* 87(5):475–486
15. Li M, Li W, Shi C et al (2014) Assessment of precipitable water vapor derived from ground-based Beidou observations with precise point positioning approach. *Adv Space Res*
16. Wang G, de Jong K, Zhao Q, et al (2013) Multipath analysis of code measurements for BeiDou geostationary satellites. *GPS Solut* 1–11

17. Zhou SS, Cao YL, Zhou JH et al (2012) Positioning accuracy assessment for the 4GEO/5IGSO/2MEO constellation of compass. *Sci China Phys Mech Astron* 55(12):2290–2299
18. Shi C, Zhao Q, Hu Z et al (2013) Precise relative positioning using real tracking data from COMPASS GEO and IGSO satellites. *GPS Solut* 17(1):103–119
19. Li M, Shi C, Zhao Q, Liu J (2011) Multi-GNSS precision orbit determination. *Acta Geod Cartogr Sin* 40:26–30
20. Xu A, Xu Z, Ge M, Xu X, Zhu H, Sui X (2013) Estimating zenith tropospheric delays from BeiDou navigation satellite system observations. *Sensors* 13:4514–4526

Chapter 12

A Parallel Processing Strategy of Large GNSS Data Based on Precise Point Positioning Model

Yang Cui, Zhiping Lu, Hao Lu, Jian Li, Yupu Wang
and Lingyong Huang

Abstract With continuous increasing of the data scale of GNSS observations network, the computing pressure of data processing is growing. The undifferenced precise point positioning (PPP) model is one of the main strategies of GNSS network data processing. With the increasing of stations' scale, the processing time of PPP pattern also increases linearly, the traditional serial processing pattern need to consume a large amount of computing time. As the PPP model is not related, this model has good characteristics of parallel processing between stations. This paper established a distributed parallel processing strategy based on the PPP model, which can not only improve the efficiency of data processing, but also enhance the efficiency of hardware performance. However, due to the high concurrency of data access and processing, the parallel programming is faced with great challenges which can cause immeasurable results. In this paper, by analyzing the flow characteristics of the PPP method, a parallel GNSS data process model at multi-core and multi node level was set up, and a lightweight parallel programming model was adopted to realize the parallel model. Through a large number of data tests and experiments, high efficiency of parallel processing of GNSS data based on the PPP model was achieved. The experiment shows that, under the environment of four multi-core nodes, the parallel processing is at least six times faster than the traditional serial processing.

Keywords Large GNSS network · Undifferenced model · PPP · Parallel programming · Task parallel library · WCF

Y. Cui (✉) · Z. Lu · H. Lu · J. Li · Y. Wang · L. Huang
School of Surveying and Mapping, PLA Information Engineering University,
Zhengzhou 450052, China
e-mail: cuiyangtj@126.com

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015
Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_12

139

12.1 Introduction

The computational burden of large global navigation satellite systems (GNSS) networks data processing is nowadays already a big challenge, so the theory of rapid processing and application of large network data, especially the real time data processing, has received more and more concern and attention [1–4]. Currently in the data processing for joined multiple ground tracking networks, the double difference (DD) method which is based on the baseline network solutions and the undifference (UD) method based on precise point positioning (PPP) technology are two main treatment strategies [2]. For n stations, the calculation time of double difference is $O(n^3)$. With the increase of stations' scale, the processing time increases in geometric ratio. It is difficult to meet the needs of integrated processing of large scale data. The UD method based on PPP pattern completely breaks the current computational bottleneck of DD network solution, and it gradually becomes the trend of large GNSS network data processing. Although the algorithm complexity of UD method is $O(n)$, the computational cost is still great as the number of stations increases rapidly.

On the other hand, performance of computer hardware platform updates quickly, multi-core CPU and many-core GPU have become the mainstream of current general computer architecture. Ordinary personal computer has a “supercomputing” performance, the corresponding parallel computing technology and its application attract more and more attention. With the rapid development of computer network and communication technology, computing platform and environment have passed through a series of new changes, new type of distributed network computing solution platform emerge, such as Web services, grid computing and cloud computing. Establishing a parallel distributed computing platform by using multiple computers in the network becomes the preferred method of large-scale geodetic observational data processing [5–9].

Traditional GNSS data processing procedures are generally written for single-processor architecture, and the utilization efficiency of current hardware platform computing performance is very low. In order to achieve GNSS data distributed processing in a multi-node network environment, parallel design for the traditional model is a primary problem [7]. As UD calculation procedure for each station has a natural feature of parallel processing, this paper implements distributed computing of UD model based on windows communication foundation (WCF) technology under multi-node network environment, and multi-core parallel computing is achieved by using task parallel library (TPL) technology in a compute node. By using such parallel treatment strategy under multi-core and multi-node network, fast parallel processing of large GNSS network data is realized.

12.2 PPP and Network Parallel Design

12.2.1 PPP Workflow

The PPP approach combines precise clocks and orbits products to calculate very precise positions up to few centimeter level with a single receiver which can be double or single frequency. PPP provides a new solution for wide-area high precision positioning, which forms a set of complete theories, methods and products. Due to the fact that it can directly obtain high precision coordinates and is not restricted to the reference station distribution, PPP has been widely applied. Data processing of single station with precise point positioning includes five parts:

1. Data reading module. The main objects in this part include observation files, precise ephemeris and clocks files, ERP file and antenna file.
2. Data preprocessing module. After reading the raw data, a few steps are implemented, including basic quality checking for observational data, calculate the coordinates of the observation satellites corresponding the signal transmission time, cycle slip detection and marking.
3. Error correction module. Correct and reasonable error correction is the key to achieve high precision positioning outcomes. Referencing to the latest specification of IERS 2010, the original carrier and pseudo-range observation error are corrected. This module includes tropospheric effects correction, carrier phase wind-up effect correction, antenna phase centre correction, relativistic correction, earth deformation effects (solid tides, ocean loading, and pole tide) correction, etc.
4. Parameter estimation module. By using object-oriented programming model, the parameters prior information of current epoch are extract automatically from the previous epoch estimation information, and the new estimation information is updated.
5. Results analysis and output modules. The estimation information of current epoch is analyzed and checked, current results are output and the calculation of the next epoch is prepared.

According to the PPP processing flow above, a set of precise point positioning program has been developed by using C# language, Fig. 12.1 shows the core code for the program domain module, and this module contains various error correction models, data preprocessing and other core code. 45 IGS stations' observational data in the first day of 2013 are selected and processed by using this program. The positioning results are compared with the coordinate solution file which was released by IGS, and the difference is shown on Fig. 12.2. The static positioning results are basically within 2 cm, and which are similar to the positioning accuracy of the similar software.

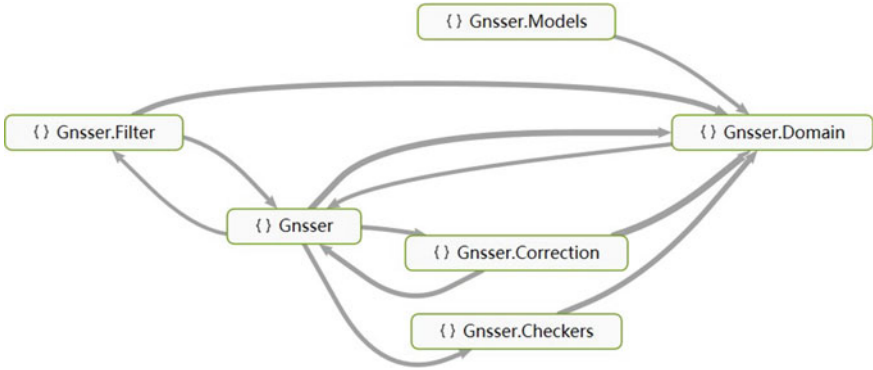


Fig. 12.1 PPP core code

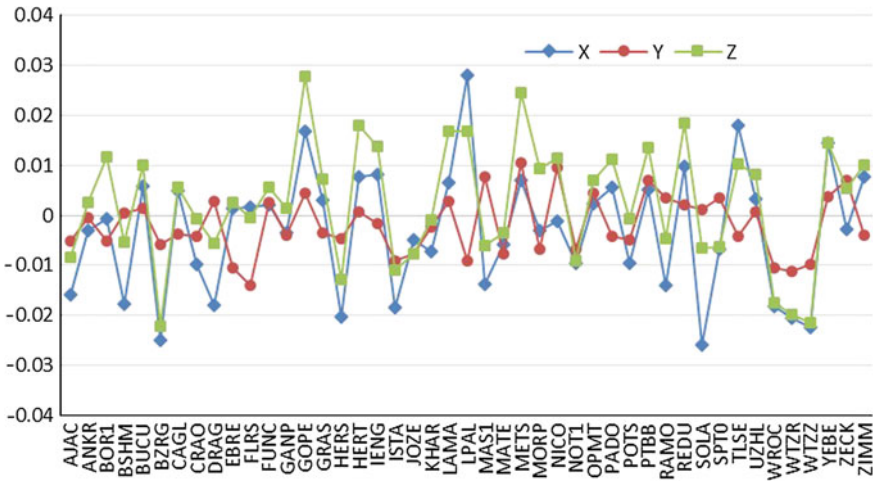


Fig. 12.2 Static positioning calculation results of 45 IGS station

12.2.2 Parallel Network Design of PPP

As the number of the GNSS stations increases, the processing time of UD strategy will greatly increase, and the timeliness of data processing is low. Currently, GNSS data acquisition and resources (network, server, etc.) are distributed, and the distributed computing environment which is composed by existing multiple multicore computers via internet network provides a new hardware platform for GNSS data processing, as shown in Fig. 12.3. The multicore CPU hardware structure of each node belongs to symmetric multi-processor shared-memory architecture, which can be implemented with multi-threads (or process) to give full play to the performance of the hardware. Physically parallel execution can be

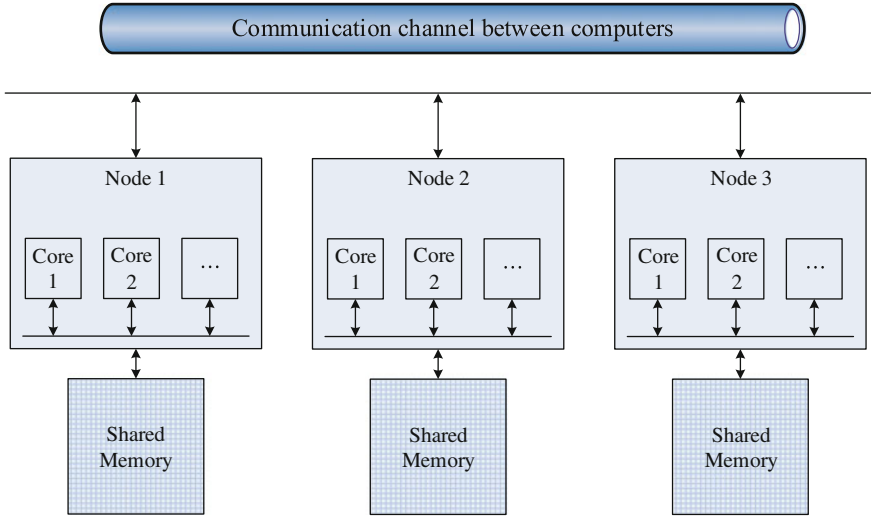


Fig. 12.3 Multi-core and multi-node distributed computing environment

realized on the multicore multithread, which is different from the traditional single-core multithread. Multiple computers which are connected via the Internet can collaborate to processing GNSS data by using open protocols and consistent programming interface.

Due to the fact that UD method has good features in parallel, coarse-grained UD processing of individual station as the minimum unit of large scale GNSS data computing tasks can avoid design of complex parallel procedure and algorithm. In order to ensure load balancing of computing tasks in distributed computing environment, as well as to achieve parallel computing between nodes, a load balancing task allocation method was designed as follows: assuming that current Internet has m available compute nodes and n UD computing tasks. The control center gathers information of each node, such as CPU frequency, memory size, number of physical cores and other information, and then it adopts percentile system to evaluate computers performance, assuming their scores are s_1, s_2, \dots, s_m . When $m > n$, the control center selects n top-ranking nodes, when $m < n$, all of the nodes are taken to participate calculation. Assuming that $q = \min(m, n)$, then the node which gets the score of s_i is assigned with k_i parallel computing tasks:

$$k_i = \left[\left(n / \sum_{j=1}^q s_j \right) \cdot s_i \right] \quad (12.1)$$

where $[]$ is the rounding operation. According to this automatic assignment policy, the distribution of all tasks is dynamic and controllable, and load balancing among all compute nodes is guaranteed. Usually the number of UD computing tasks is larger than the available computing nodes, and current personal computers are

dual-core or quad-core. In order to fully utilize the computing performance of resource, each node can be performed further parallel computing by using multi-core characteristics.

12.3 Parallel Implementation Based on WCF and TPL Technology

As an integrator of the development of Microsoft distributed applications, WCF combines all technologies associated with distributed system under the .NET framework, such as .NET Remoting, ASMX, WSE and so on [10]. WCF is simple and easy to operate, which supports for multiple protocols, including TCP, HTTP, cross-process and customize protocols. It allows developers to build a cross-platform, secure, reliable and transactional solutions, which can be collaborated with existing systems. When using WCF to achieve distributed PPP computing, the traditional PPP applications only need to be established and published as web services. The node which is deployed PPP services is called the server, and each client can call the PPP computing services via the Internet. Through the GNSS services management container, the client can find, access and manage all available GNSS network services. The server provide real-time GNSS services through the Internet, which can avoid a large amount of local data being transmitted to the remote client or fusion end.

Each PPP services are implemented by using parallel programming mode. Compared with single-threaded programs, the development and debugging of the same functions multithreaded programs are much more complex. In 2010, Microsoft released the Parallel Extensions which is applicable to the .NET framework4.0. By using the new task parallel extensions library TPL, the users can use a lightweight task-based programming model to describe the parallel development, and it is easier to write codes which take advantage of multicore processors capability. Users can write extension codes which is along with the increase of available cores, instead of dealing with complex managed threads. The tectonic unit of parallel program is upgraded from the “thread” to the “task” by parallel extensions. The difficulty of parallel computing program development greatly reduced, and the efficiency of parallel development improved.

In the .NET Framework4, a new task-based parallelism namespace (System.Threading.Tasks) is introduced in TPL. By using the methods provided by the parallel static class, parallel computing can be achieved in load balancing. Software developers do not need to build, manage, control and synchronization the complex multiple threads form the bottom. For many PPP computing tasks, each PPP computational task is seen as a “Task”, and the “Task” can be parallelized by circulating parallel class. TPL can achieve parallel execution of a fixed number independent loop iterations without changing the original code inside the loops, which enables all tasks kept busy calculated.

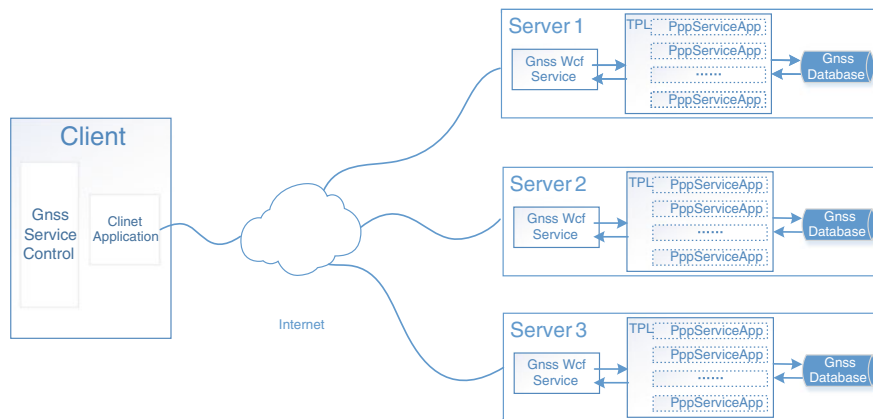


Fig. 12.4 Flowchart of PPP network parallel service implemented based on WCF and TPL technology

Figure 12.4 is the flow chart of UD network parallel service implemented based on WCF and TPL technology. The client application calls the UD multi-core parallel computing services which is provided by multiple computing nodes in the Internet, thus realizing co-processing of large scale GNSS data across multiple nodes. Communication among computing node are simple instructions, avoiding a log of raw observation data transmission in the network. Each node implements a coarse-grained multi-core UD parallel computing without a heavyweight thread code.

12.4 Experiment and Analysis

12.4.1 Experimental Environment and Data

In order to build a distributed computing environment, the GNSS WCF UD service software is installed on four computers, and their location is in Longhai Road campus of Zhengzhou Information Engineering University. Local client which is in the center campus is parallel to invoke computing services provided by these servers. Configuration of the four compute nodes are basically similar (frequency 2.0–2.4 GHz, memory 2–4 GB, Windows 7 operating system), and the physics core are quad-core. The development of WCF UD computing services adopted Visual Studio 2013 C# .NET framework, and the client is a desktop application.

160 IGS stations are randomly selected as the test data. Observation date is January 1, 2013, and sampling rate is 30 s. All data are stored on a public FTP server at the local area network for ease of management. Each computing task contains the name of the stations, observation date and other information.

After receiving the instruction of computing task, the compute node analyses the task and automatically downloads stations data from FTP server, and then parallel UD computing is implemented under the multi-core environment.

12.4.2 Analysis and Comparison

There are two solutions to dealing with UD computing task of different scales: Scheme 1, traditional serial processing in one node; Scheme 2, distributed parallel processing in four nodes. The statistics of computing time of different tasks scale are shown in Fig. 12.5.

In Fig. 12.5, blue squares represent the time overhead of serial computing scheme, and red circles indicate the time overhead of distributed computing scheme. The figure shows that the computational efficiency of distributed processing scheme is higher than that of traditional centralized processing. Thanks to multi-core parallel and multi-node parallel processing strategy, the parallel speedup ratio achieved 6.0 when using four nodes, which is six times faster than traditional serial computing model. For example, serial processing of 160 stations needs 161 min, but the distributed parallel processing only requires 27 min. If more computing nodes are used, the speedup will be greater and the computational efficiency will be higher. With the increase of GNSS data scale, the improvement of processing efficiency of distributed computing is more pronounced, and the computing pressure on the client also greatly reduces. This is the new trend that current cloud computing and other new technologies are chasing.

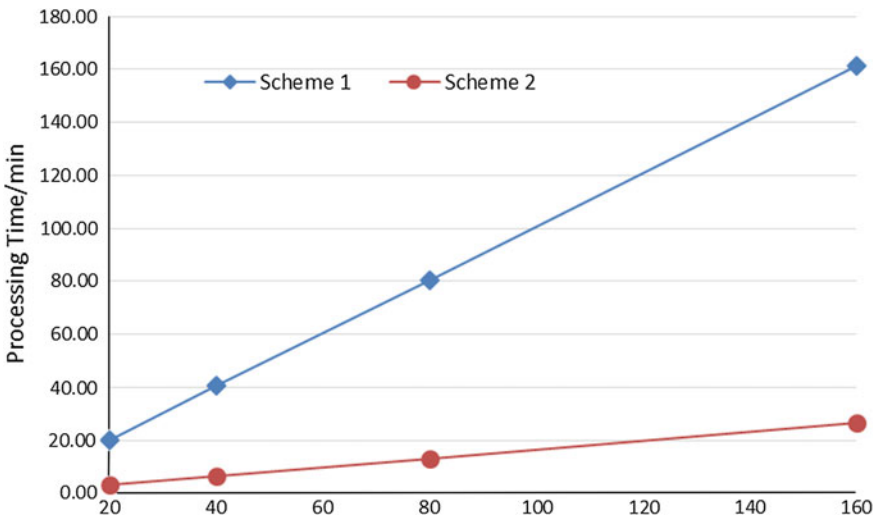


Fig. 12.5 Comparison of the two schemes time overhead

12.5 Conclusion

By analysing the characteristics of the UD model flow, a distributed parallel processing strategy of large GNSS data based on the UD pattern is proposed in this paper, and multi-core and multi-node distributed parallel PPP processing are achieved by using lightweight parallel programming model provided by WCF technology and TPL technology. Finally, the distribution UD co-processing of large-scale GNSS data based on multiple nodes is realized, which not only validates the high efficiency of distributed parallel GNSS data processing, but also improves the efficiency of resource utilization and reduces the computing pressure on the client.

Acknowledgments This paper is supported by the National Natural Science Foundation of China, No. 41274015; the National 863 Program of China, No. 2013AA122501; State Key Laboratory of Geo-information Engineering, NO. SKLGIE2014-M-1-5 and NO. SKLGIE2014-M-1-6.

References

1. Ge M, Gendt G, Dick G, Zhang FP, Rothacher M (2006) A new data processing strategy for huge GNSS global networks. *J Geodesy* 82:199–203
2. Chen H, Jiang W, Ge M et al (2014) An enhanced strategy for GNSS data processing of massive networks. *J Geodesy* 88:857–867
3. Chen J, Zhang Y, Xie Y et al (2014) Rapid data processing of huge networks and multi-GNSS constellation. *Geomat Inf Sci Wuhan Univ* 39(3):253–257 (in Chinese)
4. Cai H (2010) Application research of method of large GNSS network realtime data rapid solution. PhD thesis, University of Wuhan, Wuhan, China. (in Chinese)
5. Chen Z (2014) Research on the key techniques of distributed processing on large scale GNSS observation data. PhD thesis, PLA Information Engineering University, Zhengzhou, China (in Chinese)
6. GPS Dancer (2012) IAG dancer: global solutions for all receivers in the world. <http://www.gpsdancer.org/documents/igs2012.pdf>. Accessed 20 Jan 2014
7. Boomkamp H (2010) Global GPS reference frame solutions of unlimited size. *Adv Space Res* 46(2):136–143
8. Cui Y, Lu Z, Chen Z et al (2013) Research of parallel data processing for GNSS network adjustment under multi-core environment. *Acta Geodaetica Cartogr Sin* 42(5):661–667 (in Chinese)
9. Kai Y (2013) Research on high performance GNSS normal equation processing method based on OpenMP. *J Geodesy Geodyn* 33(1):141–144 (in Chinese)
10. Hillar GC (2012) Professional parallel programming with C#: master parallel extensions with .NET4. TSinghua Univeristy Press, Beijing (in Chinese)

Chapter 13

New Results of Multi-GNSS Orbits Validation Based on SLR Observations

Jinchao Xia, Geshi Tang, Chao Han, Jianfeng Cao,
Hongzheng Cui and Xie Li

Abstract Nowadays there are more than 30 GNSS satellites equipped with laser retro-reflectors. Orbits of these satellites can be independently validated with a high precision using range observations by satellite laser ranging (SLR). China has completed the Asia-Pacific area COMPASS Satellite Network, and is committed to the establishment of the international GNSS Monitoring and Assessment System (IGMAS). Orbit validation is one of the key factors in the development of IGMAS, so this paper validated almost all GNSS satellites carrying reflectors to see whether the IGMAS orbits is qualified, including all 24 of GLONASS, 4 of COMPASS, 4 of GALILEO based on the rapid orbits from BACC IGMAS Analysis center. According to IERS2010 convention, correction models such as station coordinates, laser propagation delay models were introduced to establish the measurement model. The validation of the BACC orbits shows a mean deviation around 5–10 cm for the GLONASS satellites, 10–15 cm for the COMPASS MEO/IGSO satellites, 2.65 m for the COMPASS GEO satellites, and 20–25 cm for the GALILEO satellites.

Keywords GNSS · COMPASS · IGMAS · Satellite orbits

J. Xia (✉) · G. Tang · J. Cao · H. Cui · X. Li
Aerospace Flight Dynamics Laboratory, Beijing Aerospace Control Center (BACC),
No. 26, Beiqing Road, Haidian District, Beijing 100094, China
e-mail: xiajinchao2006@126.com

C. Han
School of Space Technology, Beijing University of Aeronautics and Astronautics,
Beijing 100083, People's Republic of China

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015
Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_13

13.1 Introduction

Satellite Laser Ranging (SLR) is a precise range measurement system consisted of two parts-retro reflectors on a satellite and a ground station that produces short laser pulses. It measures the flight time interval of the pulses from the ground station to the target satellite and back again, in order to calculate the range between the station and the satellite. Since the first SLR experiment in the 1960s, many countries have developed SLR systems. Today over 50 stations of the ILRS network over the world are actively ranging over 60 satellites as well as 5 lunar reflectors [1]. After five decades development, SLR technology has achieved sub-centimeter precision in a single measurement and becomes one of the most precise technologies of the space geodesy means [2].

SLR has become a distinctive technique to validate GNSS orbits since it was proposed to track to GPS satellites, which provides an opportunity to compare GPS and laser system directly [3]. The reasons are as follows. First, compared with GNSS microwave observation, SLR allows an independent way of external validation without ambiguities since SLR's signal is not affected by the ionosphere [4]. Second, this type of satellite, usually irregular shaped, is orbiting at a high altitude of about or greater than 20,000 km, which means the solar radiation pressure is the largest non-conservative perturbation acting on it, resulting the solar radiation hard to model. Third, for some GNSS constellations, the tracking network coverage are limited to a region and far from sufficient, but SLR stations over the world may play an important role to improve the tracking coverage. Therefore SLR is a very important means of external orbit comparison. SLR observations to GNSS satellites may be used for different purposes. First, Satellite Laser Ranging (SLR) observations allow for an independent validation of orbits derived using microwave measurements as provided by IGMAS or *the International GPS Service (IGS)*. Second, laser and microwave observations can be combined to determine orbit [3, 5, 6]. The Chinese COMPASS system has already launched a regional navigation service and pursues build-up of a global system till 2020. In the development process of IGMAS center, the validation to GNSS using SLR observations may be very useful especially when the constellation is not completed.

There are two GPS satellites (GPS 35, 36) equipped with retro-reflector, as well as all GLONASS satellites, multiple COMPASS (M1, M3, G1, I3, and I5) and Galileo satellites (101, 102, 103, 104, and new launched 201). As a key feature, lots of satellites in the new constellations (i.e. Galileo, COMPASS, QZSS and IRNSS) are equipped with laser ranging reflector arrays enabling high-precision two-way ranging measurements. This paper compared values between SLR ranges and computed ranges derived using IGMAS orbits (BACC supplied GPS, GLONASS, COMPASS, GALILEO orbits). Up to now, the SLR observations to satellites were mainly used for dynamical satellites like LAGEOS I, LAGEOS II, CHAMP, or a few GNSS satellites like GPS 35 and 36 or all GLONASS satellites and so on [4–10]. Some are used for combination with GNSS observations for orbit determination, but seldom work were done for multi-GNSS orbit validation.

13.2 SLR Data and Preprocessing

13.2.1 SLR Data

SLR observations are collected by ILRS (International Laser Ranging Service). In the procedure of observation, every station collects raw observations, generates normal points in an interval (5 min for GNSS satellites), and then submits all together into a unified global ILRS release. The resulting range measurements are available at the Crustal Dynamics Data Information System (CDDIS) as normal point data [11].

The latest normal point data format is the Consolidated Range Data (CRD) [12]. It is a flexible, extensible format for the ILRS full rate, sampled engineering, and normal point data. CSTG Normal Point Data is a history format. As for new data, only CRD data is available.

In this paper, all the SLR data for all satellites has been performed in 20-day duration, from Aug 1, 2014 to Aug 20, 2014, which is in a routine operation time of BACC IGMAS center. During this time, SLR stations observed 30 satellites (see Fig. 13.1). Most satellites are observed by more than 15 stations. Some satellites were observed by more than 20 stations, while 3 satellites were observed by less than 10 stations. Because GPS 35 and 36 was no longer observed by ILRS, so there are no SLR observations. BACC supplied orbits of nearly all satellites of 4 systems (GPS, GLONASS, COMPASS and GALILEO) using microwave observations. In order to validate the orbits of BACC GNSS, the rapid ephemerides of BACC were used in this paper (Fig. 13.2).

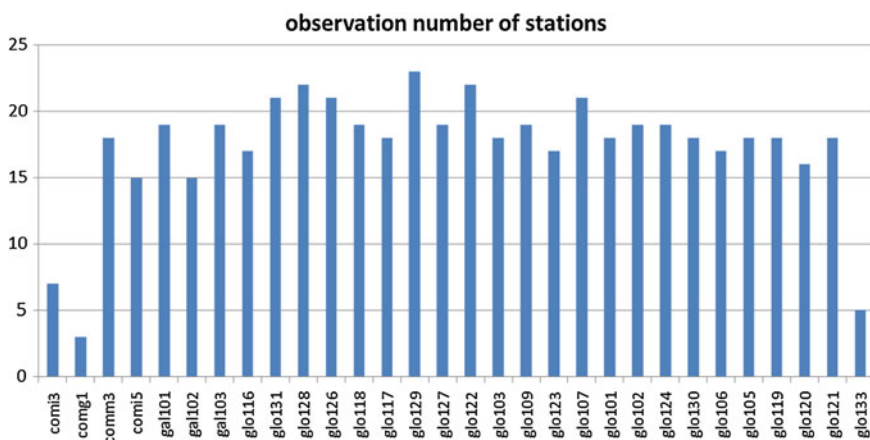


Fig. 13.1 Observation number of stations of every GNSS satellites

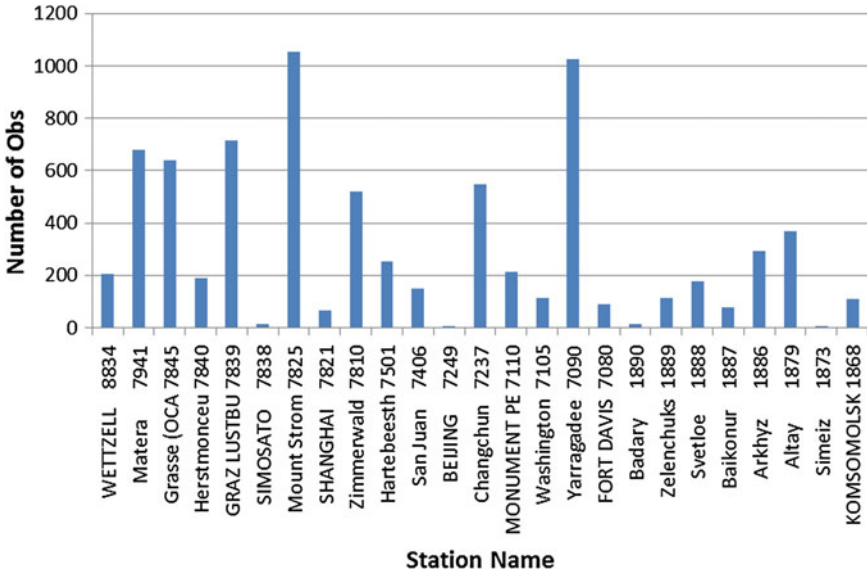


Fig. 13.2 Observation number of stations during validation

13.2.2 SLR Data Preprocessing

After getting the raw data, it's necessary to preprocess it to eliminate poor qualified data and form a convenient data structure for orbit validation. The data structure of a normal point may include laser emission time t , fly time Δt , station ID, satellite ID, station temperature, humidity, air pressure, observation quality and rms. In the preprocessing, SLR data with big rms (>5 cm) should be deleted, as well as the data with elevation angle less than 10° . In addition, the observation of a station should be stable, the observing and calculating difference should be within a certain threshold range according to the precision of SLR system, e.g. within 1 m or less. If the threshold is exceeded, this observation should be considered as a gross error.

Take the i -th station with position $P_i(x_i, y_i, z_i)$ for example, which is observing satellite $S(X^s, Y^s, Z^s)$, so the distance between the P_i and S can be demonstrated as ρ :

$$\rho = \sqrt{(X^s - x_i)^2 + (Y^s - y_i)^2 + (Z^s - z_i)^2} \quad (13.1)$$

For orbit validation, we should obtain three types of data: observed data R_i , station coordinates P_i after eccentricity correction, precise orbit of satellite S with precise CoM. Station position can be derived using ITRF coordinates. Satellite position of the reflecting time can be interpolated using the ephemeris by Chebyshev or Lagrange polynomial interpolation (9-order Chebyshev in this work). We also need some other supporting information, like: (1) IERS C04 EOP data for conversion between the earth-fixed frame and the inertial frame; (2) JPL-DE421 ephemeris

for the calculation of the general relativity corrections and coordinate system conversion; (3) SLR station coordinates: ITRF2008 coordinates and velocity, and eccentric station.

The methods of orbit validation of different GNSS systems are almost the same. The main computing model is on the station coordinates corrections and the propagation delay model. The sore difference is the satellite-related calculations. Because station position is in the ITRF frame, satellite ephemeris should be consistent with it. As different satellite-fixed reference systems have different frames, a transformation should be applied to them.

13.3 SLR Data Measurement Model

The validation of GNSS orbits is based on the difference between the SLR observed range and the computed range. So the computed range should be precisely calculated, as well as the observed range. This means precise correction to station, CoM and propagation delay.

13.3.1 Station Correction

Station's position are influenced by station motion model, as well as solid tide, pole tide, ocean loading tide and so on [13, 14].

Station coordinates can be derived from ITRF2008 coordinates $P_{2005.0}$ and velocity V . The unit of t_{obs} is converted to year, while ecc is the station eccentricity:

$$P = P_{2005.0} + V \times (t_{obs} - 2005.0) + ecc \quad (13.2)$$

Station displacement by solid tide is a necessary correction. The main considerations are cyclical elasticity by the Sun and the Moon gravitation. Wahr tidal model is introduced. Generally, It's a function of the observing time (t), station (r_{Sta}), the position of the Moon (r_{Moon}) and the Sun (r_{Sun}) and tidal Love number (p_{Love}):

$$\Delta P_{stide} = dP_{Solid}(t, r_{Sta}, r_{Moon}, r_{Sun}, p_{Love}) \quad (13.3)$$

The displacement caused by the ocean tide loading is also introduced. Ocean tide loading is the elastic response of the crust to the ocean tides. At a point at moment t , three-dimensional displacement components caused by ocean loading is a function of 11 components tides:

$$\Delta P_{otide} = dP_{Otide}(t, A_{cj}, \omega_j, \chi_j, f_j, \mu_j, \Phi_{cj}) \quad (13.4)$$

Here, j denotes 11 component tides, while f_j, μ_j depends on the longitude of the ascending node of the Moon; ω_j and χ_j are the angular velocity and the astronomical argument at the moment $t = 0$ h. A_{cj}, Φ_{cj} are the amplitude and phase of the station of j component tide, respectively.

The displacement caused by polar tide can be denoted as a function like:

$$\Delta P_{ptide} = dP_{ptide} \left(t, r_{Sta}, g_{Sta}, x_p, y_p, p_{Love} \right) \quad (13.5)$$

Constants like the rotational angular velocity of the Earth (ω) are also concerned; g_{Sta} is the gravity of this station; (x_p, y_p) is the mean polar motion, which can be read from EOP file.

So an accurate station position may be:

$$P = P_{obs} + \Delta P_{stide} + \Delta P_{otide} + \Delta P_{ptide} \quad (13.6)$$

13.3.2 Laser Propagation Correction

The laser propagation delay consists of two main parts: troposphere delay and general relativity effects, so corrections of these should be considered. In addition, CoM should be calculated.

Tropospheric delay in zenith direction may reach a few meters; an effective troposphere model is the Marini-Murray model, which is a function of the station temperature, humidity, laser wavelength, angle of observation height [9].

The tropospheric correction reads as following form,

$$\Delta \rho_{Trop} = d_{Trop}(\lambda, r_{Sta}, P_0, T, E) \quad (13.7)$$

λ is the laser wavelength. Station position r_{Sta} will be transformed to the latitude φ and elevation h of the station, respectively; P_0, T, W , are atmospheric pressure of the station (mbar), temperature (Kelvin) and relative humidity (%), respectively.

The effect of general relativity delay is as the following formula:

$$\Delta \rho_{Rel} = \frac{2GM}{C^2} \ln \left(\frac{r1 + r2 + \rho}{r1 + r2 - \rho} \right) \quad (13.8)$$

M represents a celestial body mass (the Earth), C is the speed of light, $r1, r2$ represents earth center to the satellite and station respectively, ρ is the distance from the station to satellite.

CoM of a satellite (only nominal yaw-mode orbit is involved in this paper): Center correction of the satellite is related with the direction of satellite-fixed axis in space. For GPS, COMPASS MEO and IGSO, GLONASS, GALILEO, Z axis

points earth mass center, while the solar panels is oriented to the Sun, and Y axis is perpendicular to the plane of the Sun-Earth-Satellite, X axis completes a right-hand system; As for COMPASS GEO satellite, Z-axis points center of the earth, and Y axis is perpendicular to the instantaneous orbital plane. The formula is as follows:

$$\Delta\rho_{Offst} = \frac{[C](R-r) \cdot S_p}{|R-r|} \quad (13.9)$$

[C] is transformation matrix from J2000.0 geocentric inertial coordinate system to satellite fixed coordinate system; R , r are the position of the station and the satellite in J2000.0 inertial system; S_p is the satellite-fixed vector of the reflector.

13.4 Validation of GNSS Orbit

13.4.1 Computed Range

The computed range value is necessary to be derived for orbit validation. When dealing with a normal point of pulse emission time t_0 , associated with IGMAS ephemeris, we cannot directly calculate the arrival satellite time t_1 and back to station time t_2 . Since the station and the satellite are all moving, this is an iterative process in the inertial frame. Time t_1 is calculated as follows:

- (1) At time t_0 , the inertial coordinates of the station O_0 , and the position of the satellite S_1 , so we get an approximate value of $t_1 = t_0 + (S_0 - O_0)/c$, c is the speed of light;
- (2) Store S_1 of Step (1) as S_{1old} , and interpolate the new position S_1 with time t_1 . Calculate the distance D from S_1 to S_{1old} , meanwhile calculate the new $t_1 = t_0 + (S_{1new} - O_0)/c$;
- (3) Repeat Step (2) till $D < \text{threshold}$. Threshold can be set to 1 mm.

After t_1 is calculated, a similar process can be done to determine t_2 . It should be noted that, inertial coordinate of the station O_2 should be calculated at t_2 . This calculation is a pure geometric process, relating only the launch time t_0 , station location and satellite ephemeris, no need to consider the light propagation delay.

After getting the value of t_2 , $t_2 - t_0$ is the calculated value of laser pulse flight time. The residual O-C expression as following formula can be used:

$$\rho_o - \rho_c = \frac{\Delta t \cdot c}{2} - \sum \Delta\rho - \frac{t_2 - t_0}{2} \cdot c \quad (13.10)$$

Where $t_2 - t_0$ is the calculated time of flight, Δt is the observed time of flight. $\sum \Delta\rho$ means summary of all the range corrections.

According to this formula, after data preprocessing mentioned in Sect. 13.2, and corrections of position and delay in Sect. 13.3, we can begin orbit validation.

13.4.2 Results and Analysis of Orbit Validation

The resulting residuals are an indicator of the microwave-observation-based orbit errors of the radial direction (laser shooting direction). Without considering stations' time bias and other systematic bias, this is a direct way of orbit validation. As seen from Fig. 13.3, the comg1 (COMPASS GEO) shows a mean deviation of 2.65 m, while the other COMPASS satellites show deviations less than 15 cm (comi3, comi5, and comm3); the GALILEO satellites (gal101, gal102, and gal103) show deviations around 20–25 cm; the GLONASS satellites show deviations less than 10 cm, and most of them are around 5 cm. The highest bar of COMPASS GEO

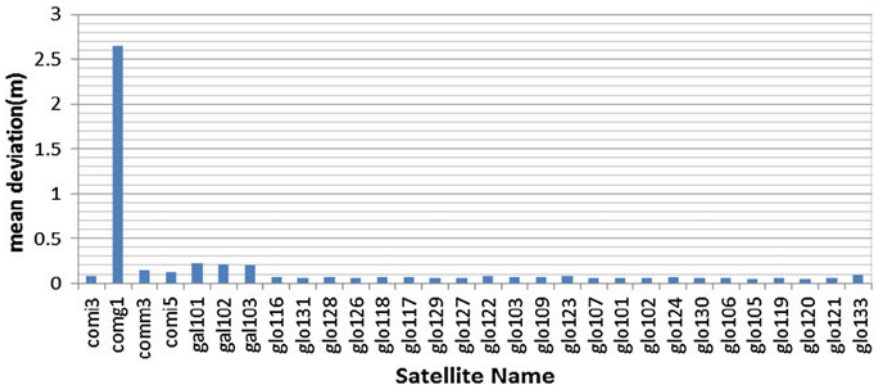


Fig. 13.3 Mean deviations of multi GNSS orbit validation

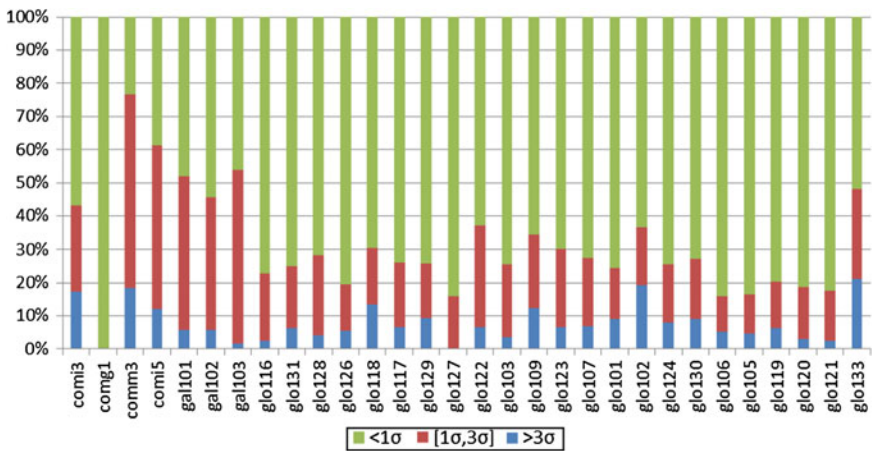


Fig. 13.4 Ratio of observations with different deviations

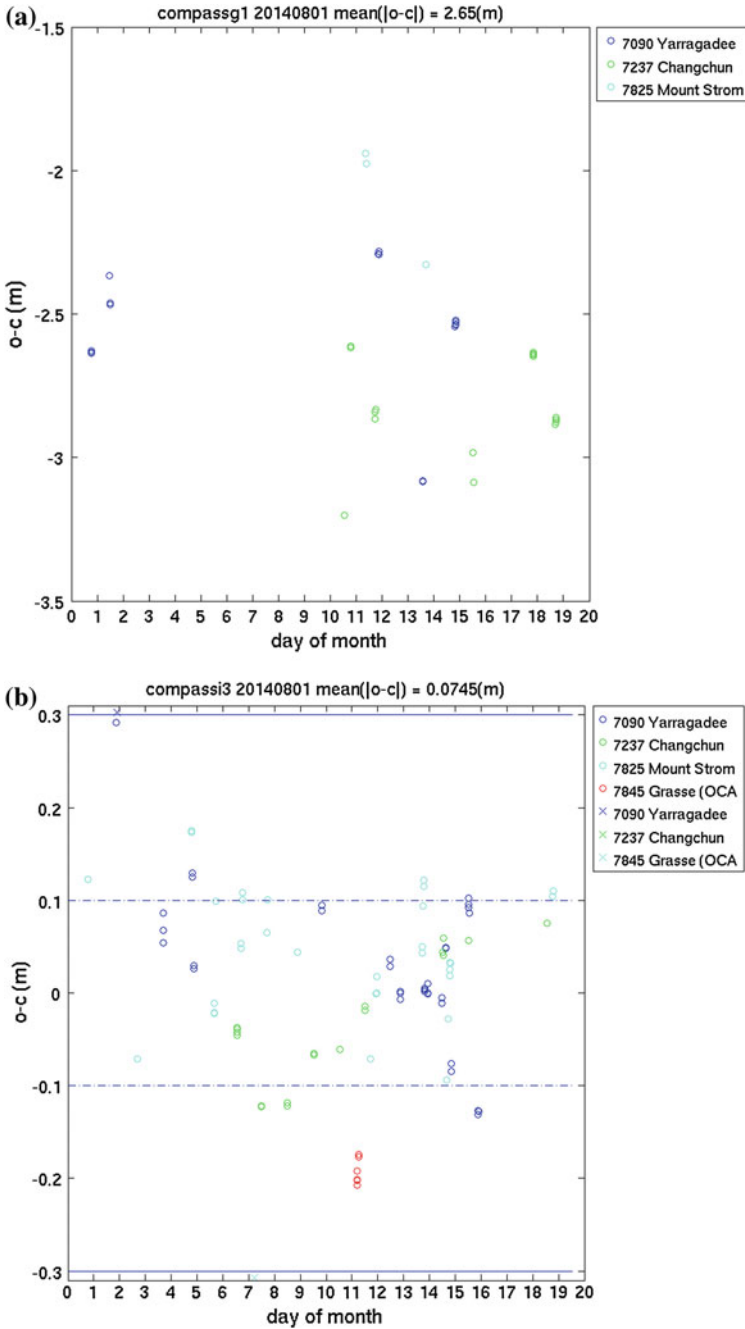


Fig. 13.5 a Validation of COMPASS-G1. b Validation of COMPASS-I3. c Validation of COMPASS-M3. d Validation of GALILEO-103. e Validation of GLONASS-105

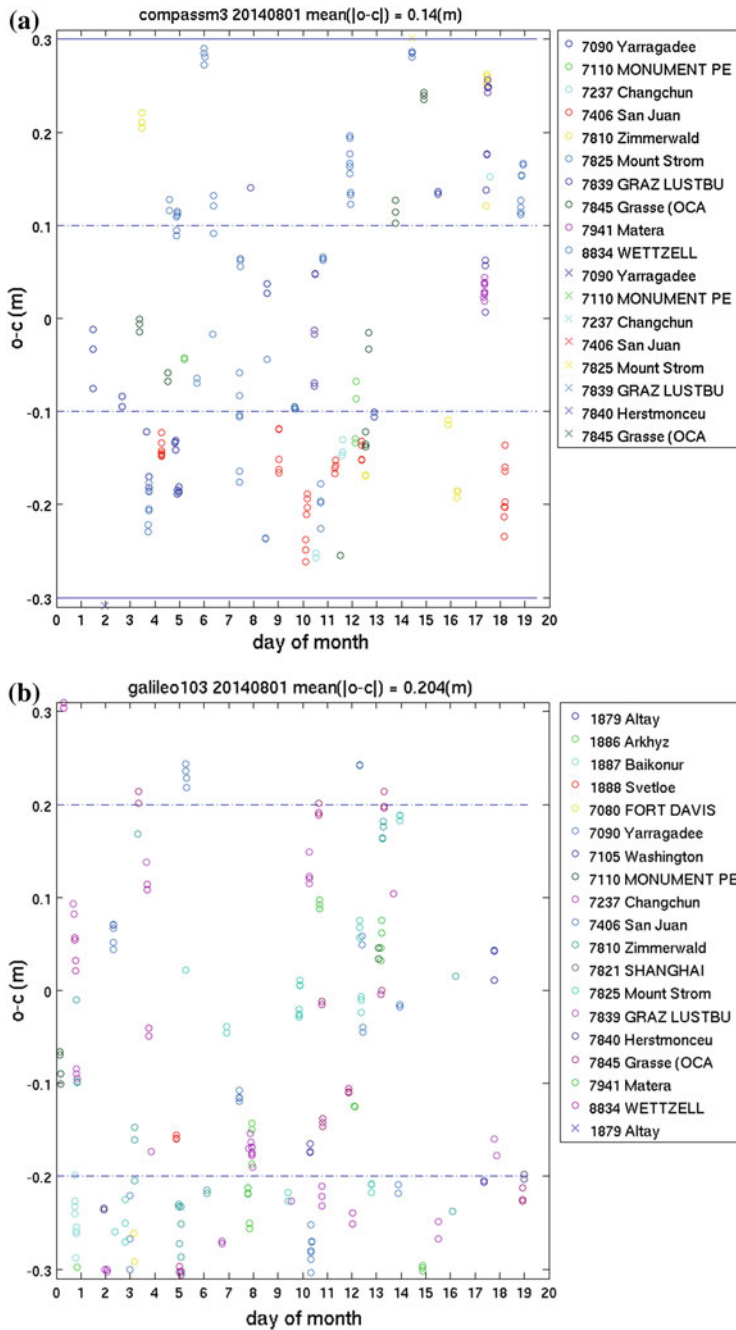


Fig. 13.5 (continued)

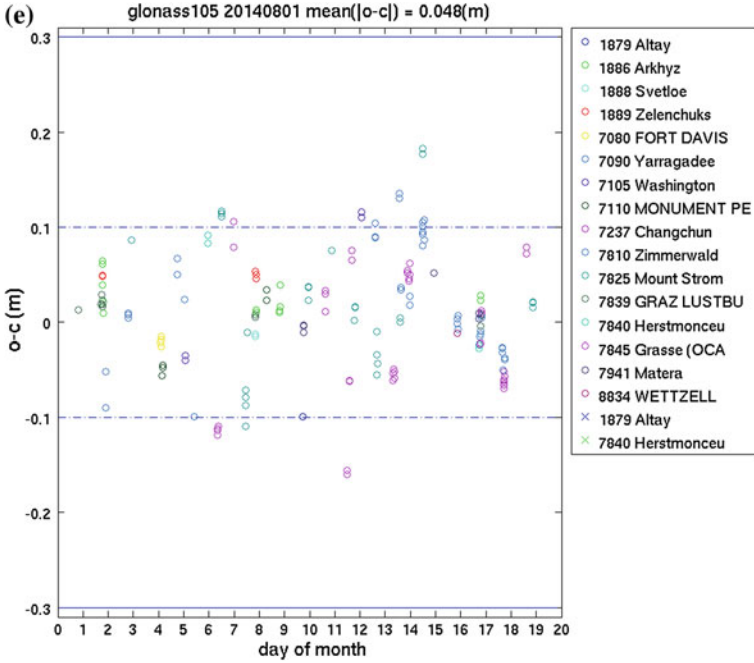


Fig. 13.5 (continued)

reflects the altitude of GEO satellite and its observation geometry. Figure 13.3 shows a mean deviation around 5–10 cm for the GLONASS satellites, 10–15 cm for the COMPASS MEO/IGSO satellites, 2.65 m for the COMPASS GEO satellites, and 20–25 cm for the GALILEO satellites. After checking the stations (7090, 7237, 7825) observing satellite COMPASS GEO G1, which were on good status, we can conclude that the inconsistency between the orbits and the SLR observation resulted from the deduced orbit.

Figure 13.4 shows the ratio of the validation results, we can see most of observations (around 90 %) are less than 3σ . Different thresholds were set for different satellites, according to altitude and observation geometry of different satellites. Here the threshold (σ) COMPASS GEO is set as 5 m, the GLONASS and COMPASS MEO and IGSO as 10 cm, the GALILEO as 20 cm, respectively. Here, the mean deviation of each satellite is expressed as

$$mean(|\rho_o - c|) = \frac{1}{n} \cdot \sum_{i=1}^n |\Delta\rho_i| \quad (13.11)$$

In Eq. (13.11), n is the number of observations with deviation less than 3σ .

Some detailed validations of satellites are shown in Fig. 13.5a–e. Figure 13.5a reflects the radial deviation between COMPASS-G1. Because it is a GEO satellite

orbiting at a high altitude, the number of observation is a little less, compared with other satellites. Figure 13.5b, c show validations of COMPASS-I3 and COMPASS-M3, and the former is centralized in the threshold, while the latter is a little diverse between 3 times of threshold.

Figure 13.5d, e show validations of GALILEO-103 and GLONASS-105, and the former is a little diverse with quite lot observations below the lower threshold line, while the latter is centralized within the upper and lower threshold.

In a word, the method of validating GNSS satellite orbit using SLR observations was studied in this paper. We performed orbit validation for 30 GNSS satellites during 20 days of data from BACC (one of IGMAS analysis center). The results show we achieved the standard of IGMAS. The validations performed here have to be extended to a long time span. Such an extension of this study will allow for the measurement of differences between the SLR and GNSS observations. The results have the potential to judge the orbit of IGMAS center.

References

1. Noll C, Pearlman M, Torrence M (2013) ILRS Station Reporting (2013). http://cddis.gsfc.nasa.gov/docs/2013/ilrs_stations_lw18_20131111.pdf. Accessed 30 Dec 2014
2. Zhao Y, Liu NL (1999) Satellite laser ranging: status and development. *Bull Surv Mapp* 12:23–26 (in Chinese)
3. Zhu SY et al (1997) Apropos laser tracking to GPS satellites. *J Geodesy* 71:423–431
4. Zhao G et al (2013) Precise orbit determination of COMPASS satellites using satellite laser ranging. *CSNC* 245:221–229
5. Kizilsu G et al (2000) SLR precision analyses for LAGEOS I & II. *Earth Planets and Space* 52 (10):789–794
6. Thaller D et al (2011) Combination of GNSS and SLR observations using satellite co-locations. *J Geodesy* 85(5):257–272
7. Qu F et al (2003) Precise orbit determination of GPS35 satellite using SLR data. *Acta Geodaet Cartograph Sin* 32(3):224–228 (in Chinese)
8. Xianping Q et al (2005) Evaluation of CHAMP satellite orbit with SLR measurements. *Geomatics Inf Sci Wuhan Univ* 30(1) (in Chinese)
9. Liu YQ et al (2007) Evaluation of GPS35 satellite precise orbit with SLR measurements. *Eng Surv Mapp* 16(2):36–38
10. Li M, Shi C et al (2011) Multi GNSS precision orbit determination. *Acta Geodaet Cartograph Sin* 40(S1):26–30
11. Normal point data. ftp://cddis.gsfc.nasa.gov/pub/slr/data/npt_crd/. Accessed 30 Dec 2014
12. Ricklefs RL, Moore CJ et al (2014) Consolidated laser ranging data format (CRD) version 1.01. http://ilrs.gsfc.nasa.gov/docs/2009/crd_v1.01.pdf. Accessed 30 Dec 2014
13. McCarthy DD, Petit G (2004) IERS conventions (2003). IERS Technical Note 32. <http://www.iers.org/IERS/EN/Publications/TechnicalNotes/tn32.html>. Accessed 30 Dec 2014
14. Petit G, Luzum B (2014) IERS Conventions (2010). IERS Technical Note 36. <http://www.iers.org/IERS/EN/Publications/TechnicalNotes/tn36.html>. Accessed 30 Dec 2014

Chapter 14

Study in BDS Uncombined PPP Ionospheric Delay Estimation and Differential Code Biases

Huarun Wang, Hongzhou Chai, Min Wang, Zongpeng Pan
and Yang Chong

Abstract Station-satellite DCB may be absorbed into the ionospheric delay in uncombined PPP resolution because of the direct usage of the precise clock offset production based on the ionosphere-free combination estimation. BDS station-satellite DCB is separated in this paper with the ionospheric delay estimation of uncombined PPP, analysis in results and stability is also introduced. Slant ionospheric delay is inverted with ionosphere observation and separated DCB at last.

Keywords BDS · Uncombined PPP · Station-Satellite DCB · Ionospheric delay

14.1 Introduction

In recent years, the technology of Precise Point Positioning (PPP) is widely applied using the precise orbit and clock offset production without establishing base station to achieve relatively higher accuracy of positioning [1, 2]. Traditional PPP applies ionosphere-free combination thus abandoning partial original information and enlarging the combination noise, uncombined PPP can avoid these drawbacks by estimating the slant ionospheric delay which is regarded as an unknown parameter in the observation equation. Owing to the precise clock offset estimated by ionosphere-free combination, station-satellite DCB is directly involved into the uncombined PPP estimation, and absorbed by ionospheric delay, receiver clock offset and integer ambiguity [3, 4]. A comparison of coordinate correction and

This work is supported by National Natural Science Foundation of China (41274045).

H. Wang (✉) · H. Chai · M. Wang · Z. Pan · Y. Chong
Geography Space Information Institute, Information Engineering University,
Zhengzhou 450052, China
e-mail: 593855433@qq.com

clock offset between uncombined PPP and traditional PPP is given firstly in this paper, then a separation of station-satellite DCB is processed by ionosphere modeling with the ionospheric observation containing station-satellite DCB as well as the theory of demarcation with carrier phase smoothed pseudorange [5–7]. Meanwhile analysis of stability in station-satellite DCB corresponding to different constellations is also given and the slant ionospheric delay of uncombined PPP is reached at last with station-satellite DCB separated.

14.2 Principle of Uncombined PPP and Ionosphere Observation

The influence of code hardware delay can be contained into clock offset because of the fact that precise satellite orbit and clock offset production are estimated by ionosphere-free PPP combination, its magnitude can be represented as Eq. (14.1):

$$\Delta B = \frac{f_1^2 B_1 - f_2^2 B_2}{f_1^2 - f_2^2} \quad (14.1)$$

where f_j is frequency, $B_{r,P_j}, B_{P_j}^s$ is station and satellite pseudorange code hardware delay, $B_1 = B_{r,P_1} + B_{P_1}^s$, $B_2 = B_{r,P_2} + B_{P_2}^s$ which should be deducted from code observation equation are the comprehensive effect on two frequency points respectively. Ignoring higher order ionospheric influence, the uncombined PPP code and carrier phase observation equation based on precise orbit and clock offset can be represented as follows:

$$\begin{aligned} P_{r,j}^s &= -\mu_r^s \cdot \Delta r + cdt_r + M_r^s \cdot zpd_r + \gamma_j \tau + \varepsilon_{P_j} \\ L_{r,j}^s &= -\mu_r^s \cdot \Delta r + cdt_r + M_r^s \cdot zpd_r - \gamma_j \tau + \lambda_j N_{r,j}^s + \varepsilon_{L_j} \end{aligned} \quad (14.2)$$

where μ_r^s , Δr are line-of-sight direction unit vector and station coordinate correction respectively, dt_r is receiver clock offset, M_r^s is troposphere zenith path delay mapping function, τ is ionospheric observation obtaining station-satellite DCB meanwhile $\gamma_j = f_1^2/f_j^2$, λ_j is wave length, $N_{r,j}^s$ is integer ambiguity, ε_{P_j} , ε_{L_j} are code and carrier phase noise respectively, I is slant ionospheric delay on B1 frequency point and meets the relation:

$$I = \tau + \frac{f_2^2(B_2 - B_1)}{f_2^2 - f_1^2} \quad (14.3)$$

In this paper, kalman filter model is adopted in uncombined PPP estimation. For static positioning, coordinate correction is set unchanged parameter, troposphere zenith path delay is set random walk model, clock offset and ionospheric observation meet

white noise process, integer ambiguity is set unchanged without cycle slip, yet valued huge noise when cycle slip happens [8].

Carrier phase smoothed pseudorange is used to estimate ionospheric delay and separate hardware delay biases in traditional geometry-free mode, geometry-free combination is defined as:

$$\begin{aligned}\Phi_I &= \Phi_1 - \Phi_2 = (\gamma_2 I - \gamma_1 I) + (\lambda_1 N_1 - \lambda_2 N_2) \\ P_I &= P_2 - P_1 = (\gamma_2 I + B_2) - (\gamma_1 I + B_1)\end{aligned}\quad (14.4)$$

Also carrier phase smoothed pseudorange is set as ionospheric observation $\tilde{\Phi}_{I,arc}$ which is represented as [9]:

$$\begin{aligned}\Phi_{I,arc} &= \Phi_I - \langle \Phi_I - P_I \rangle_{arc} = (\gamma_2 I + B_2) - (\gamma_1 I + B_1) \\ &= \frac{f_1^2 - f_2^2}{f_2^2} I + (B_2 - B_1)\end{aligned}\quad (14.5)$$

where $\langle \rangle_{arc}$ represents the observation average value in a continuous arc of a specific satellite, and $\tilde{\Phi}_{I,arc} = \frac{f_2^2}{f_1^2 - f_2^2} \Phi_{I,arc}$, the relation can be written as:

$$I = \tilde{\Phi}_{I,arc} + \frac{f_2^2 (B_2 - B_1)}{f_2^2 - f_1^2}\quad (14.6)$$

Obviously the two ionosphere observations share a common format containing ionospheric delay and station-satellite DCB except for original observation and error influence. Station-satellite DCB can be absorbed by ionospheric delay, receiver clock offset, integer ambiguity according to the linear relation in uncombined PPP observation equation, whereas can only be absorbed by ionospheric delay in geometry-free mode, thus there may exist some differences in station-satellite DCB estimated with the two kinds of ionosphere observation.

14.3 Ionosphere Modelling and Station-Satellite DCB Separation

According to the linear relation between ionosphere observation and station-satellite DCB, there exists rank-deficient in equations formed with every satellite's observation data on each epoch, and the station-satellite DCB can be separated through ionosphere modeling. Slant ionospheric delay can be mapped into vertical zenith direction with ionosphere single layer model and the correlation between ionospheric delay and satellite elevation angle [10].

$$VTEC = MF(z) \cdot STEC \quad (14.7)$$

The mapping function is represented as:

$$MF(z) = \cos\left(\arcsin\left(\frac{R}{R+H}\sin(\alpha z)\right)\right) \quad (14.8)$$

Where z is elevation angle, R is earth radius ($R = 3671$ km), H is the altitude away from earth surface ($H = 506.7$ km, $\alpha = 0.9782$), ionosphere model applies spherical harmonics function.

$$VTEC = \sum_{n=0}^{n_{\max}} \sum_{m=0}^n \tilde{P}_{nm}(\sin \beta)(a_{nm} \cos ms + b_{nm} \sin ms) \quad (14.9)$$

Where β is the geocentric latitude of the ionosphere pierce point (IPP), s is the sun-fixed longitude of IPP, n_{\max} is maximum order (defaulted as 2), \tilde{P}_{nm} is normalized Legendre polynomials, a_{nm} , b_{nm} are ionosphere model coefficients which is estimated every 2 h, slant ionospheric delay on B1 frequency point meets the relation with total electron content as:

$$STEC = \frac{f_1^2 \cdot I}{40.28} \quad (14.10)$$

Function can be easily got as:

$$\frac{MF(z) \cdot f_1^2}{40.28} I = \sum_{n=0}^{n_{\max}} \sum_{m=0}^n \tilde{P}_{nm}(\sin \beta)(a_{nm} \cos ms + b_{nm} \sin ms) \quad (14.11)$$

The ionosphere model coefficients can be estimated with Eq. (14.11) according to least squares method, also station-satellite DCB and ionospheric delay can be obtained then.

14.4 Results and Analysis

14.4.1 Results and Analysis of Uncombined PPP Estimation

Test data with 30 s intervals is selected from MGEX station CUT0 on 240th day in 2013 in this paper, kalman filter estimation with ionosphere-free combination and uncombined PPP are both applied. The initial coordinate is set as true reference value provided by authority institution, owing to the lack of BDS satellite antenna phase centre correction, only the polar motion and solid earth tide are considered in the error correction. In static mode, the result of position estimation is shown as Fig. 14.1.

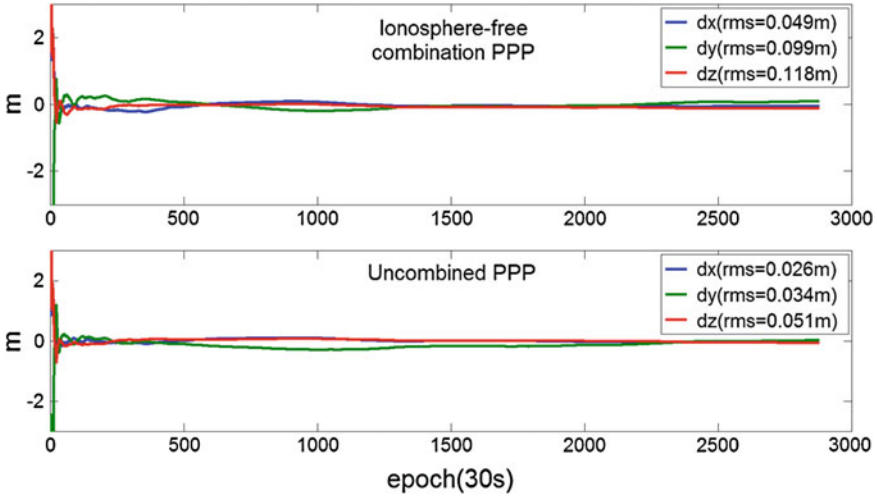


Fig. 14.1 Position correction series of PPP in CUT0 station

Figure 14.1 shows the comparison of traditional ionosphere-free combination PPP and uncombined PPP in position estimation, the RMS on X, Y, Z direction in a single day is 0.049, 0.099, 0.118 m respectively in uncombined PPP's result, comparatively the latter reaches the result of 0.026, 0.034, 0.051 m. Analysis from convergence rate shows that the accuracy RMS on X, Y, Z direction in half an hour is 0.069, 0.272 and 0.287 m in uncombined PPP, the other reaches 0.052, 0.164 and 0.152 m. Obviously, uncombined PPP performs better than ionosphere-free PPP in both positioning correction estimation and convergence rate because the former remains an original noise whereas the latter one enlarges the noise by nearly three times.

Figure 14.2 shows the P1 original observation and receiver clock offset estimation and mutual differences in the two PPP modes.

It can be easily seen that P1 pseudorange shows a seasonal variation because of seasonal leap of receiver's clock, also the clock offset estimated shares the same current. Although results of the two PPP modes present a good consistency, their mutual differences RMS reaches 0.439 m which equals 1.464 ns. According to former analysis, this consequence may be caused by absorption by receiver clock due to the linear relation between station-satellite DCB and clock offset.

14.4.2 Test of Station-Satellite DCB Separation

Here observation data of C05, C09, and C11 is selected to calculate ionospheric observation with the method introduced before, Figs. 14.3, 14.4 and 14.5 show their series respectively.

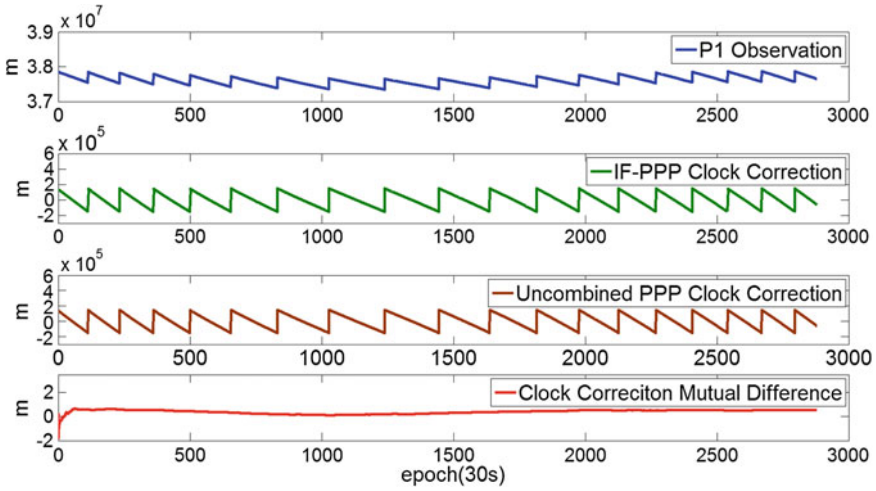


Fig. 14.2 P1 observation and clock offset and mutual difference

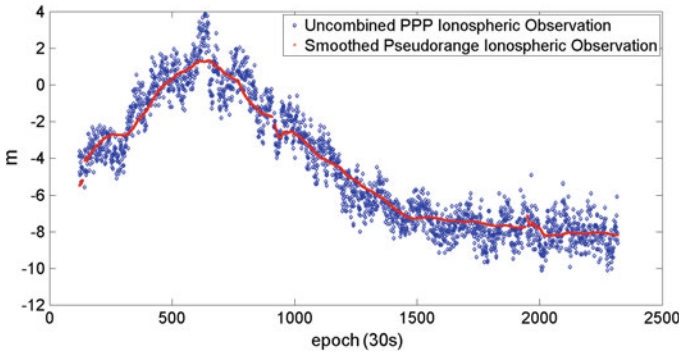


Fig. 14.3 Ionospheric observation series of C05 satellite

Obviously, ionospheric observation calculated by uncombined PPP which is characterized as white noise is much noisy, it seems dispersive but still conforming with the result from smoothed pseudorange. It also can be seen that the ionospheric observation from the two methods shows different situation for different constellations, for instance the results estimated from C05 and C09 share little discrepancy, however the calculated values from C11 exist obvious systemic deviation. The reason is probably that station-satellite DCB in uncombined PPP can be absorbed into ionospheric delay, clock offset and integer ambiguity but can only be absorbed into ionospheric delay in smoothed pseudorange method.

After obtaining ionospheric observation, station-satellite DCB can be separated through ionosphere modeling with the method introduced in Sect. 14.3. Because BDS satellite’s DCB production hasn’t been released yet, DCB of satellite and

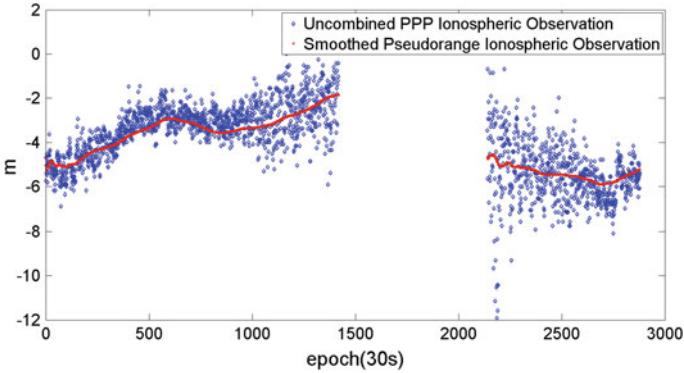


Fig. 14.4 Ionospheric observation series of C09 satellite

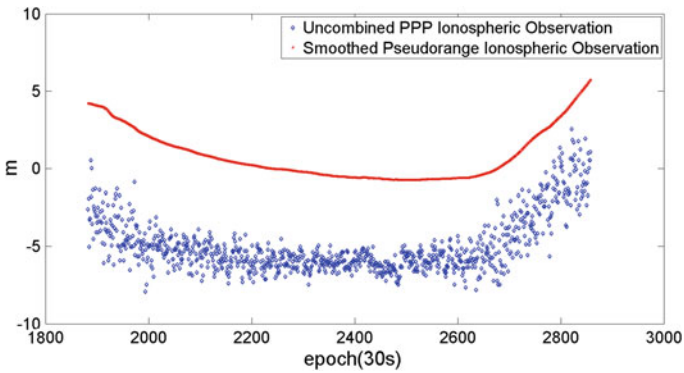


Fig. 14.5 Ionospheric observation series of C11 satellite

receiver are united into one parameter to estimate rather than separating the satellite DCB with constraints. Figure 14.6 shows the comparative statistical results of station-satellite DCB separated by the two ionospheric observation calculated from CUT0 station’s data on 240th day in 2013, and C01 satellite doesn’t join the estimation due to the lack of precise orbit that day.

It can be seen that station-satellite DCB estimated by the two kinds of ionospheric observation from most non-MEO satellites are near except for C08, the values from MEO satellites differ largely in general. Figure 14.7 gives a statistical variation of the station-satellite DCB estimated corresponding to each satellite in CUT0 station on 236–240th day.

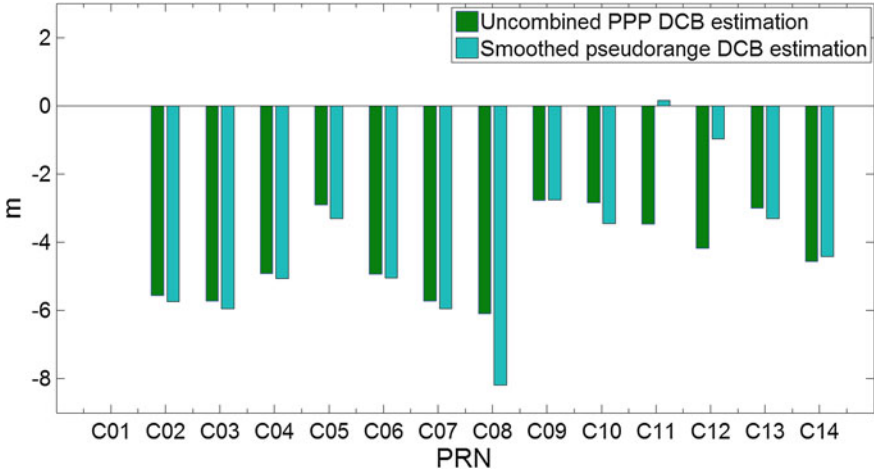


Fig. 14.6 Comparison of station-satellite DCB estimation on 240th day in CUT0 station

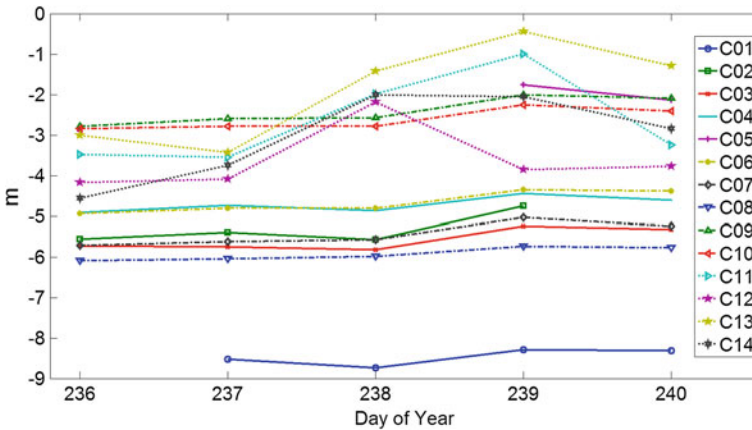


Fig. 14.7 Station-satellite DCB estimation of every satellite on 236–240th day in CUT0 station

Table 14.1 gives the specific station-satellite DCB estimation and standard deviation in the 5 days.

Figures 14.8 and 14.9 shows the variation of part of GPS station-satellite DCB estimated and the standard deviation respectively in these days.

Naught accrues when there is no precise ephemeris related to the corresponding satellite which is removed in least squares estimation. From the figures and table, BDS station-satellite DCB varies stably, the estimation stability of GEO and IGSO

Table 14.1 Station-satellite DCB estimation and standard deviation (unit: m)

PRN day	1	2	3	4	5	6	7
236	-8.310	0.000	-5.326	-4.609	-2.136	-4.378	-5.245
237	-8.287	-4.742	-5.235	-4.441	-1.756	-4.348	-5.030
238	-8.739	-5.569	-5.813	-4.863	0.000	-4.806	-5.573
239	-8.523	-5.400	-5.744	-4.741	0.000	-4.806	-5.618
240	0.000	-5.553	-5.719	-4.915	-2.896	-4.934	-5.716
Std	0.211	0.390	0.266	0.193	0.581	0.271	0.288
PRN day	8	9	10	11	12	13	14
236	-5.775	-2.089	-2.402	-3.239	-3.769	-1.291	-2.842
237	-5.741	-2.011	-2.253	-0.991	-3.852	-0.446	-2.054
238	-5.979	-2.572	-2.787	-1.983	-2.187	-1.413	-2.008
239	-6.037	-2.593	-2.779	-3.547	-4.087	-3.432	-3.746
240	-6.087	-2.778	-2.846	-3.472	-4.175	-3.004	-4.557
Std	0.156	0.338	0.267	1.121	0.815	1.254	1.104

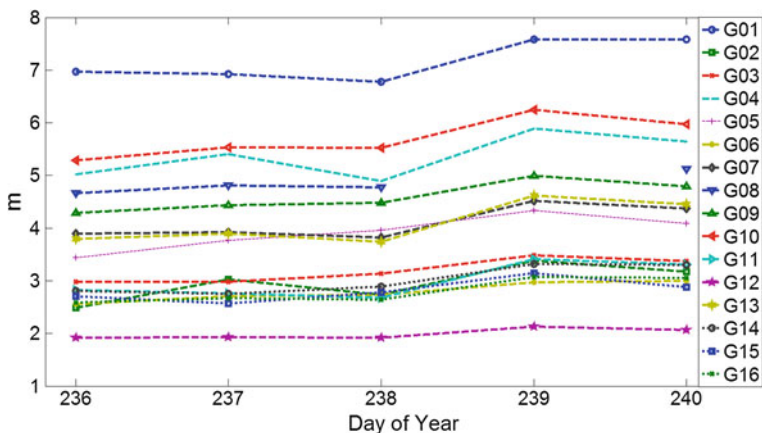


Fig. 14.8 Station-satellite DCB estimation of GPS satellite on 236–240th day in CUT0 station

satellites is equivalent with the standard deviation of 0.296 m, MEO satellites have a worse result with the standard deviation of 1.073 m, however GPS station-satellite DCB is stable in general with the standard deviation of 0.274 m which is common with BDS’s GEO and IGSO but better than BDS’s MEO, the reason is maybe that magnitude of station-satellite DCB absorbed into clock offset and integer ambiguity differs from BDS constellations.

Figures 14.10 and 14.11 gives the ionospheric delay series of C05 and C11 respectively which is obtained from ionospheric observation removing the station-satellite DCB.

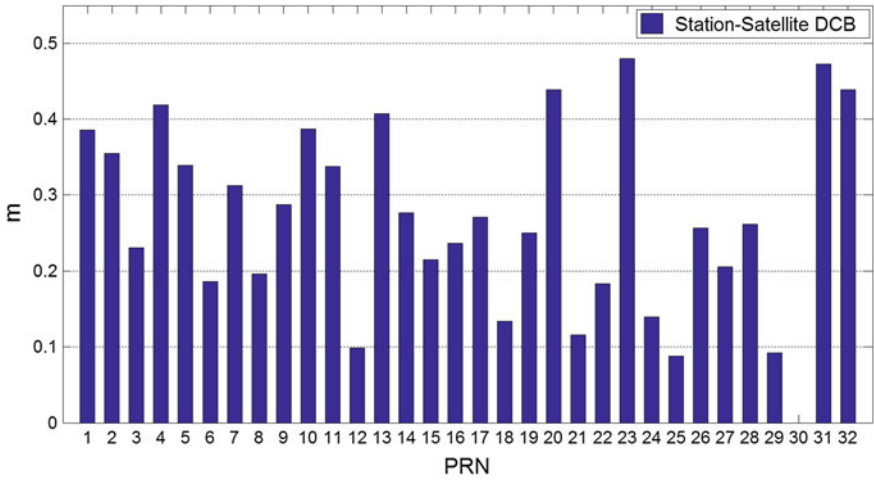


Fig. 14.9 Station-satellite DCB estimation standard deviation of GPS satellite on 236–240th day in CUT0 station

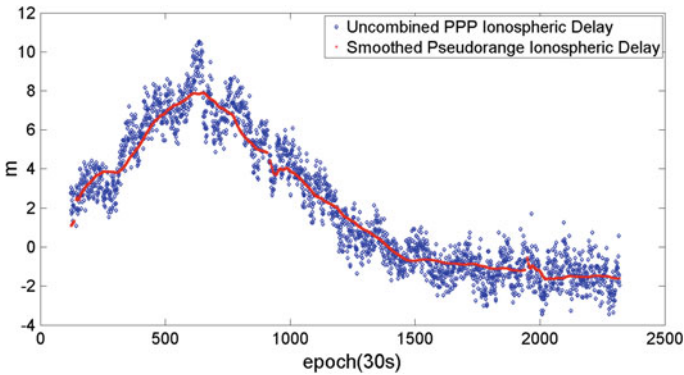


Fig. 14.10 Ionospheric delay series of CO5 satellite after removing station-satellite DCB

From the figures, the ionospheric delay after removing station-satellite DCB with the two different methods performs a good consistency, and the variation trend shows good agreement with satellite constellations and properties that has been known. Notably, some ionospheric delay values estimated in partial period of time is negative which are inconsistent with real value, this maybe result from the uncorrected influence such as antenna phase center and antenna altitude offset, some other station systemic errors which haven't been taken account of may also cause the negative values.

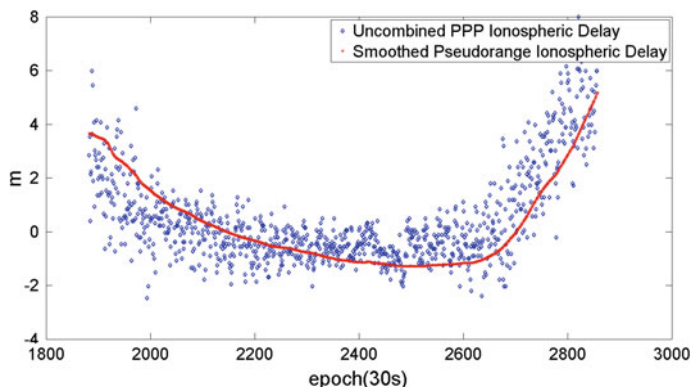


Fig. 14.11 Ionospheric delay series of C11 satellite after removing station-satellite DCB

14.5 Conclusions

Uncombined PPP has an advantage of positioning accuracy and convergence rate compared to traditional ionosphere-free combination PPP. The separated station-satellite DCB from the two kinds of ionospheric observation exists differences because station-satellite DCB in uncombined PPP can be absorbed into ionospheric delay as well as receiver clock and integer ambiguity. Also the stability of station-satellite DCB estimated from uncombined PPP in different BDS constellations is variant, stability corresponding to GEO and IGSO seems equivalent with standard variation of 0.296 m, whereas the MEO's is worse with standard variation of 1.073 m.

Due to the uncorrected systemic corrections such as antenna phase center offset and antenna altitude, although the ionospheric delay inverted coincides with the known ionosphere space-time properties, there exists negative values in some period of time and a systemic bias in general.

Acknowledgments GNSS Research Centre of Wuhan University is gratefully acknowledged for providing BDS data and precise ephemeris. This work is supported by China National Natural Science Foundation of China (No: 41274045).

References

1. Zhang BC, Ou JK, Yuan YB, Zhong SM (2010) Precise point positioning algorithm based on original dual frequency GPS code and carrier phase observations and its application. *Acta Geodaetica Cartogr Sin* 39(5):478–483
2. Li M (2012) Precise point positioning with BeiDou navigation satellites system. In: The 3rd China satellite navigation conference

3. Zhang BC, Ou JK, Yuan YB, Li ZS (2011) Calibration of slant total electron content and satellite-receiver's differential code biases with uncombined precise point positioning technique. *Acta Geodaetica Cartogr Sin* 40(4):447–453
4. Zhang BC, Ou JK, Li ZS, Yuan YB (2011) Determination of ionospheric observables with precise point positioning. *Chinese J Geophys* 54(4):950–957
5. Jin R, Jin S, Feng G (2012) M_DCB: Matlab code for estimating GNSS satellite and receiver differential code biases. *GPS Solut* 16(4):541–548
6. Wu XL, Ping JS, Liu L (2011) Hardware delay calculation of regional satellite navigation system. *Geomat Inf Sci Wuhan Univ* 36(10):1218–1221
7. Liu J (2011) Study of ionosphere prediction models and discussion of pre-seism ionosphere anomalous. Information Engineering University, Zhengzhou
8. Yan W (2012) The Research and application of GNSS system simulation and precise point positioning. Institute of Geodesy and Geophysics Chinese Academy of Sciences, Wuhan
9. Spits J (2012) Total electron content reconstruction using triple frequency GNSS signals. deLiege University
10. Yuan YB (2002) Study of the theory and method in ionosphere monitoring and delay correction based on GPS. Institute of Geodesy and Geophysics Chinese Academy of Sciences, Wuhan

Chapter 15

Research of Key Technology on the Combination of GPS, VLBI, SLR and DORIS Solution for Station Coordinates and ERPs

Min Li, Tian-he Xu and Hang Yu

Abstract This paper mainly focuses on the combination of station coordinates and ERPs and the realization of TRF based on four technologies GPS, VLBI, SLR and DORIS, and discussed some key points including the combination model, datum definition methods and strategies to the validity of solutions in detail. By comparing our results with ITRF08, the difference in determining the center of mass is 2.4–7.7 mm, the uncertainly error is 1.25 mm; the consistently in scale is 1.10 ± 0.06 ppb; small difference still remains in rotation because of the uncertainly of solutions. As to ERP, the combined solution showed a little difference with C04 and the optimal solution in xpo and ypo, the WRMS compared to C04 is 0.077 and 0.055 mas separately, LOD and UT are fused with DORIS solutions, the combined solution of LOD generally keep a consistency with C04, the WRMS is 0.026 ms, but UT show a large difference and unconsistence with C04 and the optimal solution because it could only be determined by VLBI.

Keywords Multi-technique combination · Terrestrial reference frame · Datum definition · Local-ties · Scale factor

M. Li (✉) · H. Yu
Chang'an University, Xi'an, Shanxi, China
e-mail: liminbmw760@163.com

T. Xu
State Key Laboratory of Geo-Information Engineering, Xi'an, Shanxi, China

T. Xu
State Key Laboratory of Geodesy and Earth's Dynamics, Wuhan, Hubei, China

T. Xu
Xian Research Institute of Surveying and Mapping, Xi'an, China

15.1 Introduction

The establishment of International Terrestrial Reference Frame (ITRF) is based on four techniques-GPS (Global Positioning System), VLBI (Very Long Baseline Interferometry), SLR (Satellite Laser Ranging) and DORIS (Doppler Orbit determination and Radio positioning Integrated on Satellite). The observation data with multi-years are analyzed and the corresponding accumulated solutions with multi-years including station coordinates and ERPs are realised by each technique analysis-center, then their solutions are combined according to their property and advantage at normal equation level by IERS (International Earth Rotation and ReferenceSystems Service), we may get a high accurate and long-term stable terrestrial reference frame as well as the corresponding high accurate EOP series. At present, the latest TRF is ITRF08, it's accuracy is better than ITRF05 after resolving the previous data with the latest model and software, the long-term stability of its origin can approach 1 cm, and the accuracy of scale factor is better than 1.2 ppb (parts per billion, 10^{-9}).

This paper mainly focuses on the combination of station coordinates and ERPs based on four techniques, and discussed some key points including the combination model, datum definition methods and the optimization of co-locations coordinates. As to the difficulties including the balance between the constraints added to each technique and the acquisition of optimal solution, we have found a better way to fuse each technique and their co-locations after trying several projects and get better results while comparing with ITRF08.

15.2 Combination Model

The solution in week or day of each technique contains a reference datum which can be called TRF, the single TRF and the combined TRF can be connected by Helmert 14 parameters that are 7 datum transformation parameters and their velocities, then, we may get the combination model in regard to station coordinates and velocities [1], detailed description of the formula please refer to the first reference context.

15.3 Datum Definition for TRF

As we have discussed before, datum definition means determination of fourteen parameters including three translations' (origin definition), one scale and three rotations' (direction definition) and their velocities if we consider the change of TRF. ERP connect celestial reference frame and terrestrial reference frame, to solve ERP with any technique means definition and use of TRF, and the transformation

between one TRF solved by one technique and the other TRF and combined TRF is achieved by the above 14 parameters which are directly connected to the transformation with different ERP series.

In fact, different techniques have different abilities for TRF monitoring and realization. GPS, SLR and DORIS are satellite techniques and they are sensitive to origin and scale, VLBI can get a higher accuracy than other techniques on the determination of scale [2]. The direction cannot be determined by any techniques and it's made by given agreements. So we need to define a datum for the combined TRF.

On this paper, the datum is defined according to the strategies of ITRF2008 realization [3]:

- (1) Origin: the three translation parameters and their velocities between the combined TRF and SLR's are set to zero at the reference epoch.
- (2) Scale: the scale factor parameter and its velocity are set to zero between the combined TRF and the average of VLBI's and SLR's at the reference epoch.
- (3) Direction: the three rotation parameters and their velocities on the core stations between the combined TRF and ITRF2008 are set to zero at the reference epoch.

15.4 Constraints for Datum Definition

The datum of ITRF is defined with proper constraints and they are mainly strong constraints, loose constraints, minimum constraints, inner constraints and regularized constraints [1, 4], which and how to use constraints is hard to determine and absolutely important to the accuracy of TRF.

On this paper, the inner constraints and minimum constraints are used for intra-technique and inter-technique combination separately.

15.4.1 *Minimum Constraints*

Minimum constraint is added to each technique before inter-technique combination. As VLBI cannot determine the origin and direction, so the constraints are added and the translation and rotation parameters between the VLBI solution and ITRF are set to zero on the core stations so that its origin and direction are linked to ITRF but scale; SLR cannot determine the direction, so the rotation parameters are set to zero and the direction is linked to ITRF but origin and scale; it's the same with GNSS and DORIS, their seven parameters are all set to zero.

15.5 Determination of the Relative Weight Factor

When it comes to the liner L-S adjustment, the post-processing variance factor can be regard as the full assessment of a sub-net, and it's based on known relative weight factor of different normal equations. While in inter-technique combination, even though the normal equations of a sub-net is known, the relative weight factor of different techniques' normal equations is still unknown, that's to say, they need to be determined [5].

As to the inter-technique normal equation systems, the rescaling model [6] can be expressed as:

$$\left(\sum_{i=1}^4 \frac{1}{\sigma_i^2} A_i^T P_i A_i \right) \cdot x = \sum_{i=1}^4 \frac{1}{\sigma_i^2} \cdot A_i^T P_i l_i \quad (15.1)$$

The coefficient and constant matrix are shown as $A_i^T P_i A_i$ and $A_i^T P_i l_i$. $1/\sigma_i^2$ is the weight factor, σ_i is the primary variance factor which can be determined using Helmert variance components estimation:

While iteration after k times, we have:

$$\sigma_i^{2(k+1)} = \frac{v_i^{T(k)} \cdot P_i \cdot v_i^{(k)}}{r_i^{(k)}} \quad (15.2)$$

$$r_i^k = n_i - \frac{1}{\sigma_i^{2(k)}} \cdot \text{tr} \left(\left(\sum_{j=1}^4 \sigma_j^{2(k)} \cdot A_j^T P_j A_j \right)^{-1} \cdot A_i^T P_i A_i \right) \quad (15.3)$$

$$v_i^{T(k)} \cdot P_i \cdot v_i^{(k)} = x^{T(k)} \cdot A_i^T P_i A_i \cdot x^{(k)} - 2 \cdot x^{T(k)} \cdot A_i^T P_i l_i + l_i^T P_i l_i \quad (15.4)$$

Among the above three formulas, v_i is the residual vector between the combined solution and sub-net i , the weight matrix, degree of freedom, number of observations are shown as P_i , r_i , n_i , x is the unknown parameters of the combined solution. Just keep iteration until the ratio of the post-processing variance factor between each sub-net get close to 1. Currently many analysis-centers have adopted this method to rescale the sun-net, such as IVS combination center BKG, IGS analysis-center and ITRF combination center IGN, Bernese software [7] and so on.

In fact, while inter-technique combination, the iteration is hard to converge because of the differences of number of stations and accuracy of each sub-net, so usually the methods of Helmert variance component estimation and experienced weighting are combined [8] which will be discussed later.

15.6 Selection of Co-location Sites and Local-Ties

While inter-technique combination, a large number of high accuracy and well distributed co-location sites together with local-ties are particularly important for the realization of ITRF.

Co-location sites [1] refer to stations set up at the same site or close together with two or more space geodetic techniques, their distance is called local-ties which is measured with classic method or GPS.

The accuracy of co-location sites and local-ties is crucial to ITRF, they can unify all the TRFs and thereby eliminating the systematic errors between different techniques. Usually the distance is at several hundreds magnitude, in fact it can reach 30 km due to insufficient number of the uneven distributed stations, so we need to establish the tracking network and resurvey the local-ties [1]. Currently the uncertainty of the accuracy of local-ties can reach 1–3 mm (some more than 3 mm), they are supplied in SINEX format with variance–covariance matrix information.

15.7 Repeatability of Station Coordinates

The stability of TRF can be measured by repeatability of station coordinates [9], the repeatability refer to the precision a particular TRF can reach with another set of observation data under the same datum. On this paper, we have computed the weekly or daily repeatability for single technique and the combination which are connected with Helmert 14 transformation parameters, and the post residuals v are used to compute the repeatability:

$$r_{ista} = \sqrt{\frac{\sum_{i=1}^{nsol} v_i^2}{nsol - 1}} \quad (15.5)$$

$nsol$ is the number of solutions for station $ista$ number i , note that residuals v_i and repeatability σ_i contain three components of coordinates, $r_{ista,North}$, $r_{ista,East}$, $r_{ista,Height}$ need to be computed separately, so we still need to compute the repeatability for each component r_{sol} , it's expressed as:

$$r_{sol,North} = \sqrt{\frac{\sum_{ista=1}^{nsta} n_{ista} \cdot r_{ista,North}^2}{\sum_{ista=1}^{nsta} n_{ista}}} \quad (15.6)$$

n_{ista} is the number of techniques co-located at the site, the repeatability is the weighted average value consider all the stations co-located together. Solutions of 1 year are sufficient for the determination of the repeatability of three components of all the stations.

15.8 Combination Analysis

15.8.1 Characteristics of Single Technique Solution

It's necessary to analysis each technique solution before combination, the reference epoch is day 184.5 (day of year) year 2013; their characteristics are shown in Table 15.1:

15.8.2 Intra-technique Combination

For one technique, the process mainly include gross error detection and rejection, unification of a priori value, parameter elimination, normal equations reconstruction, variance component estimation and rescaling factor determination, different kinds of minimum constraints attached with ITRF solutions added to each technique, normal equations stacking and computation.

The covariance matrix of IGS solutions in GPS week 1765, 1769 and 1770 are abnormal, and they need to be re-analyzed.

The four-technique solutions are all based on ITRF08 excepted that UT and LOD for DORIS are UT1R-UTC and LODR, they are UT1-UTC and LOD but deducted with harmonic tidal item of the solid Earth, the corresponding correction formula can be refered to IERS conventions 2010 [10].

Table 15.1 Characteristics of single technique solution

Technique	GNSS	VLBI	SLR	DORIS
Data input	Weekly	Daily (discontinuities)	Weekly	Weekly
Number of stations	616	99	95	101
Solution type	Variance–covariance	Normal equations	Variance–covariance	Variance–covariance
Constraint	Minimum	None	Loose	None
ERPs	Pole motion (velocity), LOD	Pole motion (velocity), LOD, UT	Pole motion, LOD	Pole motion (velocity), LOD, UT

IGS solutions have already been added with minimum constraints (translation, rotation and scale); IVS solutions are free of datum and we can add any constraints if necessary; ILRS has released two kinds of solutions including *ilrsa* (official solution, provided by main combination center ASI) and *ilrsb* (available from alternate combination center DGFI), as they differ in precision, we combine them. DORIS has no official combination solutions, solutions with loose constraints are provided by five analysis centers and their products are combined by us. Due to the particularity and less number of stations of SLR, we can't get a stable and high accuracy solution by adding minimum constraints (such as NNR condition), therefore loose constraints (usually the standard deviation with 1 m is enough) are added [11]. Finally VLBI, SLR and DORIS are added with minimum constraints according to their own properties we have mentioned before, in addition the systematic errors from analysis centers for SLR and DORIS need to be excluded.

15.8.3 Inter-technique Combination

Since we only consider solutions for 1 year, the velocity of station is not computed, thus Helmert 7 parameters are considered. Then we need to change ERP intervals, define datum, co-adjustment with local-ties information, normal equations stacking and computation.

Taking into account that the ERP intervals differ in various techniques, especially the day interrupted solution for VLBI, thus we need to change the intervals for all techniques to 12 h, at the same time, polar motion parameters need to be added with loose constraints, usually standard deviation with 100 mas is enough [7].

For co-location sites, the following problems are existing currently [1]:

- (1) Instruments change frequently leads to discontinuities of GPS station coordinate series, and it has limited the reliability of local-ties greatly;
- (2) VLBI and SLR co-located sites are not well averaged distributed;
- (3) The number of VLBI and DORIS co-located sites is very small and the data is also limited.
- (4) Parts of SLR and GPS stations are not co-located together;
- (5) The number of SLR and DORIS co-located sites is very small.

So, we have collected all the co-location sites used for ITRF08 realization, a total of 285 co-location sites stored in 53 SINEX files [12], but how to use them properly is worth studying and we will discuss in detail below.

15.9 Combination Results

15.9.1 Weight Factor

In order to get statistically satisfactory combined solution, we input each post variance factor which is used by Helmert variance component estimation together with experienced weighting to adjust each technique solution. After computation, the initial variance factor has turned to be: GPS 0.034, VLBI 0.67, SLR 8.12, DORIS 4.05, so the scale factor for each technique is 29.41, 1.49, 0.12, 0.25, GPS and VLBI are up weighted, SLR and DORIS are down weighted, the final variance factor for the combined solution is 5.75.

The stations selected are essential to obtain a stable and reliable weight factor, the standard deviations of some non-core stations are large and that will significantly down the weight factor, therefore the stations are excluded. The main reason of large variance factor for SLR and DORIS is discontinuity and big jump of their station coordinates, such as fewer numbers of observation days and poor solution accuracy, like SLR sites 7110 and 7403 (located at southern hemisphere), DORIS sites BETB, GR3B, JIVB, SANB, STJB and STKB.

15.9.2 Selection of Co-location Sites

Since only pole motion and local-ties connect four techniques, they have a direct impact on precision of station coordinates, datum parameters and variance factor, therefore, we have to consider several points while selecting co-location sites: (1) precision of local-ties; (2) effect of co-location systematic errors; (3) effect of co-location sites on datum parameters and ERP [3, 11].

After checking the precision of ITRF2008 local-ties, the results are shown in Table 15.2:

For establishing and maintaining high-precision terrestrial reference frame, the tie discrepancies between different technique reference points should be less than 10 mm [13], obviously quite few co-location sites meet requirements.

Local-ties inevitably contain systematic errors caused by GPS. We can find errors which can reach 2.3 mm related to un-standardized radoms covered on GPS

Table 15.2 Accuracy of Local-ties between different techniques (G/V/S/D: GPS/VLBI/SLR/DORIS)

Technique	G-V	G-S	G-D	V-S	V-D	S-D
Number	36	35	38	27	18	17
Local-ties (mm)	<6	17	15	13	2	2
	6-10	9	10	5	4	3
	>10	10	10	20	21	12

antenna in elevation at sites co-located with GPS such as COR1 and NYA1 by investigating the residuals of local-ties [3, 14, 15].

Although the precision of some local-ties meet requirement, unavoidable deformation of local reference frame will occur when integrated with techniques differ large in precision [11]. For example statistics show that the precision of three components of coordinates of co-location sites are 1.8 and 10–20 mm for SLR and VLBI, these sites will absolutely cause deformation of SLR and VLBI nets if they are used, so it’s particularly important to rescale the normal equations of co-location sites, and at the same time, more additional factors are waiting to be considered.

In order to prevent deformation of nets cause by local-ties, we have designed five projects, the arguments are rescaling factors of local-ties, the variables are similar-linked transformation residuals RMS between each technique solution and the combined solution [11, 16, 17], 4 parameters transformation residuals RMS for polar motion between GPS and the combination, difference of scale factors of the combination contributed by VLBI and SLR separately, the variance factor of the combination. We select polar motion as it’s the only parameter connecting four techniques and GPS is best for its determination; scale is most sensitive to deformation of local reference frame [3]; variance factor is related to deformation residuals. The five rescaling factors are: ratios between minimum stds of co-location sites and variance factor of single TRF, ratios between maximum stds of co-location sites and variance factor of single TRF, selection of local-ties with precision 0.1, 0.5 and 1 mm. the results are in Table 15.3:

We can see from the above results:

- (1) The appraisal results from four parts show that we can obtain smaller deformation residuals of local reference frame, polar motion and scale differences and variance factor if we select the ratio between the maximum stds of co-location sites and variance factor of local reference frame or the local-ties with precision 0.5 mm.
- (2) Higher or lower precision of local-ties will effect the deformation of TRF, especially the local-ties with precision 0.1 mm can cause deformation of VLBI

Table 15.3 Results of five projects (G/V/S/D: the same as Table 15.2, C: the Combination, P: Projects)

	Similar transformation residuals RMS/mm				Polar motion differences (G-C) RMS/mas		Scale difference/ppb	Variance factor
	G-C	V-C	S-C	D-C	xpo	ypo		
P1	2.3	2.9	6.1	5.1	0.083	0.058	0.55	5.50
P2	1.6	1.4	3.9	2.2	0.069	0.050	0.41	3.15
P3	2.4	2.9	7.4	4.1	0.091	0.070	0.52	22.55
P4	1.2	1.6	3.4	1.9	0.070	0.039	0.39	6.25
P5	2.1	2.7	6.6	2.3	0.066	0.059	0.33	2.20

and SLR nets, as the variance factor significantly became larger; the precision of local-ties also effects the scale of the combined TRF, that's to say scale connects local-ties tightly.

15.9.3 Datum Parameters

The origin of the combined TRF is defined by SLR, the translation parameters at reference epoch 13:184.5 are shown in Table 15.4, and we can see that it's quite small, the maximum value is 0.22 mm; scale differences defined by the average of VLBI and SLR compared with each of them are shown in Table 15.4, it's obviously small. In order to evaluate the precision of origin and scale, we compare our combined solutions with ITRF08 and find that the consistency level is 1.25 mm. As to direction, we select 179 core stations used by ITRF08 including 107 GPS stations, 27 VLBI stations, 15 SLR stations and 12 DORIS stations [3], small variation of rotation parameters is found when comparing with ITRF08, that is mainly related to uncertainty of solutions [18]. Detailed comparison results are shown in Table 15.5.

For a good TRF the similar transformation residuals should be zero (or very small) with its inner TRFs. Finally to the combined TRF in order to reflect the internal precision and find whether deformation occurs, we have compared it with four techniques using similar transformation, RMS of residuals are shown in Table 15.6, the deformation between the four-single TRF and the combination is quite small, SLR and DORIS are larger, maybe it's because of their lower-precision of solutions.

Table 15.4 Translation and Scale difference between single technique and the combination

	T_X	T_Y	T_Z	Scale
SLR	-0.09 ± 0.13	0.22 ± 0.13	0.17 ± 0.13	0.91 ± 0.30
VLBI	–	–	–	-0.25 ± 0.28

Table 15.5 Datum difference between ITRF08 and the combination

T_X /mm	T_Y /mm	T_Z /mm	D/ppb	Rx/mas	Ry/mas	Rz/mas
$-2.46 \pm$	$-2.90 \pm$	$7.71 \pm$	$1.10 \pm$	$0.03 \pm$	$0.03 \pm$	$0.04 \pm$
1.25	1.25	1.25	0.06	0.03	0.03	0.03

Table 15.6 Similarity-linked transformation residuals between single technique and the combination

Technique	3D RMS (mm)
GPS	2.05
VLBI	2.22
SLR	5.69
DORIS	3.34

15.9.4 Station Coordinates

The stability of TRF can be measured by repeatability of station coordinates [9], parts of co-location sites of single technique and the combination are shown in Fig. 15.1, by comparison we can see:

- (1) Repeatability of station coordinates for GPS changes a little before and after the combination, they almost are at the same level, poor repeatability for VLBI, SLR and DORIS before combination and big difference among the stations are explained: a limited number of stations and their uneven distribution; few observation days resulting in station coordinates discontinuities, i.e., only about 20 on average among the almost more than 50 SLR stations worldwide can be used for daily observation and several stations appear quite few times; uneven distribution of VLBI and SLR stations for north and south hemispheres (especially for VLBI) resulting in very poor repeatability in North direction of several stations located at south hemisphere. After combination we get a good repeatability of worldwide stations, the three components of all the stations are more balanced (except few jumps in elevation direction), mainly due to a more uniform distribution of both hemispheres stations, technique fusion has formed advantages complementary, the differences in precision between four techniques are partially eliminated and the stability of solutions are obviously improved.
- (2) The repeatability of GPS co-location sites has been deteriorated and the stability has been reduced but not VLBI and SLR after combination, the reason is probably that after the determination of origin by SLR, the other techniques are all effected by SLR and the three components have made a certain degree of translation especially for GPS that can be seen in Table 15.7; further more the repeatability and stability in direction N, E and U are all greatly improved for DORIS, which maybe the result of co-adjustment with GPS, VLBI and SLR co-location sites [19].
- (3) We find that station coordinates can't well be determined by single technique will improve a lot in precision after combination such as GPS station YAR2, VLBI station 7224, most SLR stations and almost all the DORIS stations.
- (4) Totally speaking, the repeatability of the four technique co-location sites in direction horizon and elevation is within 5 mm, an average improvement of 3.29 and 2.81 mm in North and East, direction elevation is within 15 mm, an average improvement of 0.98 mm.

Fig. 15.1 Repeatability of station coordinates for single technique and the combination (from *left to right*: 22 GPS stations, 6 VLBI stations, 10 SLR stations, 6 DORIS stations)

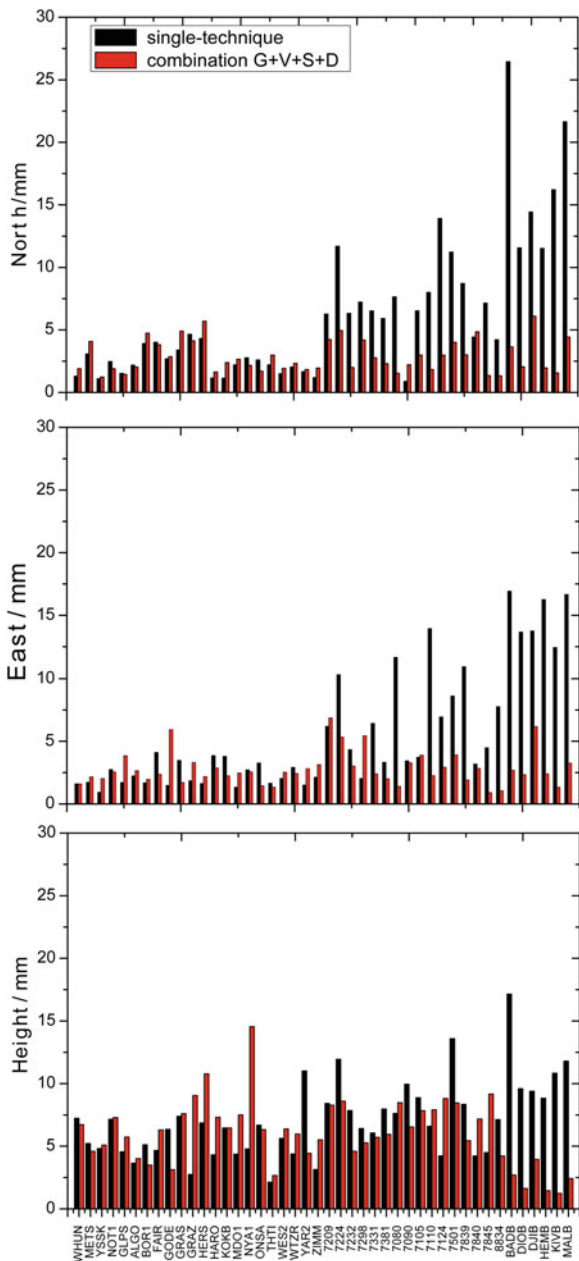


Table 15.7 Repeatability of station coordinates for four techniques and the combination

Repeatability technique/co-location sites	Single technique (mm)			Combination (mm)		
	North	East	Height	North	East	Height
GPS (22)	2.41	2.30	5.40	2.74	2.56	6.41
VLBI (6)	7.32	5.45	8.11	3.41	4.19	6.40
SLR (10)	7.27	7.49	7.51	2.61	2.45	7.41
DORIS (6)	16.97	14.96	11.27	3.29	3.04	2.24

15.9.5 ERPs

Because of the inconsistent resolution of ERPs such as half week for VLBI, in order to compare and evaluate the precision and stability of ERPs, we have transferred the resolution into half-day by parameter transformation and adding additional constraints before normal equations stacking, LOD and UT1-UTC is interpolated into daily resolution. The optimal solution of single technique and the combination are compared to EOP C04/IERS2010; the results are shown in Figs. 15.2, 15.3, 15.4 and 15.5, weighted RMS of statistical results for unbiased differences (systematic errors excluded) compared to the official solution are shown in Table 15.8:

We can see from the results: GPS solution is official and can make a good comparison with other techniques, its polar motion parameters are in consistent with IERS08 C04. Our LOD and UT solution is somehow certain of inconsistent with C04, as the VLBI observations are discontinue we have introduced the DORIS

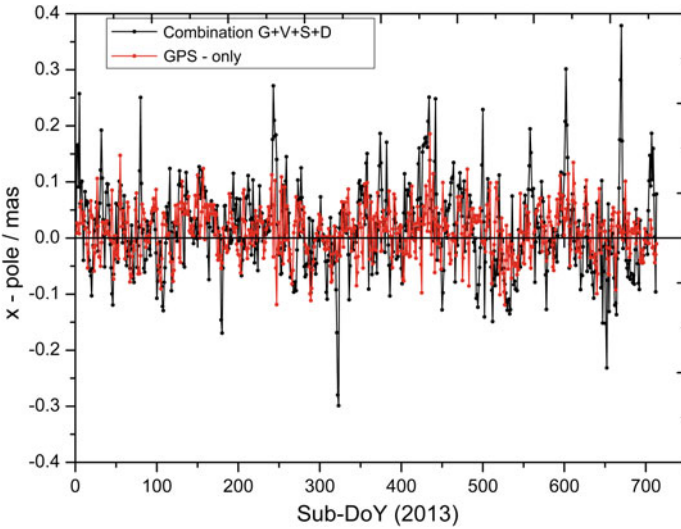


Fig. 15.2 Difference between C04 and GPS solution and the combination for XPO

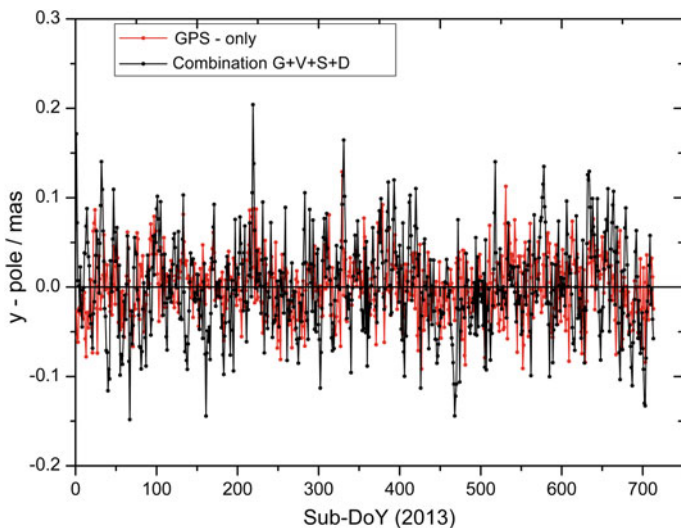


Fig. 15.3 Difference between C04 and GPS solution and the combination for YPO

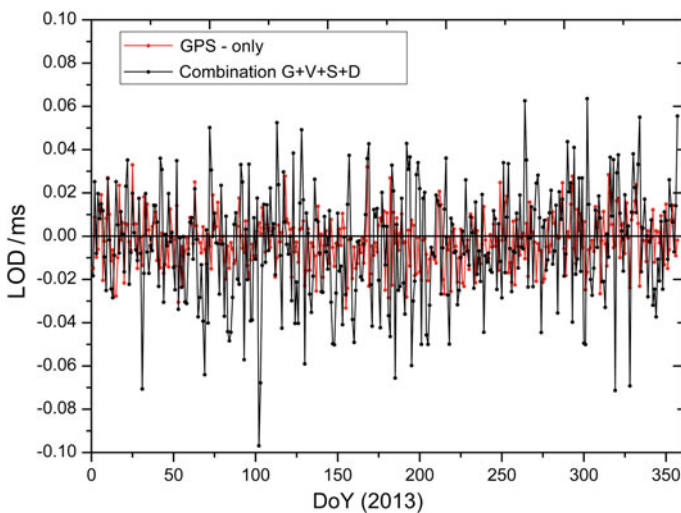


Fig. 15.4 Difference between C04 and GPS solution and the combination for LOD

solutions which resulting in inconsistency with C04 though pre-processing of systematic errors while ITRF08 only takes VLBI solutions and thus the systematic errors related to satellite technology are fully avoided.

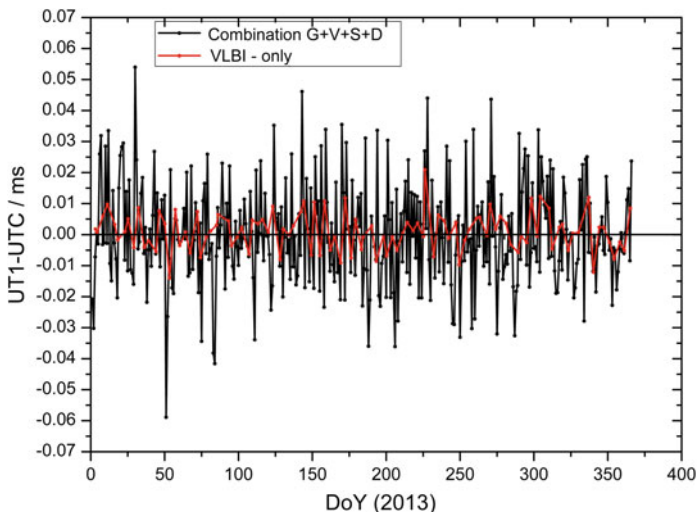


Fig. 15.5 Difference between C04 and VLBI solution and the combination for UT1-UTC

Table 15.8 Difference between C04 and the optimal solution and the combination for XPO, YPO, LOD and UT

Parameter	Technique	Max	Min	Mean	WRMS
XPO (mas)	GPS	0.182	-0.125	0.011	0.046
	VLBI	0.298	-0.235	0.048	0.123
	SLR	0.580	-0.440	0.118	0.208
	DORIS	0.600	-0.592	0.220	0.224
	Com	0.391	-0.301	0.015	0.077
YPO (mas)	GPS	0.126	-0.102	0.001	0.035
	VLBI	0.228	-0.198	0.065	0.117
	SLR	0.578	-0.558	0.040	0.163
	DORIS	0.594	-0.599	-0.046	0.258
	Com	0.209	-0.149	-1.5E-4	0.055
LOD (ms)	GPS	0.034	-0.033	-0.002	0.012
	VLBI	0.056	-0.060	0.003	0.022
	SLR	0.092	-0.096	-0.003	0.030
	DORIS	0.087	-0.085	0.004	0.028
	Com	0.065	-0.099	-0.004	0.026
UT1-UTC (ms)	VLBI	0.021	-0.015	0.001	0.006
	DORIS	0.071	-0.069	0.008	0.021
	Com	0.055	-0.059	-8.4e-5	0.017

15.10 Summary

On this paper, we have mainly discussed the strategy of optimal combination, the accuracy and consistency of the combined solution in every aspects, our results are in good accordance with ITRF08 in station coordinates, ERPs (except for and UT) and datum parameters, as we are the first time for GPS/VLBI/SLR/DORIS combination, the following questions are waiting to be solved:

- (1) Impact of excessive constraints. We have added a variety of loose constraints to each technique before combination, such as SLR and DORIS, considering two facts: quite small number of stations and unstable of TRF; making up the defect for each technique in datum definition, so we have added minimum constraints without eliminating the loose constraints that may enhance the solutions but also effect or change its own datum information (for example attaching the translation, rotation and scale parameters to ITRF08 with minimum constraints).
- (2) Errors of interpolating ERP series. While ERP series linear interpolation and smooth, we don't consider the impact of high-frequency white noise and have to analysis the impact of such errors on the combined TRF.
- (3) Co-location sites selection. At present so many urgent problems for co-location sites: few sites, uneven distribution, low accuracy of local-ties, observation data out of time, so the worldwide sites need to be resurveyed and parts instruments need renovation.
- (4) Construction of nonlinear and mm-level TRF [4, 20]. First the existing space geodesy ground monitoring networks should be glomerated or at least more dense, particularly in the southern hemisphere, refurbishment and upgrading of the equipments; the second is to continue to optimize the combination mathematical model and the methods of benchmarks establishment, adding station seasonal changes and simultaneous analysis its inner multiple geophysical signal.
- (5) Multi-GNSS determining TRF. Although the measurement accuracy of GPS is better than GLONASS/BeiDou/Galileo, we have found in addition to the annual and semi-annual terms, but also hidden 1.04 cpy (cycle per year) term of false signals of nearly 1 year closely related to the GPS satellite constellations [21]. So it's quite important to introduce the other GNSS to reduce GPS technique-related errors, but the multi-GNSS data fusion methods still requires serious consideration.

Acknowledgments This work was supported by Natural Science Foundation of China (41174008) and the Open Foundation of State Key Laboratory of Geodesy and Earth's Dynamics (SKLGED2013-4-2-EZ) and State Key Laboratory of Astronautic and Dynamics (2014ADL-DW0101).

References

1. Zou R (2009) The research of key technology on the realization and maintenance of TRF. Wuhan University, Wuhan (in Chinese)
2. Altamimi Z, Collilieux X (2007) ITRF2005: a new release of the international terrestrial reference frame based on time series of station positions and Earth orientation parameters. *J Geophys Res* 112(B9):B09401, 25–30
3. Altamimi Z, Collilieux X, Metivier L (2011) ITRF2008: an improved solution of the international terrestrial reference frame. *J Geodesy* 85:457–473
4. Liu J, Wei N, Shi C (2013) The latest situation and progress of ITRF. *Nat Inf* 35(4):243–250 (in Chinese)
5. Bähr H, Altamimi Z (2007) Variance component estimation for combination of terrestrial reference frames. Universitäts verlag, Karlsruhe
6. ftp://cddis.gsfc.nasa.gov/vlbi/ivsdocuments/bkg2014a.dsnx_itrf2013.txt. Accessed 10 Nov 2014
7. Dach R, Hugentobler U (2007) Bernese GPS software version 5.0. Astronomical Institute, University of Berne
8. Yingyan C (2012) A brief introduction to ITRF2008. *Geodesy Geodyn* 32(1):47–50 (in Chinese)
9. Thaller D (2007) Inter-technique combination based on homogeneous normal equations systems including station coordinates, Earth orientation and troposphere parameters. Scientific Technical Report, GeoForschungsZentrum Potsdam, 15 Aug 2007
10. Dennis DM, Petit G (2010) IERS conventions 2010 IERS Technical note, No. 32. IERS
11. Seitz M, Angermann D (2012) The 2008 DGFI realization of the ITRS: DTRF2008. *J Geodesy* 86:1097–1123
12. <ftp://itrf.ign.fr/pub/itrf/itrf2008/ITRF2008-TIE-SNX.tar>. Accessed 12 Nov 2014
13. Altamimi Z (2012) ITRF and co-location sites, IERS technical note, No. 33. IERS
14. Sarti P, Sillard P (2004) Surveying co-located space-geodetic instruments for ITRF computation. *J Geodesy* 78:210–222
15. Nan J, Tianhe X, Yan X (2013) Research of receiver radoms on precise positioning. *J Wuhan Univ* 38(5):566–570 (in Chinese)
16. Coulot D, Arnaud P (2010) Global optimization of core station networks for space geodesy: application to the referencing of the SLR EOP with respect to ITRF. *J Geodesy* 84:31–50
17. Drewes H (2009) Reference systems, reference frames, and the geodetic datum-basic considerations. In: Sideris M (ed) *Observing our changing Earth*. IAG Symposia, vol 133, pp 3–9
18. Pesek I, Kostelecky J (2006) Simultaneous determination of Earth orientation parameters and station coordinates from combination of results of different observation techniques. *Stud Geophys Geod* 50:537–548
19. Gambis D (2006) DORIS and the determination of the Earth's polar motion. *J Geod* 80:649–656
20. Wenyao Zhu, Shuli Song, Lina He (2011) Recommendations for construction of a nonlinear international terrestrial reference frame. *Sci China: Phys Mech Astron* 54(1):164–171 (in Chinese)
21. Ray JR, Altamimi Z, Collilieux X et al (2008) Anomalous harmonics in the spectra of GPS position estimates. *GPS Solut* 12:55–64

Chapter 16

Cycle-Slip Detection and Correction Based on Polynomial-Fitting Ionosphere-Free Combination and Ionospheric Total Electron Contents Rate

Ming Liu, Hongzhou Chai and Di Li

Abstract This paper develops a new cycle slip detection and repair method that is based on polynomial-fitting ionosphere-free combination and ionosphere total electron contents (TEC) rate (TECR), and uses dual-frequency GNSS data to present the mathematical model of the two method. It proves that the new algorithm can detect and repair the cycle slips with a very high level of success by the experiments with 1 Hz or even higher rate data, especially the method of TECR. In addition, the combination of the two methods can detect and repair their insensitive cycle slips effectively. Because the new method only use carrier phase, it can not only detect the cycle slips effectively, but also repair them reliably.

Keywords Polynomial fitting · Ionosphere-free combination · Total electron contents rata · Cycle slip detection and correction

16.1 Introduction

Cycle slips occur if the GPS receiver loses phase lock of the satellite signal. However, correcting cycle slip detection is critical for high accuracy of positioning. Now, there are many researches on the cycle slip detection such as Bastos and Landau [1], Colombo et al. [4], Bisnath and Langley [2], and Lee et al. [5]. But some methods are based on the integration of the GPS and INS data, e.g. Colombo et al. [4] and Lee et al. [5], which significantly constrain their feasibility in many applications due to the cost of INS system as well as the complexity of adding an INS system to GPS. A list of the general cycle slip detection methods such as phase-code comparison, phase-phase ionospheric residual, Doppler integration, and differential phases of time have been summarized in Xu [6]. However these methods have their own limitation. The phase-code comparison method is not

M. Liu (✉) · H. Chai · D. Li
Institute of Surveying and Mapping, Zhengzhou 450052, China
e-mail: 291410820@qq.com

effective in repairing small cycle slips (e.g. 1–2 cycles) due to the low accuracy of code measurement. We have tested the Doppler integration method, and like the phase–code comparison, it cannot succeed in small cycle slips. The phase–phase ionospheric residual method, which is essentially the geometry-free linear combination, has a shortcoming of being insensitive to special cycle slip pairs and unable to check on which frequency the cycle slip happen [6].

In this paper, a new cycle slip detection and repair method that employs the ionospheric total electron content (TEC) rate (TECR) and polynomial-fitting ionosphere-free combination is proposed. Unlike other methods where the ionospheric variation is regarded as a nuisance or simply ignored, in this method, the ionospheric variation, characterized by the TEC rate, is precisely estimated and used to detect and repair cycle slip. Without cycle slips, the ionospheric physical TECR is normally bounded by a certain value. In this study, the threshold value that we choose is 0.15 TECU/s. When there are cycle slips, the ionospheric TECR will become significantly larger, with most contribution from the artificial ionospheric TECR which is resulted from cycle slips. The ionosphere-free (IF) ambiguity will also change when there are cycle slips. With the data of both ionospheric TECR change and IF ambiguity change, the cycle slips can be precisely and uniquely determined.

16.2 Polynomial-Fitting Ionosphere-Free Combination

16.2.1 Carrier Phase Observations

The observation equations of dual-frequency GPS carrier phase measurements can be written as:

$$\lambda_1\Phi_1 = \rho - (\delta t_r - \delta t_k)c + \lambda_1 N_1 - \delta_{ion} + \delta_{trop} + \varepsilon \quad (16.1)$$

$$\lambda_2\Phi_2 = \rho - (\delta t_r - \delta t_k)c + \lambda_2 N_2 - \gamma\delta_{ion} + \delta_{trop} + \varepsilon \quad (16.2)$$

The observation equations for IF combination can be written as:

$$\Phi_c = (f_1^2\lambda_1\Phi_1 - f_2^2\lambda_2\Phi_2) / (f_1^2 - f_2^2) = \lambda(f_1\Phi_1 - f_2\Phi_2) \quad (16.3)$$

$$\Phi_c = \rho - (\delta t_r - \delta t_k)c + \lambda N + \delta_{trop} + \varepsilon_c \quad (16.4)$$

$$N = f_1 N_1 - f_2 N_2, \quad \lambda = c / (f_1^2 - f_2^2) \quad (16.5)$$

where λ_1 and λ_2 are the wavelengths of the GPS L1 and L2 signals, respectively; Φ_1 and Φ_2 are carrier phase measurements on L1 and L2 frequencies, respectively; ρ is the geometrical distance between the receiver and the satellite; δt_r and δt_k are the GPS receiver and satellite clock errors, respectively; c is the speed of light in

vacuum; δ_{ion} is the ionospheric range delay on GPS L1 signal; δ_{trop} is the tropospheric range delay; N_1 and N_2 are the integer number of cycles for GPS L1 and L2 signals, respectively, which are often called ambiguities; $\gamma = f_1^2/f_2^2$ is the ratio of the squared frequencies of GPS L1 and L2 signals. The carrier phase measurements Φ_1 and Φ_2 and the ambiguities N_1 and N_2 are in the unit of cycle. The δt_r and δt_k are in the unit of second. The rest variables are all in the unit of meter.

16.2.2 Cycle Slip Determined from the Polynomial-Fitting Ionosphere-Free Combination

The polynomial-fitting ionosphere-free combination can be formed as below.

$$\Phi_c(i) = a_0 + a_1(t_i - t_0) + a_2(t_i - t_0)^2 \quad (16.6)$$

When there are cycle slips,

$$\Phi_c(i) = a_0 + a_1(t_i - t_0) + a_2(t_i - t_0)^2 + \lambda \Delta N_c \quad (16.7)$$

$$\Delta N_c = f_1 \Delta N_1 - f_2 \Delta N_2 \quad (16.8)$$

The specific steps are that use m epochs of carrier observation without cycle slips to fit the polynomial (6) when $m \geq 5$, resolve the coefficient of polynomial by using the least squares method, then calculate the mean square error with residual v_i

$$\sigma(m) = \sqrt{\frac{\sum v_i^2}{m-3}} \quad (i = 1, 2, \dots, m) \quad (16.9)$$

Extrapolate the carrier observation $\Phi'_c(m+1)$ of the next epoch by using the polynomial and compare with the real observation $\Phi_c(m+1)$

$$|\Phi_c(m+1) - \Phi'_c(m+1)| < k \cdot \sigma(m) \quad (16.10)$$

When meeting the above inequality, we consider that there are no cycle slips. Then continue the processing with the above real observation.

$$|\Phi_c(m+1) - \Phi'_c(m+1)| \geq k \cdot \sigma(m) \quad (16.11)$$

When meeting the above inequality, we consider that there are cycle slips. Then restart a new arc and resume the above processing. Combining Eqs. (16.6) and (16.7), the cycle slip can be estimated as;

$$\lambda \Delta N_c = \Phi_c(m+1) - \Phi'_c(m+1) \quad (16.12)$$

The maximum of the window's width is n , when $m > n$, remove the earliest observed value and fit the polynomial with the new one to maintain the maximum of window's width as n .

Because the method of TECR has a high level of success to detect and repair the cycle slips, we can enlarge the restrained condition to avoid mistaken judgement. In this paper, we order $k = 6$. In addition, when a period of observed values are so smooth that makes $\sigma(m)$ very small which is difficult to meet Eq. (16.10), it will lead to mistaken judgement. In order to avoid this situation, we order $\sigma(m) > 0.1$.

However, there are some problems exist in this method. First, the polynomial-fitting method needs five started data without cycle slip. Second, the method alone can't distinguish which frequency has cycle slip. Finally, there are some cycle slip pairs that this method couldn't detect, such as $\frac{\Delta N_1}{\Delta N_2} = \frac{f_2}{f_1} = \frac{60}{77}$.

16.3 Detecting Cycle Slips by Using TECR

From Eqs. (16.1) to (16.2), we can derive the ionospheric total electron contents from dual-frequency carrier phase measurements as below. It is assumed that we use the GPS data from epoch $(k - 1)$.

$$TEC_{\Phi}(k - 1) = \frac{f_1^2 \{ [\lambda_1 \Phi_1(k - 1) - \lambda_2 \Phi_2(k - 1)] - [\lambda_1 N_1 - \lambda_2 N_2] - b_i - b^p \}}{40.3 \times 10^{16}(\gamma - 1)} \quad (16.13)$$

where b_i and b^p , in unit of meters, are the inter-frequency biases of the receiver and the satellite, respectively. Other terms in Eq. (16.13) are the same as those defined in previous equations. The values of satellite and receiver inter-frequency biases are quite stable during a period of a few days. Therefore, they can be treated as constants during the cycle slip detection where the time interval between two consecutive epochs is normally as short as 1 s or at most a few minutes.

We only consider the cycle slip occurrence at epoch (k) . The GPS L1 and L2 carrier phase measurements at epoch $(k - 1)$ are regarded as free of cycle slips or the cycle slips have already been repaired if they indeed occur at that epoch. If the total electron contents $TEC_{\Phi}(k)$ estimated from epoch (k) is differentiated with that of epoch $(k - 1)$, the total electron contents rate (TECR) can be derived.

$$TECR_{\Phi}(k) = \frac{TEC_{\Phi}(k) - TEC_{\Phi}(k - 1)}{\Delta t} \quad (16.14)$$

where $TECR_{\Phi}(k)$ is the TEC rate at epoch (k) ; Δt is the time interval between epochs (k) and $(k - 1)$, which is typically 1 s for high rate GPS observations. Rearranging Eqs. (16.13) and (16.14), the cycle slips at epoch (k) can be estimated as:

$$\lambda_1 \Delta N_1(k) - \lambda_2 \Delta N_2(k) = -\frac{40.3 \times 10^{16} (\gamma - 1) \Delta t \cdot TECR_{\Phi}(k)}{f_1^2} + \lambda_1 [\Phi_1(k) - \Phi_1(k-1)] - \lambda_2 [\Phi_2(k) - \Phi_2(k-1)] \quad (16.15)$$

It is clear that all the terms in the above Eq. (16.15) are known except the TEC rate $TECR_{\Phi}(k)$. If the $TECR_{\Phi}(k)$ is also known, we can then estimate the cycle slip $\lambda_1 \Delta N_1(k) - \lambda_2 \Delta N_2(k)$. It should be noted that in Eq. (16.12) only carrier phase measurements are used; thus, it is expected that the accuracy of the estimated $\lambda_1 \Delta N_1(k) - \lambda_2 \Delta N_2(k)$ will be very high.

As illustrated in Eq. (16.11), the $TECR_{\Phi}(k)$ is estimated from the measurements at epoch (k) and its previous epoch (k - 1). But we cannot use the Eq. (16.11) because we are intending to detect cycle slips at epoch (k) and we are not sure if epoch (k) has cycle slips. The $TECR_{\Phi}(k)$ however can be estimated based on the measurements of the previous epochs. Since we are detecting and repairing cycle slips on an epoch-by-epoch basis, all the epochs prior to the current epoch (k) are free of cycle slips because their cycle slips, if any, have been repaired. Therefore at epoch (k), we can use the previous epochs to estimate the TEC rate for epoch (k). For instance, at epoch (k - 1), we can use the measurements of epochs (k - 1) and (k - 2) to derive $TECR_{\Phi}(k - 1)$. At epoch (k - 1), if the rate of the TEC rate (i.e. TEC acceleration) is also known, the $TECR_{\Phi}(k)$ at epoch (k) can be readily estimated as:

$$TECR_{\Phi}(k) = TECR_{\Phi}(k - 1) + TE \overset{\text{g}}{C} R_{\Phi}(k - 1) \cdot \Delta t \quad (16.16)$$

where $TE \overset{\text{g}}{C} R_{\Phi}(k - 1)$ is the TEC acceleration at epoch (k - 1). The determination of the $TE \overset{\text{g}}{C} R_{\Phi}(k - 1)$ can be performed as below:

$$TE \overset{\text{g}}{C} R_{\Phi}(k - 1) = \frac{TECR_{\Phi}(k - 1) - TECR_{\Phi}(k - 2)}{\Delta t} \quad (16.17)$$

In the practical implementation, both $TECR_{\Phi}(k - 1)$ and $TECR_{\Phi}(k - 2)$ are estimated using measurements of previous epochs. They are averaged to smooth the noise in the measurements [3]. Thus, more accurate $TECR_{\Phi}(k - 1)$ and $TECR_{\Phi}(k - 2)$ can be obtained.

The mean and the variance of the TECR at epoch (k) can be recursively calculated as below:

$$E[TECR_{\Phi}(k)] = \overline{TECR_{\Phi}(k)} = \overline{TECR_{\Phi}(k - 1)} + \frac{1}{k} \left[TECR_{\Phi}(k) - \overline{TECR_{\Phi}(k - 1)} \right] \quad (16.18)$$

$$\sigma^2(k) = \sigma^2(k-1) + \frac{1}{k} \left\{ \left[\text{TECR}_\Phi(k) - \overline{\text{TECR}_\Phi(k)} \right]^2 - \sigma^2(k-1) \right\} \quad (16.19)$$

In this study, we consider the TECR data basically constant over a short period (normally no more than half minute). The TEC acceleration is also calculated to account for the different ionospheric rates between epochs. The TECR calculated at epoch (k), shown in Eq. (16.14), is compared with the one estimated in Eq. (16.17). If their difference, called TECR residual, is within four times of the standard deviation, no cycle slips are assumed at epoch (k). Otherwise, cycle slips are detected, and Eq. (16.15) is used to determine the size of the cycle slip term $[\lambda_1 \Delta N_1(k) - \lambda_2 \Delta N_2(k)]$.

The variance of $[\lambda_1 \Delta N_1(k) - \lambda_2 \Delta N_2(k)]$ can be estimated as:

$$\sigma^2(k) = 12(\lambda_1^2 + \lambda_2^2)\sigma_\Phi^2 \quad (16.20)$$

In Eq. (16.20), the wavelengths λ_1 and λ_2 are in unit of meter/cycle, and the standard deviation σ_Φ is in the unit of cycle. The variance $\sigma^2(k)$ has a unit of m^2 . In practical application, the standard deviation σ_Φ is usually expressed with length unit, e.g. mm after being implicitly multiplied by its wavelength. The carrier phase measurement error inside the receiver is normally small, at about 1 mm level. Considering the noises resulting from multipath and atmospheric effects, the overall carrier phase measurements may have a few millimeters. If we take $\sigma_\Phi = 5$ mm, the standard deviation of the cycle slip term $[\lambda_1 \Delta N_1(k) - \lambda_2 \Delta N_2(k)]$ is estimated to be 5.4 mm. It clearly shows that the cycle slip term $[\lambda_1 \Delta N_1(k) - \lambda_2 \Delta N_2(k)]$ estimated from the TEC rate data has a very high accuracy. This is due to the sole use of high accuracy carrier phase measurements and no pseudorange measurements being used.

However, the method has its own faults either. First, the method alone can't distinguish which frequency has cycle slip. Moreover, there are some cycle slip pairs that this method couldn't detect, such as $\frac{\Delta N_1}{\Delta N_2} = \frac{\lambda_2}{\lambda_1} = \frac{77}{60}$.

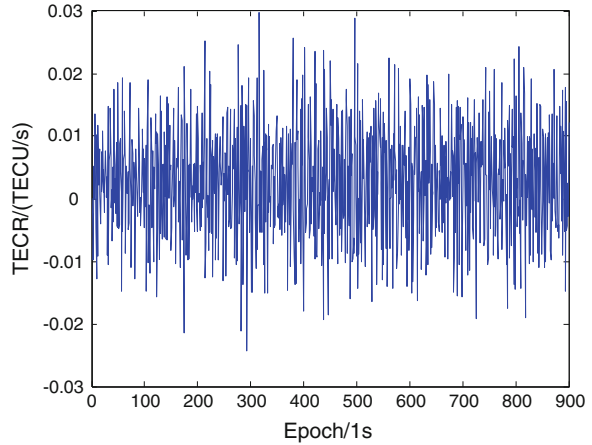
Therefore, we can use the stability of TECR to judge whether there are cycle slips. Then, combining the above two proposed methods can solve all problems.

16.4 Data Test and Analysis

The effectiveness of this method was tested using many sets of GPS observations which were recorded at an interval of 1.0 s. And analysis it by adding some cycle slips that are difficult to detect. The results show that they were essentially consistent. We used the dataset from the station coco on 15 January 2012 as example.

It has been shown that in the equatorial region like Hong Kong, the ionospheric slant TEC rate is about 0.01 TECU/s during quiet ionosphere periods, and it rises to 0.03 TECU/s during active ionosphere period. When the cycle slip pair is $(-1, -1)$, which represents -1 cycle on L1 and -1 cycle on L2 carrier phase measurements,

Fig. 16.1 The TECR of PRN2



the TECR is -0.514 TECU. Therefore we can determine whether there are cycle slips of PRN2 by using TECR, as shown in Fig. 16.1.

In this test, eight pairs of cycle slip are simulated to evaluate the effectiveness of the method. The simulated cycle slips are added to the GPS data that are then processed by our computer program implementing this algorithm. The results were shown in Figs. 16.2, 16.3 and Table 16.1.

The results in Figs. 16.2 and 16.3 clearly showed that combining the two method proposed in this paper can detect all insensitive and difficult cycle slips. The results in Table 16.1 showed that the new method has a higher level of accuracy than M-W and TECR combination method proposed by Liu [7].

Fig. 16.2 The absolute value of IF residual

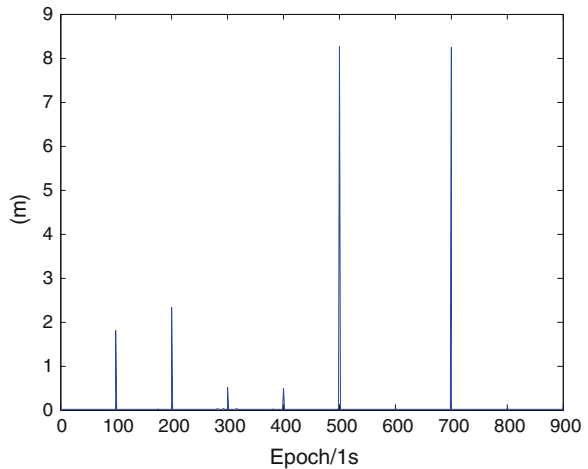


Fig. 16.3 The absolute value of TECR residual

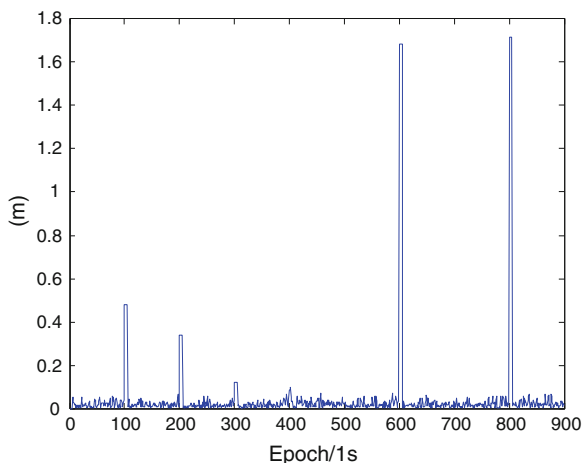


Table 16.1 The correction of cycle slip

Epoch	Cycle slip	IF and TECR combination		M-W and TECR combination	
	Pair	ΔN_1	ΔN_2	ΔN_1	ΔN_2
100	(1, 0)	0.9866	-0.0021	0.8993	-0.0701
200	(0, -1)	-0.1874	-1.1414	-0.4195	-1.3223
300	(1, 1)	1.1002	1.0874	0.7983	0.8522
400	(-1, -1)	-0.8462	-0.8609	-1.2338	-1.1629
500	(7, 9)	7.1708	9.1487	5.7406	8.0343
600	(9, 7)	8.8373	6.8829	8.3760	6.5235
700	(-7, -9)	-6.8723	-8.8899	-8.2063	-9.9293
800	(-9, -7)	-8.9910	-6.9884	-10.2155	-7.9426

16.5 Conclusions

This paper proposed a new method to detect and repair cycle slip, which is based on the joint use of ionospheric TECR and IF combination. The fundamental principle of this new method largely lies in the fact that the physical (natural) ionospheric TECR is normally considerably smaller than the artificial TECR that is caused by cycle slips. If the calculated TECR exceeds our predefined threshold, it is assumed that cycle slips happen, and the cycle slips on L1 and L2 are determined from the changes of both ionospheric TECR and IF ambiguity. The results indicate that the approach is effective in detecting and repairing cycle slips.

Acknowledgments This work is supported by China National Natural Science Foundation of China (No: 41274045).

References

1. Bastos L, Landau H (1988) Fixing cycle slips in dual-frequency kinematic GPS-applications using Kalman filtering. *Manuscr Geod* 13(4):249–256
2. Bisnath SB, Langley RB (2000) Automated cycle-slip correction of dual-frequency kinematic GPS data. In: *Proceedings of 47th Conference of CASI, Ottawa, Canada*
3. Blewitt G (1990) An automatic editing algorithm for GPS data. *Geophys Res Lett* 17(3): 199–202
4. Colombo OL, Bhapkar UV, Evans AG (1999) Inertial-aided cycle-slip detection/correction for precise, long-baseline kinematic GPS. In: *Proceedings of ION GPS-99, Nashville, TN*, pp 1915–1922
5. Lee HK, Wang J, Rizos C (2003) Effective cycle slip detection and identification for high precision GPS/INS integrated systems. *J Navig* 56(3):475–486. doi:[10.1017/S0373463303002443](https://doi.org/10.1017/S0373463303002443)
6. Xu G (2007) *GPS: theory, algorithms and applications*, 2nd edn. Springer, Berlin
7. Liu Z (2010) A new automated cycle slip detection and repair method for a single dual-frequency GPS receiver. *J Geod* 85(3):171–183

Chapter 17

BDS and GPS Ultra-Short Baseline Measurement Performance Comparison and Analysis

Kang Zhang, Jinming Hao, Yu Zhang, Qiang Li and Tuansheng Yang

Abstract Since BeiDou satellite navigation system (BDS), which is independently developed and being implemented in China, successfully realizes the coverage in the Asia-Pacific region at the end of 2012, it has been widely used in the national economy and society, with that the whole system runs stably, and is in good condition. In order to verify the precision level of the positioning solution of the second generation BeiDou system (BD-2) in the condition of short baseline, based on the measured data, through the use of self-written baseline solving program, we get the short baseline calculating results under the condition of the static and dynamic experiments respectively based on BD-2 and GPS. Then through the comprehensive comparison of the above results, we analyze the accuracy and reliability of ultra-short baseline results based on BD-2. The experimental results show that, the reliability of the short baseline solution of BD-2 is rather equal compared with GPS. Under the static condition, the precision of ultra-short baseline solution of BDS is slightly worse than GPS; under the dynamic condition, that of BDS in E and N directions are rather equal to that of GPS, while in the U direction it is not as good as GPS.

Keywords BDS · GPS · Ultra-short baseline · Precision · Reliability

K. Zhang (✉) · J. Hao · Y. Zhang
College of Navigation and Aerospace Engineering,
Information Engineering University, Zhengzhou 450001, China
e-mail: 289147188@qq.com

Q. Li
PLA 61773 Troops, Urumqi 831400, China

T. Yang
Surveying and Mapping Information Center of Guangzhou Military Region,
Guangzhou 510000, China

17.1 Instruction

Currently, research on GPS technology is relatively rich compared to other systems. In China, GPS has been widely used in transportation, agriculture, meteorology and other industries. BD-2, as China's self-developed global satellite navigation system, the rapid development of which will break the monopoly of GPS upon satellite-navigation-related industries in China. Up to 2013, Beidou has been made great progress in the military and civilian aspects, and the latest Beidou military receivers can reach positioning results in centimeter level. Domestic and foreign learners have done a lot of research on GPS high-precision dynamic measurement with long baseline and short baseline [1, 2]. Tang W in Wuhan University has done some tests and analysis upon BD-2 single-epoch baseline solution [5].

In this paper, BDS/GPS dual-system multi-frequency receivers are used to collect the ultra-short baseline data in Zhengzhou, and observation data from different systems were calculated respectively, as to analyze the accuracy and reliability of ultra-short baseline measurement with the use of the Beidou satellite navigation system. And the results of BDS were compared with that of GPS to verify the positioning performance of China's Beidou satellite navigation system at this stage.

17.2 Basic Theory

In ultra-short baseline measurement, since the distance between the base and rover is generally less than 100 meters, with the use of double-difference observables we can eliminate all the errors associated with the baseline length, such as satellite position error, ionospheric delay error, tropospheric delay error and so on [1]. The main source of error is the measurement noise and multipath effects.

Whether static or dynamic baseline solution, it belongs to the relative positioning in essence. In two receiver antennas receiving satellite signals, we use the one receiver antenna single point positioning solution as the coordinates of the base station, while the other one is performed differential calculation. Due to the short baseline length, the accuracy of the coordinates of the base station affecting the final baseline solution results can be basically negligible.

Typically, the double-difference model can be simplified as:

$$\begin{cases} \nabla\Delta\Phi_{ur}^{ij} = \nabla\Delta\rho_{ur}^{ij} + \lambda\nabla\Delta N_{ur}^{ij} + \nabla\Delta\varepsilon_{ur,\Phi}^{ij} \\ \nabla\Delta P_{ur}^{ij} = \nabla\Delta\rho_{ur}^{ij} + \nabla\Delta\varepsilon_{ur,P}^{ij} \end{cases} \quad (17.1)$$

where,

$\nabla\Delta$ represents the double-difference operator

i, j represents the No. of satellite

u, r is the No. of the receiver (station)

Φ_u^j, P_u^j respectively denote the phase and Pseudo-range observables (meter)
 N_u^j is for integer ambiguity (week)
 $\varepsilon_{u,\Phi}^j, \varepsilon_{u,P}^j$ respectively denote carrier phase noise and pseudorange noise (meter)

Suppose that the number of simultaneous observing satellites is m , we firstly perform single-differential calculation between two stations to the same satellite, so the error normal equation is obtained [6, 8]:

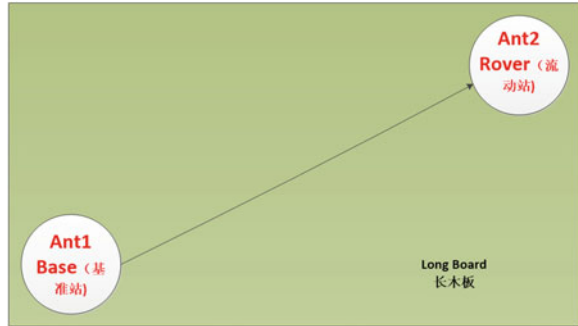
$$V = A \cdot X - L \tag{17.2}$$

In which,

$$\begin{aligned}
 V &= \begin{bmatrix} v_\Phi^1 \\ \vdots \\ v_\Phi^{m-1} \\ v_P^1 \\ \vdots \\ v_P^{m-1} \end{bmatrix}_{2(m-1) \times 1}, \quad X = \begin{bmatrix} \delta x \\ \delta y \\ \delta z \\ \Delta N_{ur}^1 \\ \vdots \\ \Delta N_{ur}^m \end{bmatrix}_{(3+m) \times 1}, \\
 L &= \begin{bmatrix} \nabla \Delta \Phi_{ur}^{i1} - \nabla \Delta \rho_{ur}^{i1} \\ \vdots \\ \nabla \Delta \Phi_{ur}^{im} - \nabla \Delta \rho_{ur}^{im} \\ \nabla \Delta P_{ur}^{i1} - \nabla \Delta \rho_{ur}^{i1} \\ \vdots \\ \nabla \Delta P_{ur}^{im} - \nabla \Delta \rho_{ur}^{im} \end{bmatrix}_{2(m-1) \times 1}, \quad \text{and} \\
 A &= \begin{bmatrix} -e_x^i + e_x^1 & -e_y^i + e_y^1 & -e_z^i + e_z^1 & -\lambda^1 & & \lambda^i \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ -e_x^i + e_x^{i-1} & -e_y^i + e_y^{i-1} & -e_z^i + e_z^{i-1} & & -\lambda^{i-1} & \lambda^i \\ -e_x^i + e_x^{i+1} & -e_y^i + e_y^{i+1} & -e_z^i + e_z^{i+1} & & & \lambda^i & -\lambda^{i+1} \\ \vdots & \vdots & \vdots & & & \vdots & \ddots \\ -e_x^i + e_x^m & -e_y^i + e_y^m & -e_z^i + e_z^m & & & \lambda^i & & -\lambda^m \\ -e_x^i + e_x^1 & -e_y^i + e_y^1 & -e_z^i + e_z^1 & & & & & \\ \vdots & \vdots & \vdots & & & & & \\ -e_x^i + e_x^m & -e_y^i + e_y^m & -e_z^i + e_z^m & & & & & \end{bmatrix}_{2(m-1) \times (3+m)}
 \end{aligned}$$

where,
 e_x^i, e_y^i, e_z^i are the sight line vector components between the rover and the No. i satellite
 ΔN_{ur}^i is for inter-stations single difference ambiguity of the No. i satellite

Fig. 17.1 Antenna configuration of ultra-short baseline



According to the principle of least squares, the corrections of the rover's unknown parameters are obtained [4]:

$$\delta X = (A^T P A)^{-1} A^T P L \quad (17.3)$$

Through the above formula, the float single-difference ambiguities can be solved, and the double-difference ambiguity can be obtained by doing subtraction between the two of single-difference ambiguity, then the integer ambiguity solution can be searched using LAMBDA algorithm [6, 7, 9].

Any two antennas compose a baseline. With the integer ambiguity estimation is successful, the baseline vector solution can be calculated out. As shown in Fig. 17.1, the baseline vector points to the 2nd antenna (Ant2) with the first antenna (Ant1) as the origin of which.

17.3 Experiment Designing and Data Processing

The experiment is carried out on October 11th, 2014 in the experimental field of satellite navigation of Information Engineering University, which is divided into static and dynamic experiments, where three different baseline lengths (all less than 20 m) are set respectively. Each session lasts about 30 min with the data sampling interval of one second. The experimental site is chosen in the open while the satellite elevation cut-off angle is set to 15 degrees in order to reduce the influence of multi-path effects. The Trimble R8 receivers produced by the US Trimble Navigation Company are used for data collection with the receiver antenna of R8-4 internal. All six sessions' data were processed and analyzed, but due to space limitations, this paper only illustrate the results of dynamic and static experiment with baseline length of about 1.3 m.

The experiment listed in this paper is done through fixing the two receivers to both ends of the long board. In the static experiment, the board is fixed on the ground stable, while in the dynamic experiment, the board is rotated and swayed back and forth continuously to simulate the movement. Both dynamic and static

experiment adopted the same measurement, and the collected multi-frequency data of BDS and GPS were both processed through afterwards batch algorithm. Due to the short baseline length, during the data processing both tropospheric and ionospheric corrections are not considered, and broadcast ephemeris is used rather than precise ephemeris.

17.4 Experiment Analysis

17.4.1 Satellite Observing Performance Analysis

In both experiments given in the paper, the number of visible satellites are separately shown in Figs. 17.2 and 17.3.

We can see that in the static experiment, BDS remains at nine satellites visible in most cases, while GPS is at seven visible, sometimes at six; in the dynamic experiment, BDS remains at nine satellites visible, while GPS is at seven visible. Overall speaking, in the experimental area, the number of satellites visible of BDS is slightly more than GPS.

Fig. 17.2 Number of visible satellites in static test with baseline length of about 1.3 m

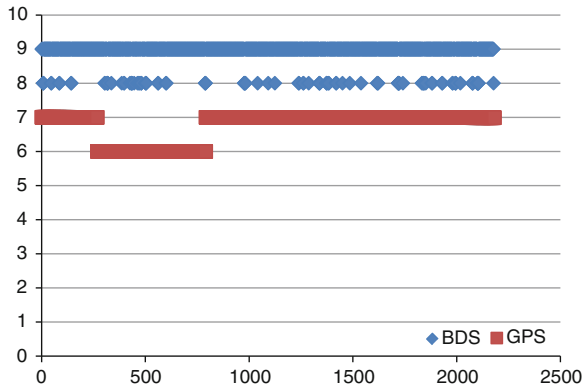
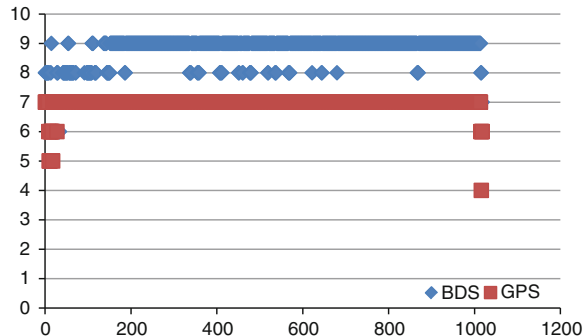


Fig. 17.3 Number of visible satellites in dynamic test with baseline length of about 1.3 m



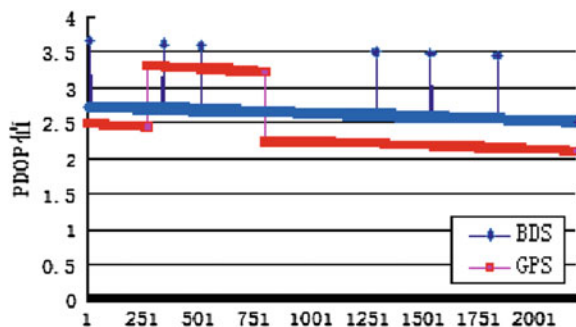


Fig. 17.4 PDOP values of BDS and GPS in static test with baseline length of about 1.3 m

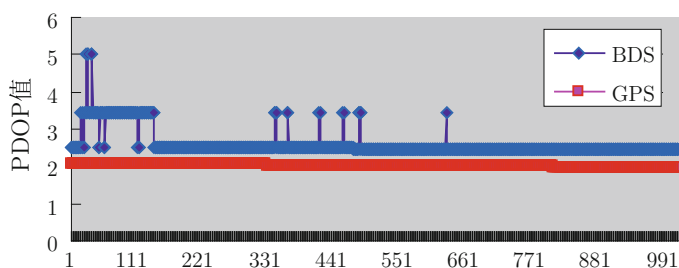


Fig. 17.5 PDOP values of BDS and GPS in dynamic test with baseline length of about 1.3 m

To further analyze the observing performance of BD-2 in Zhengzhou, the PDOP values of BDS and GPS in both static and dynamic experiment are respectively given in Figs. 17.4 and 17.5, and their carrier phase integer ambiguity RATIO values are respectively shown in Figs. 17.6 and 17.7.

Fig. 17.6 RATIO values of ambiguity resolution in static test with baseline length of about 1.3 m

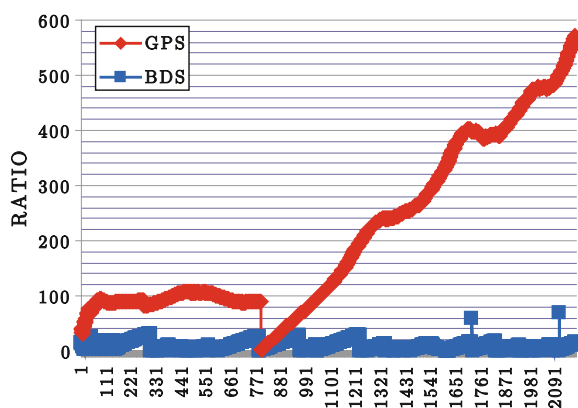
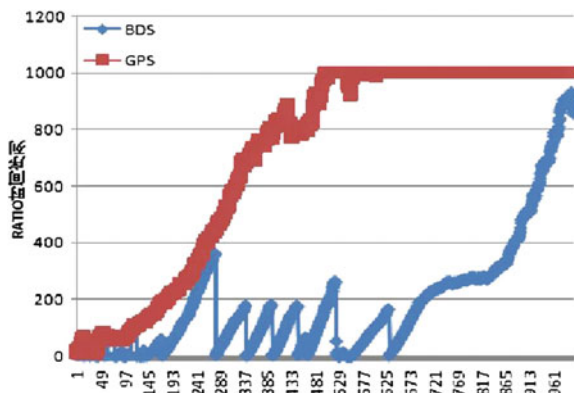


Fig. 17.7 RATIO values of ambiguity resolution in dynamic test with baseline length of about 1.3 m



From Figs. 17.4 to 17.5, it can be seen whether GPS or BDS, the PDOP value throughout the experiment remained between 2 and 3, showing that all the observations are in good condition. PDOP values of GPS tend to be lower than that of BDS, which is also reflected in the following analysis that ultra-short baseline measurement accuracy of GPS is slightly better than that of BDS. In addition it can be seen from Figs. 17.6 and 17.7 that in whether dynamic or static experiment, RATIO values of GPS are higher than that of BDS, explaining GPS ambiguity fixing more successful than BDS, which is caused by the use of a continuous process manner in ambiguity fixing and the existence of heterogeneous constellation of BDS.

17.4.2 Static Measurement Experiment

In the static experiment, entire baseline was fixed on the ground motionless, in order to analyze the static baseline measurement performance of different satellite navigation systems, the baseline measurement performance of multi-frequency of BDS and GPS were compared and analyzed in the following. The frequencies of GPS are L1 (1575.42 MHz) and L2 (1227.6 MHz), while that of BDS are B1 (1561.098 MHz) and B2 (1207.14 MHz) [3]. The measurement accuracy with baseline length of about 1.3 m of BDS and GPS are shown in Figs. 17.8, 17.9, 17.10 and Table 17.1.

From Figs. 17.8, 17.9 and 17.10, it can be seen the measurement results of both GPS and BDS are relatively stable, which are in the same order of magnitude. Though there is no significant difference, the measurement noise of BDS is yet a little bigger compared to GPS. From Table 17.1 it can be seen whether in the direction E, N or U direction, the standard deviation of BDS(E-2) is significantly worse than that of GPS (E-3) with an order of magnitude lower, which indicates that in the static case ultra-short baseline measurement accuracy of BDS is worse than GPS.

Fig. 17.8 BDS and GPS static baseline solution results with baseline length of about 1.3 m in E direction

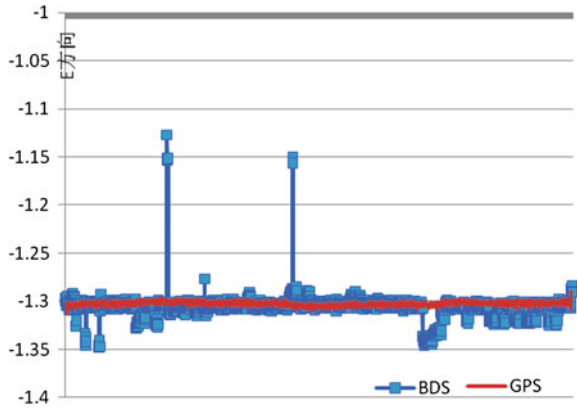


Fig. 17.9 BDS and GPS static baseline solution results with baseline length of about 1.3 m in N direction

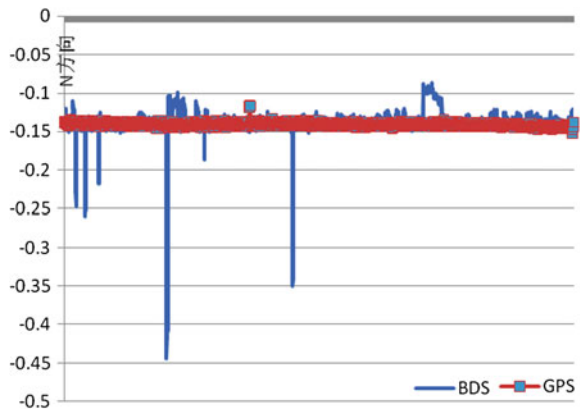


Fig. 17.10 BDS and GPS static baseline solution results with baseline length of about 1.3 m in U direction

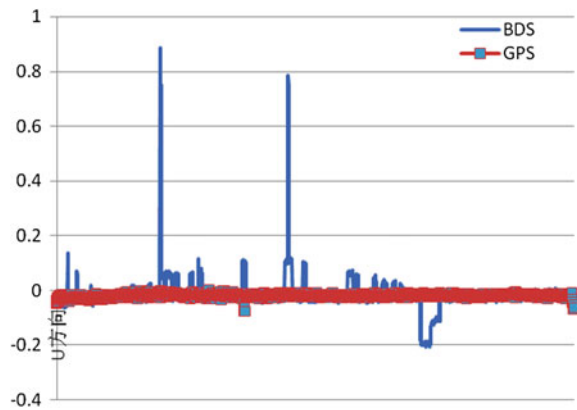


Table 17.1 BDS and GPS measurements and standard deviation in static baseline

	Mean (m)			Standard deviation (m)		
	E	N	U	E	N	U
BDS	-1.306	-0.137	-0.020	0.013	0.020	0.064
GPS	-1.303	-0.141	-0.018	0.002	0.002	0.006

17.4.3 Dynamic Measurement Experiment

In the dynamic measurement experiment, the entire baseline consisted of two receivers was overall shaken, with the baseline length remained unchanged. The dynamic baseline measurement performance of multi-frequency of BDS and GPS were compared and analyzed in the following. The measurement accuracy with dynamic baseline length of about 1.3 m of BDS and GPS are shown in Figs. 17.11, 17.12, 17.13 and Table 17.2.

From Figs. 17.11, 17.12 to 17.13, it can be seen the measurement results of BDS and GPS are both relatively stable, which are in the same order of magnitude, while the measurement noise of BDS is a little bigger compared to GPS. As can be seen from Table 17.2, in the E and N directions standard deviation of the baseline solution of BDS is in the same order of magnitude compared with GPS, while in the U direction that of BDS is an order of magnitude lower than the GPS, which shows in the dynamic case ultra-short baseline measurement accuracy of BDS is equal to GPS in the E and N directions, while worse in the U direction.

Besides, in this paper other baseline observation data collected were also processed and analyzed, the results of which are consistent with the above results, what are not listed due to limited space.

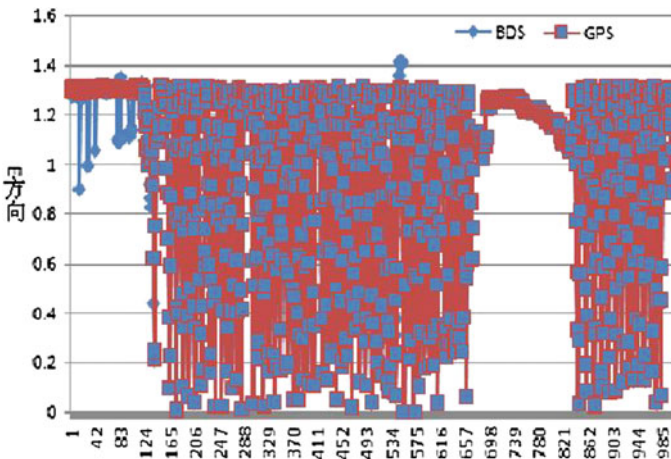


Fig. 17.11 BDS and GPS dynamic baseline solution results with baseline length of about 1.3 m in E direction

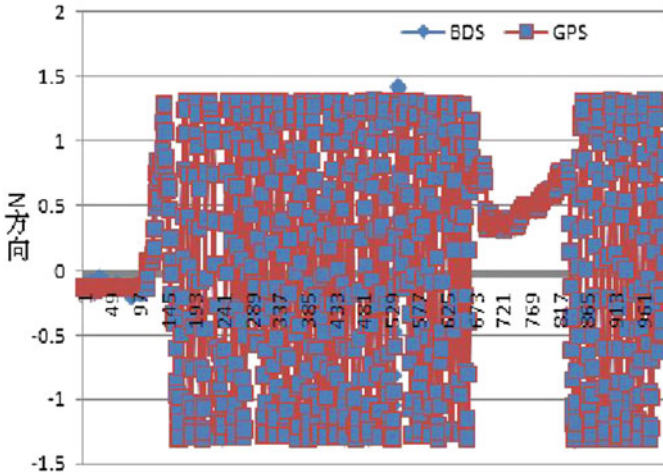


Fig. 17.12 BDS and GPS dynamic baseline solution results with baseline length of about 1.3 m in N direction

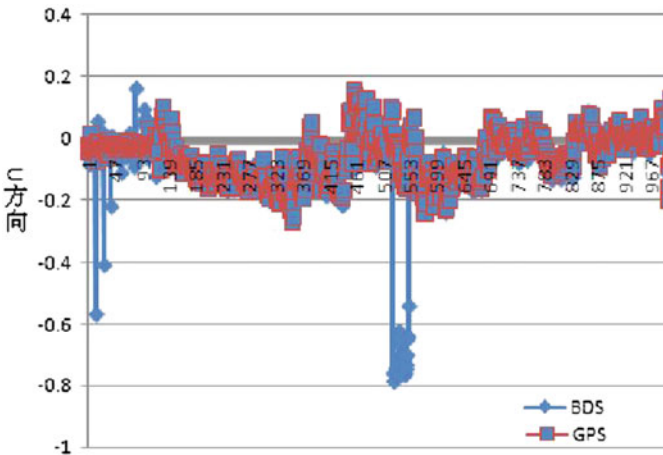


Fig. 17.13 BDS and GPS dynamic baseline solution results with baseline length of about 1.3 m in U direction

Table 17.2 BDS and GPS measurements and standard deviation in dynamic baseline

	Mean (m)			Average (m)		
	E	N	U	E	N	U
BDS	0.960	0.672	-0.087	0.379	0.4175	0.167
GPS	0.964	0.670	0.078	0.380	0.4185	0.063

17.5 Conclusions

In this paper, a detailed comparison of ultra-short baseline measurement performance with the use of BDS and GPS is made. We get the following conclusions:

- (1) As a whole, whether static or dynamic, the ultra-short baseline measurement results of BDS and GPS are quite equal, with no significant difference, the observation noise of BDS is yet a little bigger compared to GPS.
- (2) The experiments show that under the same conditions, the number of visible satellites of BDS tend to be more than GPS, in the integer ambiguity fixing of the baseline solution, the RATIO value of BDS is smaller than GPS, and the PDOP value of BDS do not differ greatly with GPS.
- (3) Under the static condition, the precision of ultra-short baseline solution of BDS is slightly worse than GPS regardless in the directions of E, N or U.
- (4) Under the dynamic condition, the precision of ultra-short baseline solution of BDS in the E and N directions are rather equal to that of GPS, while in the U direction it is not as good as GPS.

References

1. He H (2002) High-precision GPS dynamic measurement and quality control. PLA Information Engineering University
2. Horemuz M, Sjöberg L (2002) Rapid GPS ambiguity resolution for short and long baselines. *J Geodesy* 76:381–395
3. Li K, Yang L, Chai L, Ding A, Guo Z (2014) GNSS positioning theory. Coal Industry Press
4. Sui L, Song L, Chai H (2010) Error theory and measurement adjustment basis. Mapping Press
5. Tang W, Deng C, Gao L (2013) Preliminary results of single Epoch baseline solution based on Beidou navigation satellite system. *Geomat Inf Sci Wuhan Univ* 38(8)
6. Teunissen P (1993) Least-squares estimation of the integer GPS ambiguities. Paper presented at the invited lecture, section IV theory and methodology. IAG general meeting, Beijing, China
7. Teunissen P (1999) An optimality property of the integer least-squares estimator. *J Geodesy* 73 (11):587–593
8. Teunissen P, Giorgi G, Buist P (2011) Testing of a new single-frequency GNSS carrier phase attitude determination method: land, ship and aircraft experiments. *GPS Solut* 15(1):15–28
9. Wang B, Miao L, Wang S, Jun Shen (2009) A constrained LAMBDA method for GPS attitude determination. *GPS Solut* 13(2):97–107

Chapter 18

The iGMAS Combined Products and the Analysis of Their Consistency

Hongliang Cai, Kangkang Chen, Tianhe Xu and Guo Chen

Abstract Several analysis centers (AC) and a product integration and service center have been built by the international GNSS continuous Monitoring and Assessment System (iGMAS). Product integration and service center (ISC) is a reprocessing center of high-precision products, which is also the window of products distribution center. Product integration and service center analyzes the quality of product from each analysis center, and then reprocesses to generate the combined product. The consistency of combined products is very important to the users of navigation and positioning, which directly affects the service performance of combined products. At first, this paper introduces the progress in iGMAS product combination. Then, the consistency of basic combined products (orbits, Earth rotation parameters, station coordinates, and clocks) is analyzed in details. And the precise point positioning tests are implemented to verify the consistency of combined products. Finally, some suggestions are given according to the test results of the present stage. The results of PPP test show that the static PPP coordinate daily repeatability of combined orbit/clock is at millimeter level. The standard deviation of combined orbit/clock single epoch PPP difference is comparable with

H. Cai (✉)

Beijing Institute of Tracking and Telecommunication Technology, Beijing 100094, China
e-mail: caibanyu@126.com

K. Chen

School of Geology Engineering and Surveying, Chang'an University, Xi'an, Shanxi, China
e-mail: chenkkuser@126.com

T. Xu

State Key Laboratory of Geo-information Engineering, Xi'an, Shanxi, China

T. Xu

State Key Laboratory of Astronautic and Dynamics, Xi'an, Shanxi, China

T. Xu

Xian Research Institute of Surveying and Mapping, Xi'an, Shanxi, China

G. Chen

The School of Geomatic and Geodesy, Wuhan University, Wuhan, Hubei, China

© Springer-Verlag Berlin Heidelberg 2015

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015*

Proceedings: Volume III, Lecture Notes in Electrical Engineering 342,

DOI 10.1007/978-3-662-46632-2_18

the best AC orbit/clock solutions. The results of static and dynamic PPP also prove that the consistency of combined orbit/clock considering the consistency correction is improved significantly.

Keywords iGMAS · Product combination · Analysis of consistency · Precise point position

18.1 Introduction

The international GNSS continuous Monitoring and Assessment System (iGMAS) is an open technology platform to monitor and evaluate the health and key performance of global satellite navigation system (GNSS). The purpose is to generate and publish high-precision products of orbit and clock errors, tracking station coordinates, earth rotation parameters, global ionosphere delay and other products. Product integration and service center (ISC) is an important part of the international continuous monitoring and assessment system and is also the center of high-precision products reprocessing. Meanwhile it is the window of products distribution center. Currently, all iGMAS analysis centers have submitted the station coordinate, ERPs, satellite orbit and clock errors of GPS, GLONASS, BDS and Galileo to ISC. ISC analyzes the quality of all analysis centers' products on time, and provides the combination products and their precision assessment report at the same time. Due to different kinds of combined products and different methods that are used, the consistency among iGMAS combined products will not be guaranteed. Terrestrial reference frame is maintained by tracking station coordinates. If station coordinate is not consistent with the combined orbits and ERPs, it would cause some serious problems. The consistency of satellite orbit and clock products also has an important influence on the precision of precise point positioning [1–3].

The research on analysis center products combination had become a rich source since the IGS organization was founded. Many foreign researchers, such as Springer et al. have done a lot of work in this area [4–10]. The product combination models and methods of IGS analysis center (AC) orbit, clock error, station coordinates and other products have been improved continually [8–11]. Currently, the international GNSS continuous monitoring and assessment system is in the experimental stage. It will provide more complete products and service including iGMAS station related products, ionosphere scintillation index and integrity products etc. It will also face many new problems and challenges. Among them, the consistency of the iGMAS combination products is a problem that needed to solve currently. This paper reviewed the present situation and problems of iGMAS product combination. And then, the consistency of combined products was discussed in detail. This paper focused on the strategy of data processing to improve the consistency of combined product, and the tests of precise point positioning were

implemented to verify the consistency of combined orbit and clock products. Finally, some suggestions were given according to the current situation of ACs, which would make iGMAS provide better products and service and meet the needs of different users.

18.2 The Present Situation of iGMAS Product Combination

Usually, the data processing software and strategies of analysis centers are not the same. Meanwhile, the station number and station distribution of each AC in product calculation are also different. There is an obviously difference in the coordinate and time reference frame of AC' products. There are abnormal values in products of a single analysis center, especially the clock errors. Therefore, it is necessary to combine the station coordinates, ERPs, satellite orbit and clock errors submitted by analysis center and provide combined products.

Currently, iGMAS products integrated center can generate and push almost all products on time, such as final, rapid and ultra-rapid orbit, clock errors and ERPs, final station coordinates, final and ultra-rapid troposphere products, final and rapid ionosphere TEC grid, DCB and GNSS integrity etc. The algorithm of products integration has been improved and perfected continuously. Meanwhile, the product integration software has also been modified and upgraded. There are some problems found in the test phase. The precision of iGMAS station products is poor. There are no ERPs in SINEX file submitted by analysis centers, which makes it impossible for combining ERPs and station coordinates at the same time. Different PCO corrections of BDS satellites are adopted by analysis center, which has an influence on the combination and accuracy assessment of BDS satellite clock errors. The coordinates of IGS and iGMAS station in SINEX file submitted by analysis center are calculated separately, they are not self-consistent, which makes that there are many null matrix in the variance matrix when calculating the combined station coordinates. In addition, there are less same stations in different analysis center's SINEX solution, which makes the combined solution unstable.

18.3 The Analysis of iGMAS Combined Products' Consistency

Analysis center submit their all products calculated independently to ISC in single file separately on time. ISC analyzes the quality of submitted products and calculates the combined products. Usually, every combined product is calculated using different algorithm and program separately. If we do not take effective measures, the self-consistency among the combined products will be very poor [2]. It is our

impression that the combined orbits, ERPs and clock errors are of very high quality and are in general more accurate and reliable than the solutions obtained by the individual ACs. Nevertheless, there are possibly a few improvements which can be made. First of all the consistency between the combined products can, and should, be improved, especially with respect to those users who want to perform precise point positioning.

The combination of station coordinates and ERPs uses the SINEX file that submitted by analysis centers as input, through adopting the appropriate pre-processing algorithm to detect the gross errors of input data, determining the optimal weight of ACs, using the method of superposition equation to calculate the combined solution. The combined solutions include the station coordinates, ERPs and geocenter motion [10–12]. Finally, the cumulative solutions including reference epoch's station coordinates and velocity are obtained with time series stack [13–15]. It will be possible to adding ERPs in SINEX file of Analysis Center that ERPs and station coordinates are combined at the same time. It can improve the quality of the combined ERPs and ensure the consistency of station coordinates and ERPs.

The calculation model of Analysis Center final orbits is that satellite orbit, ground tracking station coordinates and ERPs are estimated at the same time using the original observation data. The Final procedures use a no-net rotation (or other removable) constraint, satisfied over the Igb08 core network. It is incumbent upon the ACs to ensure that their procedures are consistent with the expectations of the combinations. There are obvious system difference among analysis center' products, such as satellites orbit. In order to maintain the consistency between comprehensive of final product and ground station coordinates as well as ERP. The AC SINEX and ERP rotations were added to the combination model by Springer in early 2000. The introduction of RX_{snx}^{ac} , RY_{snx}^{ac} and RZ_{snx}^{ac} was proposed by Jan Kouba to maintain consistency between the IGS SINEX and final orbit combinations. The approach is predicated upon the assumption that each set of AC Final orbits, ERPs, SINEX, and clocks are each internally self-consistent. Provided the assumption is valid, then applying RX_{snx}^{ac} , RY_{snx}^{ac} and RZ_{snx}^{ac} ensures long-term consistency between the IGS terrestrial frame and the final orbits. RX_{erp}^{ac} and RY_{erp}^{ac} were introduced by T. Springer to approximate day-to-day variations in the AC orbits not captured by the AC weekly SINEX rotations. At that time, the IGS SINEX combinations were based on weekly integrations, and so the associated RX_{snx}^{ac} , RY_{snx}^{ac} and RZ_{snx}^{ac} were weekly averages. In his report on these matters, T. Springer showed that the ERP rotations significantly reduced the AC rotational biases and scatter. However, the SINEX combinations of iGMAS are based on daily integrations. There is no longer a need to include the ERP rotations in the combination model so that the final orbits are not unnecessarily harmed by the ERP rotations [9]. At the AC level, the ultra-rapid and rapid procedures transfer the terrestrial frame to the orbits by tightly fixing the a priori positions of the Igb08 reference frame stations. At the combination level, the main procedural difference between the rapid/ultra-rapid and final is that AC SINEX and AC ERP X- and Y-rotations are not applied for the iGMAS.

The satellite orbit and clock errors are combined using the weighted average method separately. The self-consistency of orbit and clock errors affects the service performance directly. In order to maintain the consistency of all IGS combined products, the compatibility corrections of satellite orbit, station coordinates and geocenter offset are applied to AC clock solutions prior [8]. Since AC clock solutions, in general, refer to different reference clocks, significant inconsistency errors can be introduced, in particular in the areas with solution gaps, or when satellite and/or station is missing from AC clock solutions. The clock combination accounts for these AC reference clock differences before the combination. It is important that the alignment of the combined clocks to a chosen reference time. Currently, the broadcast clocks are used for this purpose [16].

Due to the difference in data processing, calculating strategy and station selection etc. of analysis center, it is not appropriate to calculate troposphere combinations using the weighted average method. The final troposphere combinations of iGMAS is obtained by using precise point positioning technology to estimate stations' tropospheric delay with combined orbit and clock errors products. It can maintain the consistency between troposphere combinations and other products, such as orbit, clock errors, station coordinates and ERPs. Precise point positioning usually adopts ionosphere-free combination observation equation, the consistency of ionosphere combinations won't be discussed anymore.

Precise Point Positioning (PPP) is an effective method to evaluate the consistency of orbit and clock corrections, which usually adopts the ionosphere free combination observation equation for non-difference solution [2, 8]. The error equation can be expressed as follows:

$$\begin{aligned} v_p^j &= \rho^j + c \cdot \delta t + \delta \rho_{trop}^j - p^j + \varepsilon_p \\ v_\phi^j &= \rho^j + c \cdot \delta t + \delta \rho_{trop}^j + \lambda \cdot N^j - \lambda \cdot \phi^j + \varepsilon_\phi \end{aligned} \quad (18.1)$$

where j is the PRN number of satellite; c is the speed of light in vacuum; δt is the receiver clock error; $\delta \rho_{trop}^j$ is tropospheric delay correction; $\varepsilon_p, \varepsilon_\phi$ is the not modeling error, such as multipath and observation noise etc. p^j, ϕ^j is the ionosphere free combination observations of corresponding satellite; v_p^j, v_ϕ^j is observation error, λ is the wavelength; ρ^j is the geometric distance between the satellite position at signal transmission time and the receiver position at signal receiving time; N^j is the integer ambiguity of ionosphere free combination observations. After linearization of Eq. (18.1) is:

$$\begin{aligned} V &= AX - L \\ X &= [x \quad y \quad z \quad \delta t \quad \delta \rho_{trop} \quad N^j]^T \end{aligned} \quad (18.2)$$

where A is the corresponding design matrix; L is the constant term that corresponding observed values minus the calculated value of the theoretical schematic; X are parameters to be estimated; x, y, z is three-dimensional positional parameters;

δt is receiver clock parameters; $\delta\rho_{trop}$ is tropospheric delay parameters; N^j is integer ambiguity parameters.

When solving the equation, the position parameters of single epoch (dynamic) Precise Point Positioning are calculated separately for each epoch. In the case of static, the position parameters are regarded as constant. In the situation that without cycle slips or cycle slips is repaired correctly, integer ambiguity is regarded as a constant parameter. In the event of cycle slips, the integer ambiguity is often treated as a new constant parameter processing.

In this paper, we select the same precise point positioning software and station observation data. The precise point positioning tests were implemented using the products BDS Week 0458 (285–291 day of year in 2014). In the PPP tests, nine kinds of orbit and clock corrections were used respectively. The test results was analyzed and used to verify the consistency of integrated orbit and clock errors. The combined orbit and clock errors are obtained with independent orbit and clock errors of six iGMAS analysis centers. The nine kinds of orbits and clock corrections are as follows: combined orbit and clock with the consistency correction (ISC1), combined orbit and clock without the consistency correction (ISC2), IGS orbit and clock (IGS), orbit and clock of six iGMAS analysis Centers (AC1, ..., AC6). The static and single epoch PPP tests were respectively implemented with 25 IGS stations, which have good global distribution. The distribution of stations is as follows (Fig. 18.1):

The daily repeatability of static PPP can efficiently verify the consistency of orbit and clock errors. Figure 18.2 and Table 18.1 is the results of daily repeatability of static PPP test. The median of 25 stations coordinates daily repeatability with different orbit and clock corrections.

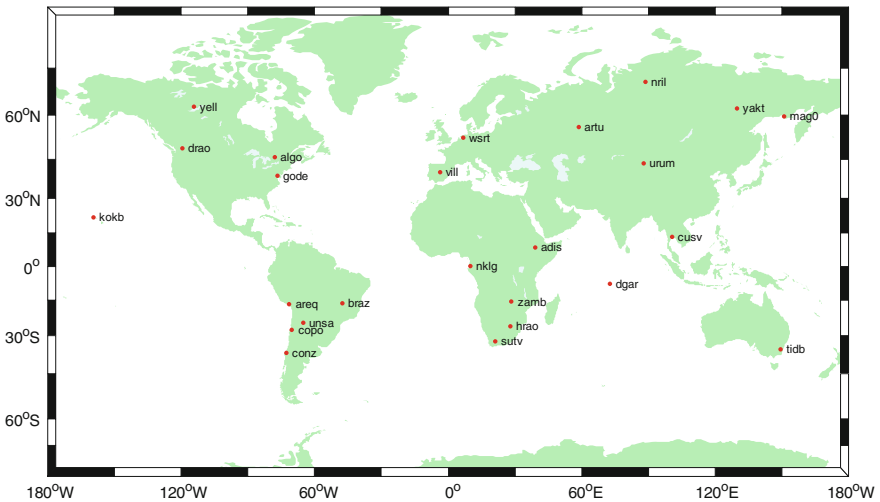


Fig. 18.1 The distribution of sites used

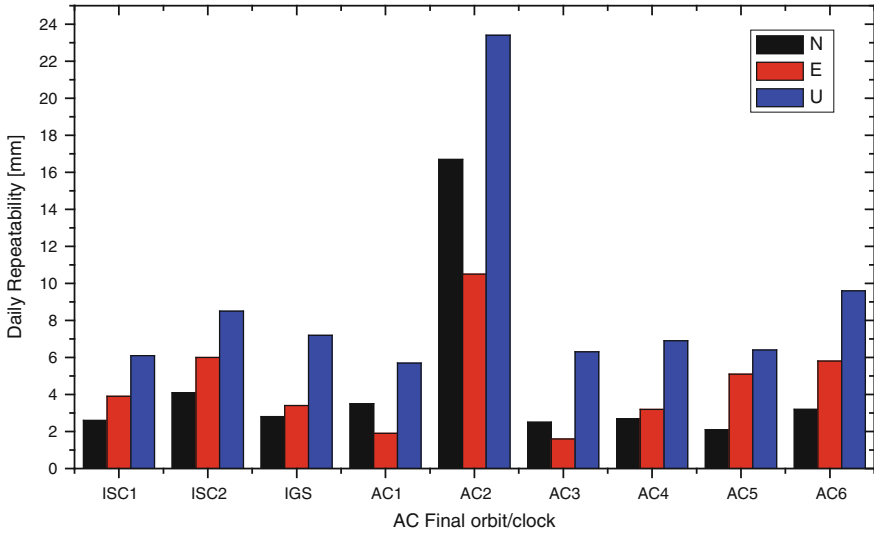


Fig. 18.2 Median daily repeatability of PPP with different orbits/clocks

Table 18.1 Median daily repeatability of PPP with different orbits/clocks (unit: mm)

Product	Period	North	East	Up
ISC1	285–291	2.6	3.9	6.1
ISC2	285–291	4.1	6.0	8.5
IGS	285–291	2.8	3.4	7.2
AC1	285–291	3.5	1.9	5.7
AC2	285–291	16.7	10.5	23.4
AC3	285–291	2.5	1.6	6.3
AC4	285–291	2.7	3.2	6.9
AC5	285–291	2.1	5.1	6.4
AC6	285–291	3.2	5.8	9.6

The results of single epoch PPP can more clearly and directly reflect the consistency and reliability of orbit and clock corrections. We selected four stations from the static PPP station data, which are ARTU, BRAZ, MAG0 and WSRT. Then we implemented single epoch PPP tests with nine kinds of orbit and clock products, which are the combined orbit and clock (ISC1) with consistency correction, combined orbit and clock (ISC2) without consistency correction, IGS orbit and clock, orbit and clock of six iGMAS analysis Centers (AC1, ..., AC6). The single epoch PPP results of four stations at first day of the BDS week were calculated using different orbit and clock products. Figures 18.3, 18.4 and 18.5 is respectively the ARTU station N, E, U component coordinate difference of single epoch PPP with respect to the IGS station coordinates solution. Figures 18.6, 18.7 and 18.8 is respectively the WSRT station N, E, U component coordinate difference of single

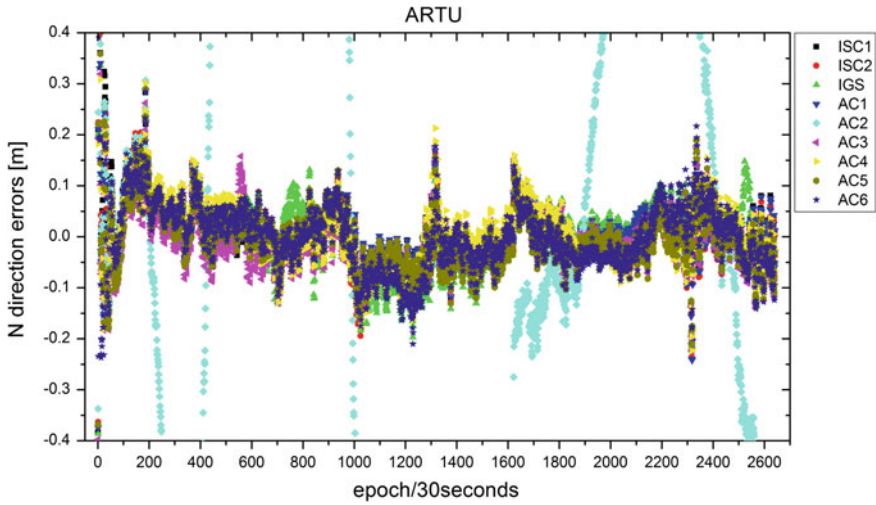


Fig. 18.3 The N component coordinate difference of single epoch PPP with respect to the IGS station coordinates solution

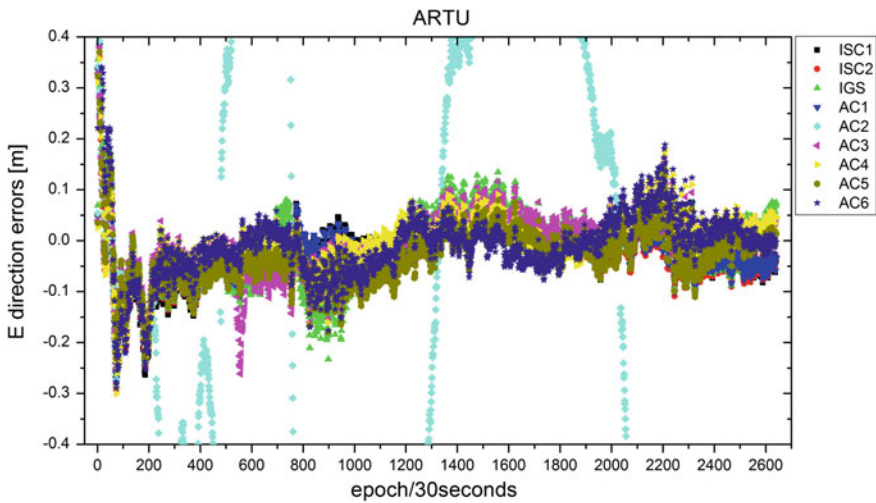


Fig. 18.4 The E component coordinate difference of single epoch PPP with respect to the IGS station coordinates solution

epoch PPP with respect to the IGS station coordinates solution. Tables 18.2, 18.3 and 18.4 is respectively the four stations standard deviation of N, E, U component coordinate difference of single epoch PPP with different orbit/clock products.

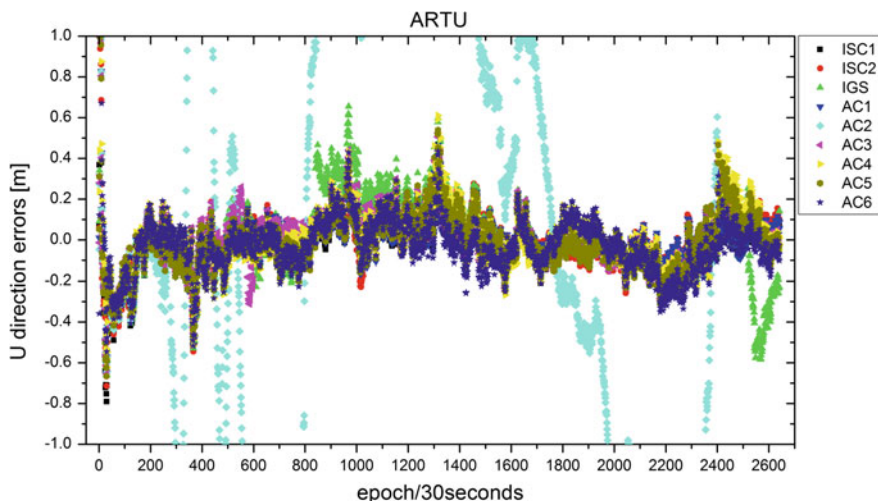


Fig. 18.5 The U component coordinate difference of single epoch PPP with respect to the IGS station coordinates solution

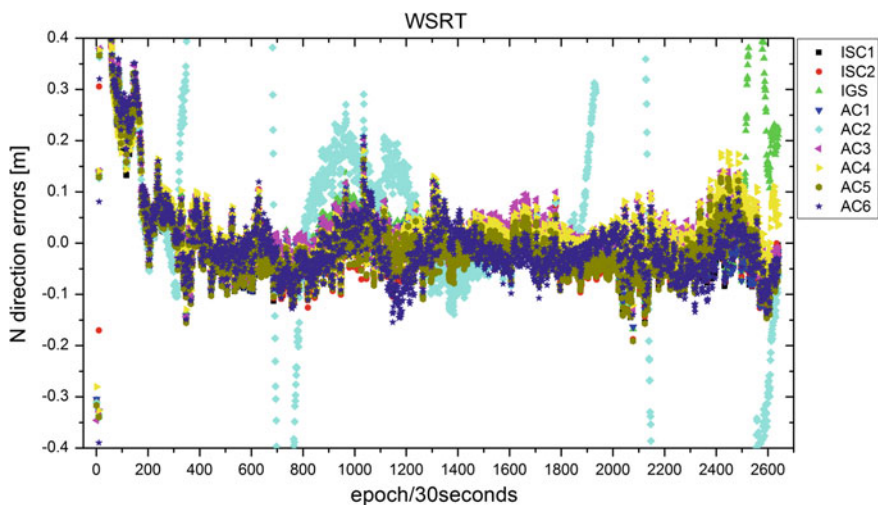


Fig. 18.6 The N component coordinate difference of single epoch PPP with respect to the IGS station coordinates solution

From the results above, we can see that

- (1) In the results of static PPP tests we can see that the median of daily repeatability using different orbit/clock products is at millimeter level except AC2. Among them, the result of AC3 and AC1 is best, then the ISC1 and IGS. The median of daily coordinate components repeatability of ISC1 is better

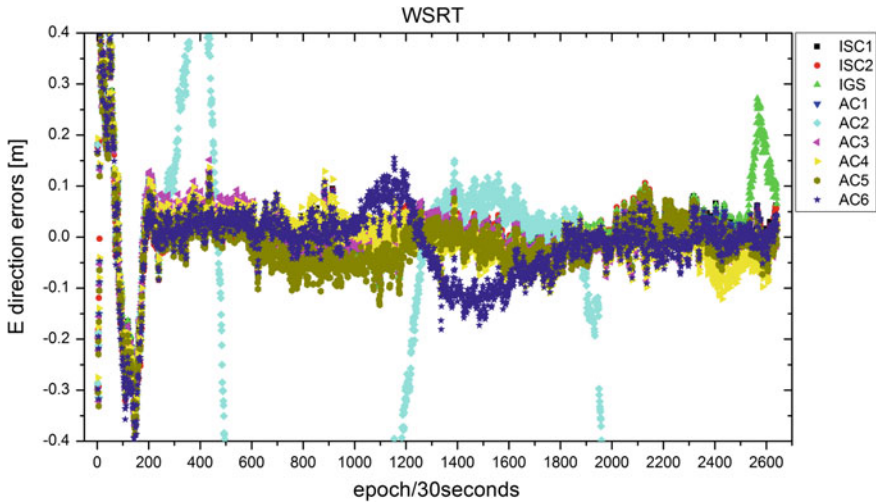


Fig. 18.7 The E component coordinate difference of single epoch PPP with respect to the IGS station coordinates solution

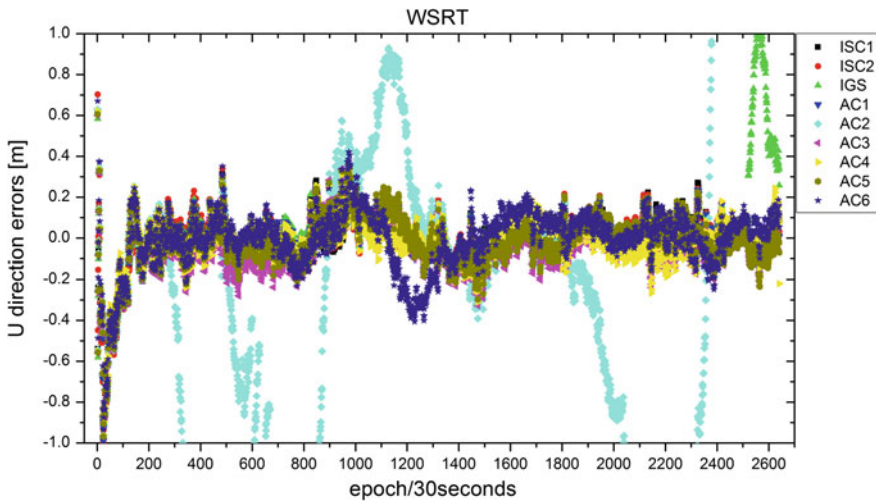


Fig. 18.8 The U component coordinate difference of single epoch PPP with respect to the IGS station coordinates solution

than ISC2 obviously, which proved that the consistency of ISC1 orbit and clock have significantly improved than ISC2.

- (2) We can see that from the N, E, U component coordinate difference of single epoch PPP with different orbit/clock products in Figs. 18.3, 18.4, 18.5, 18.6, 18.7 and 18.8 the difference of AC1 and ISC1 is smaller and more smooth.

The difference of AC2 is larger than others. Meanwhile, the results of AC2 single epoch PPP have no convergence. The divergence appeared in the end of IGS single epoch PPP difference sequence. It is supposed that there are some abnormal values in the satellite clock corrections. In addition to AC2, the results of single epoch PPP get convergence within 200 epochs. After convergence, the absolute value of E and N component coordinate difference is less than 10 cm, and the absolute value of U component coordinate difference is less than 20 cm.

- (3) Tables 18.2, 18.3 and 18.4 is the standard deviation of the N, E, U component coordinate difference of single epoch PPP with different orbit/clock products. We can see that the standard deviation of N component coordinate difference of AC1, ISC1 and AC3 is about 5 cm, which is smaller than others. The standard deviation of E component coordinate difference of AC1, ISC1 and AC4 is within 6 cm, and the U component is within 11 cm. The standard deviation of AC2 N, E, U component coordinate difference is all larger than 1 m, which is consistent with the difference sequence in Figs. 18.3, 18.4, 18.5, 18.6, 18.7 and 18.8. The standard deviation of N, E, U component coordinate difference of ISC1 single epoch PPP is all better than ISC2, which increased 10 % on average.

Table 18.2 The standard deviation of N component coordinate difference of single epoch PPP (unit: m)

Product	ARTU	BRAZ	MAG0	WSRT	Mean
ISC1	0.051	0.047	0.044	0.058	0.050
ISC2	0.057	0.054	0.049	0.060	0.055
IGS	0.065	0.048	0.059	0.085	0.064
AC1	0.049	0.043	0.044	0.057	0.048
AC2	1.381	0.188	1.508	1.474	1.138
AC3	0.050	0.055	0.044	0.058	0.052
AC4	0.059	0.057	0.044	0.057	0.054
AC5	0.052	0.064	0.044	0.062	0.055
AC6	0.062	0.073	0.048	0.066	0.062

Table 18.3 The standard deviation of E component coordinate difference of single epoch PPP (unit: m)

Product	ARTU	BRAZ	MAG0	WSRT	Mean
ISC1	0.049	0.065	0.051	0.050	0.054
ISC2	0.051	0.068	0.059	0.059	0.060
IGS	0.063	0.061	0.076	0.054	0.064
AC1	0.043	0.058	0.051	0.051	0.051
AC2	1.444	0.140	2.463	1.018	1.266
AC3	0.057	0.068	0.057	0.054	0.059
AC4	0.051	0.053	0.054	0.057	0.054
AC5	0.046	0.072	0.053	0.055	0.056
AC6	0.054	0.067	0.057	0.072	0.063

Table 18.4 The standard deviation of U component coordinate difference of single epoch PPP (unit: m)

Product	ARTU	BRAZ	MAG0	WSRT	Mean
ISC1	0.105	0.131	0.104	0.082	0.105
ISC2	0.112	0.136	0.110	0.091	0.112
IGS	0.146	0.130	0.152	0.147	0.144
AC1	0.098	0.129	0.102	0.081	0.103
AC2	3.158	1.137	4.326	1.654	2.569
AC3	0.119	0.170	0.114	0.096	0.125
AC4	0.130	0.116	0.103	0.085	0.108
AC5	0.118	0.174	0.100	0.095	0.122
AC6	0.115	0.206	0.166	0.117	0.151

- (4) The results of static PPP and single epoch PPP tests both proved that the consistency between combined orbit and clock considering consistency correction is better than without consistency correction. The models and methods of iGMAS products integrated and service center used to improve the combination products consistency is correct and effective.

18.4 Conclusions and Recommendations

ISC is a reprocessing center of iGMAS high-precision products. Meanwhile it is also the center of products distribution, which provides services directly to a variety of users. Therefore, it is necessary to ensure the reliability and consistency of the combined products. The data processing strategies and methods for improving the consistency among the combined products were introduced in this paper. The consistency of combined products was analyzed in detail. The results of static PPP and single epoch PPP tests fully proved that the consistency between combined orbit and clock products is comparable with the best products of iGMAS analysis centers.

In order to maintain the consistency of integrated products, the products submitted by analysis centers, such as orbit, clock corrections and station coordinates, should be internal self-consistent, which is a prerequisite for the method of product consistency correction introduced in this paper. It is necessary for each AC to generate station coordinates, ERPs, orbit and clock products etc. using the same set of software. To make the combined iGMAS orbit and coordinates, essentially the iGMAS reference frame, consistent we need to include the Earth Rotation Parameters in the SINEX files. The inclusion of these parameters will allow it to obtain combined ERPs which are consistent with the reference frame. The AC orbits can then easily be made consistent with this reference frame. At the same time the accuracy of the combined ERPs should be improved obviously. One other product can be estimates of the Earth's Center of Mass, which could be included in the SINEX files. These estimates may also be useful for the orbit combination and

clock combination. To apply PPP technique to upcoming missions with Low-Earth-Orbiting (LEO) satellites carrying GNSS receivers, one will need nearly continuous knowledge of the satellite clocks. It will be mandatory to have access to satellite clock estimates with a 30 s sampling rate. If satellite clocks are provided with a sampling of 30 s, people might also become interested in obtaining 30 s station clock estimates. This would especially be interesting in the framework of the time transfer project. However 30 s station clock estimates might be much more difficult to do than 30-second satellite clock estimates. Some ACs, such as IGG and TLC, have already provided 30 s sampling rate final and rapid clock products, the other ACs are needed to join in as soon as possible. Currently, iGMAS analysis centers only submit BDT products to ISC. The observation data of almost all tracking stations is referenced to GPST. The loss of accuracy in clock interpolation from BDT to GPST is great, especially the station clock. In order to avoid the loss of precision caused by the interpolation calculation and improve the usability of the integrated products, the GPST products should also be submitted and combined.

Acknowledgments This work was supported by Natural Science Foundation of China (41174008) and the Open Foundation of State Key Laboratory of Geodesy and Earth's Dynamics (SKL-GED2013-4-2-EZ) and State Key Laboratory of Astronautic and Dynamics (2014ADL-DW0101).

References

1. Jan K, Pierre H (2001) Precise Point positioning using IGS orbit and clock products. *GPS Solut* 5(2):12–28
2. Springer TA, Zumberge JF, Kouba J (1998) The IGS analysis products and consistency of the combined solutions. In: IGS Analysis Center Workshop, Darmstadt, 9–11 Feb 1998
3. Kouba J, Mireault Y (1997) Analysis Coordinator report international GPS service for geodynamics (IGS) 1996 annual report, pp 55–100
4. Springer TA, Beutler G (1993) Towards an official IGS orbit by combining the results of all IGS processing centers. In: Proceedings of the 1993 IGS workshop, Bern, Switzerland, pp 242–250, 24–26 March 1993
5. Beutler G, Kouba J, Springer TA (1995) Combining the orbits of the IGS analysis centers. *Bull Geodesique* 69:200–222
6. Kouba J, Mireault Y, Lahaye F (1995) 1994 IGS Orbit/clock combination and evaluation, Appendix I of the analysis coordinator report. In: International GPS Service for Geodynamics (IGS) 1994 Annual Report, pp 70–94
7. Altamimi Z, Boucher C, Sillard P (2002) New trends for the realization of the international terrestrial reference system. *Adv Space Res* 30(2):175–187
8. Kouba J, Springer Tim (2001) New IGS station and satellite clock combination. *GPS Solut* 4(4):31–36
9. Griffiths J (2012) A review of the IGS final orbit combination procedures (2012), report from the analysis coordinator. <http://acc.igs.org/>
10. Ferland R, Piraszewski M (2009) The IGS-coordinates, Earth rotation parameters and apparent geocenter. *J Geodesy* 83:385–392
11. Gambis D, Biancale R et al (2009) Combination of Earth orientation parameters and terrestrial frame at the observation level. In: Drewes H (ed) *Geodetic reference frames*. Springer, Berlin, pp 3–9

12. Shi C, Zou R, Yao Y, Li M (2008) Systematic error analysis in data combination based on SINEX solution. *J Wuhan Univ* 33(6):P608–P611 (in Chinese)
13. Altamimi Z, Collilieux X, Métivier L (2011) ITRF2008: an improved solution of the international terrestrial reference frame. *J Geodesy* 85(8):457–473
14. Altamimi Z, Métivier L, Collilieux X (2012) ITRF2008 plate motion model. *J Geophys Res* 117
15. Yao Y (2004) Research on the algorithm and realization of post-processing for GPS precise positioning and orbit determination. Wuhan Univ, Wuhan (in Chinese)
16. Chen K, Xu T (2014) A new strategy on precise clock combination of IGS analysis centers. In: China satellite navigation conference (CSNC) 2014 proceedings, vol 2, pp 255–268

Chapter 19

Enhanced RTK Integer Ambiguity Resolution with BeiDou Triple-Frequency Observations

Tao Li, Kongzhe Chen and Jinling Wang

Abstract With the availability of BeiDou Navigation Satellite System (BDS) triple-frequency observations, more strategies can be carried out to improve integer ambiguity resolution (IAR) performance for both short and long baseline RTK. In this paper, we first present the intrinsic natures of IAR for short and long baseline RTK. The double differenced mathematical models for short and long baseline are specified first. From the model perspective, the IAR performances of dual-frequency and triple-frequency are analyzed and compared. To improve the AR performance using triple-frequency observations, the integer least-squares (ILS) success-rate can be increased by first partial fixing the Extra Wide-lane (B2&B3) and then the Wide-lane (B1&B2, B1&B3) integer ambiguities in a geometry-based model because of their relatively long wavelengths. For short baselines, the IAR at each carrier can be resolved conditioned on the resolved WL integer ambiguities. For long baselines, two ionosphere-free combinations (B1&B2, B1&B3) can be formulated, and only the integer ambiguity vector on each carrier needs to be fixed with more redundant observations. It turns out that with triple-frequency observations, the performances of both short and long baseline RTK can be improved.

Keywords Beidou navigation satellite system · Triple-frequency RTK · Integer ambiguity resolution

T. Li (✉) · K. Chen
Unicore Communications, Inc., Beijing, China
e-mail: taoli@unicorecomm.com

K. Chen
e-mail: kongzhechen@unicorecomm.com

J. Wang
The University of New South Wales, Sydney, Australia
e-mail: jinling.wang@unsw.edu.au

19.1 Introduction

GNSS Real-Time Kinematic (RTK) technique has been popular applied in areas such as surveying, precise navigation, etc. As the key to centimeter-level or even millimeter-level accuracy, the double differenced ambiguity vector in the carrier phase observations requires to be fixed to integers. For short baselines, the double differencing strategy greatly mitigates or eliminates the nuisance parameters such as troposphere delay and ionosphere delay. This is, however, extremely difficult for long baseline scenarios, in which the atmosphere delays are quite different due to geographic difference, especially for the ionosphere delay, which is mixed with the integer ambiguities and hinders the integer ambiguity resolution (IAR).

Traditionally, dual-frequency observations are adopted to formulate linear combinations to fully eliminate the first-order ionosphere delay in long baseline RTK, such as the ionosphere-free (IF) linear combination implied in Bernese [2]. With the modernization of GNSS, triple-frequency RTK has been researched in numerous publications, such as the Three-Carrier Ambiguity Resolution (TCAR) model suggested in Forssell et al. [5] and Vollath et al. [23], and the Cascade Integer Resolution (CIR) model developed in Hatch et al. [10] and Jung et al. [12]. These two models have been studied and compared with LAMBDA in Teunissen [20], which concludes that those two models are both examples of integer bootstrapping, rather than the optimal integer least-squares (ILS) [19]. In Feng [4] and Li [13], a number of linear combinations with the aim of longer wavelength, reduced/eliminated ionosphere delay and decreased observation noise are proposed and it is suggested that the ambiguity resolved Extra Wide-lane (EWL) and Wide-lane (WL) models can be applied as precise code observations. These researches analyze the performance of triple-frequency RTK performance from a mathematical model point of view.

With the development of BeiDou Navigation Satellite System (BDS), three carrier signals are currently being broadcasted in the Asia-Pacific region, namely the B1, B2 and B3, with frequencies of 1561.098, 1207.140 and 1268.520 MHz [1, 16, 25]. Providing the actual gathered triple-frequency observations, Tang et al. [17] proposed the modified TCAR model and compared it with the other models and the results seem quite optimistic given a prior information of the ionosphere-delay. However, obtaining such a prior information with certain accuracy is hardly possible in real-time applications. A recent publication by Zhao et al. [26] takes advantage of the ambiguity resolved EWL and WL carrier phase observations and three code observations to pursue an optimal linear combination to estimate the ionosphere-delay so that the integer ambiguities at each carrier can be resolved, yet the accuracy of the estimated ionosphere-delay is extremely essential for IAR at each carrier. It is apparently that triple-frequency IAR is still an open problem for both short and long baseline RTK solutions, and of course more investigations are required.

This contribution is organized as follows. We first present the intrinsic natures of IAR for short and long baseline RTK. The double differenced mathematical models for short baselines (in which ionosphere delay is assumed to be absent) and long baselines (in which ionosphere delay is to be modeled) are specified. From the

model perspective, the IAR performances of dual-frequency and triple-frequency (including both short and long baselines) are analyzed and compared in terms of ILS success-rate, which maximizes the success-rate of IAR. To improve the IAR performance using BDS triple-frequency observations, a Z-transformation, which transforms the original ambiguity vectors of B1, B2 and B3 to EWL (B2&B3) and WL (B1&B2), are applied. The ILS success-rate can be increased by a partial fixing of the EWL integer ambiguities first because of its relatively long wavelength and less impacted by the ionosphere-delay, and followed by a partial fixing of the WL integer ambiguities conditioned on the resolved EWL ambiguities. Then, for short or medium baselines, the integer ambiguity vector at each carrier can be fixed conditioned on the resolved WL integer ambiguities. For long baselines, the two IF combinations (B1&B2 and B1&B3) are formulated and with the resolved WL integer ambiguities, only the integer ambiguity vector of B1 (or B2, B3) needs to be fixed after reparametrizing the ambiguity component so that more redundant observations are obtained. To evaluate the proposed scheme, BDS triple-frequency data from various baseline lengths are surveyed and the IAR performances for the original ambiguity resolution and the partial fixing are all statistically compared. It turns out that with the proposed strategy, RTK performance with triple-frequency observations can be improved for both short and long baselines. The IAR can sometimes be conducted epoch-wisely even for long baselines. In the end, the conclusions are given.

19.2 Mathematical Models

To minimize the impacts of troposphere delay and the ionosphere delay, a double differencing between satellites and receivers is used. Without loss of generality, the double differenced mathematical models can be described as follows [11, 14]:

$$\Delta\nabla P_k = \Delta\nabla\rho + \frac{f_1^2}{f_k^2}\Delta\nabla I + \Delta\nabla v \quad (19.1)$$

$$\Delta\nabla L_k = \Delta\nabla\rho - \frac{f_1^2}{f_k^2}\Delta\nabla I + \lambda_k\Delta\nabla N_k + \Delta\nabla\varepsilon \quad (19.2)$$

where P and L are the code and carrier phase observations in unit of meter; $\Delta\nabla$ is the double differencing operator; ρ represents the geometric distance and the non-dispersive troposphere delay; I is the ionosphere delay; N_k is the integer ambiguity vector on the k th frequency, with the corresponding wavelength of λ_k . v and ε are the observation noise for code and carrier phase, respectively. The remaining multipath is assumed to be assimilated in the observation noise.

The stochastic models of the above equations can be defined as:

$$D = \sigma_0^2 \Sigma^{-1} \quad (19.3)$$

where D is the variance-covariance matrix of the code and carrier phase observations. It is assumed there is no correlation between code and carrier phase observations and the noise ratio between code and carrier phase is 100. Even though research has indicated that B3 has a much smaller noise than that of B1 or B2 [17], we presume that each carrier shares the same noise. σ_0^2 is the a priori variance and Σ is the weight matrix.

After linearization of Eqs. (19.1) and (19.2), the unknown parameters in the geometry based models can be estimated in ways of either least-squares or Kalman filter. Hereafter, we make use of the least-squares estimation, yet the conclusion holds the same for Kalman filter. The estimated integer ambiguity vector is the so-called float ambiguities (or real-valued ambiguities). To resolve the integer ambiguities, the ILS estimation has been proved to be optimal in terms of maximizing the ILS success-rate and it is recommended to be used for IAR.

19.2.1 Short Baseline RTK

In the case of short baselines, the double differenced atmospheric delays can be greatly mitigated or even eliminated. In such a case, the IAR performance is rarely impacted by the ionosphere delay. For one satellite pair, geometry-free models of (19.1) and (19.2) can be expressed in a matrix form as follows:

$$\begin{bmatrix} \Delta \nabla P_k \\ \Delta \nabla L_k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & \lambda_k \end{bmatrix} \begin{bmatrix} \Delta \nabla \rho \\ \Delta \nabla N_k \end{bmatrix} \quad (19.4)$$

It is obvious that in Eq. (19.4), the integer ambiguity vector is only affected by the observation noise. With further extension, the triple-frequency observations of Eq. (19.4) can be explicitly given as:

$$\begin{bmatrix} \Delta \nabla P_1 \\ \Delta \nabla P_2 \\ \Delta \nabla P_3 \\ \Delta \nabla L_1 \\ \Delta \nabla L_2 \\ \Delta \nabla L_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & \lambda_1 & 0 & 0 \\ 1 & 0 & \lambda_2 & 0 \\ 1 & 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \Delta \nabla \rho \\ \Delta \nabla N_1 \\ \Delta \nabla N_2 \\ \Delta \nabla N_3 \end{bmatrix} \quad (19.5)$$

Applying the models as shown by Eqs. (19.3, 19.4 and 19.5), the IAR performance of dual-frequency and triple-frequency of a short baseline can be evaluated according to the ILS success-rate.

19.2.2 Long Baseline RTK

When the baseline length is relatively long (for instance, longer than 20 km), the atmospheric delays sometimes cannot be eliminated by double differencing due to the geographic differences, especially for the ionosphere delay, which differs for each satellite and is almost inseparable with the integer ambiguities in a single epoch. In a similar manner as the short baseline RTK, the mathematical models for one satellite pair can be explicitly given as:

$$\begin{bmatrix} \Delta\nabla P_1 \\ \Delta\nabla P_2 \\ \Delta\nabla P_3 \\ \Delta\nabla L_1 \\ \Delta\nabla L_2 \\ \Delta_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & \frac{f_1^2}{f_2^2} & 0 & 0 & 0 \\ 1 & \frac{f_1^2}{f_3^2} & 0 & 0 & 0 \\ 1 & -1 & \lambda_1 & 0 & 0 \\ 1 & -\frac{f_1^2}{f_2^2} & 0 & \lambda_2 & 0 \\ 1 & -\frac{f_1^2}{f_2^2} & 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \Delta\nabla\rho \\ \Delta\nabla I \\ \Delta\nabla N_1 \\ \Delta\nabla N_2 \\ \Delta\nabla N_3 \end{bmatrix} \tag{19.6}$$

Equation (19.6) implies that in the long baseline scenarios, the ionosphere delay needs to be modeled and estimated. However, it is shown that the ionosphere delay and the integer ambiguity vector are highly correlated for a satellite pair. Even though the ionosphere delay can be estimated with the assistance from the code observations, the accuracy is not good enough and the residual of the ionosphere delays still deteriorate the ambiguity vector from integer. As a consequence, the IAR performance suffers severely from the magnitude of the ionosphere delay.

In Fig. 19.1, the ILS success-rates for short baseline (dual and triple-frequency) and long baseline (dual and triple-frequency) RTK are plotted as functions of the code noise. As can be seen from the above figure, the IAR performance for short baseline and long baseline are quite different. The short baseline assumes that there is no ionosphere delay and the corresponding ILS success-rate is extremely close to

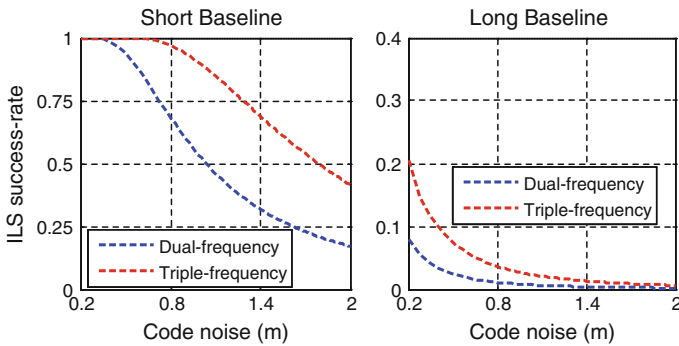


Fig. 19.1 Single epoch IAR performance with geometry-free models for short baseline and long baseline, code noise ranges from 0.2 to 2.0 m

1.0 for both dual-frequency and triple-frequency when the code observation noise is small. With the increasing of the code observation noise, the ILS success-rate gradually decreases and overall, the success-rate ranges from nearly 0.2–1.0. However, in case of long baseline, due to the existence of the ionosphere delay, the ILS success-rate drops drastically. As is shown in the right figure of Fig. 19.1, the maximum success-rate is no more than 0.25 while the minimum success-rate approaches to 0. This clearly explains why IAR becomes troublesome when conducting long baseline RTK or under strong ionosphere disturbance.

Meanwhile, Fig. 19.1 also demonstrates the ILS success-rate differences between dual-frequency and triple-frequency. Apparently, the ILS success-rate of triple-frequency outperforms dual-frequency no matter for short baselines or long baselines, which implies that, from the model's perspective, when applying triple-frequency observations, the IAR performance is supposed to be better and more reliable.

19.3 Enhanced Triple-Frequency Models

The previous section has concluded that the ionosphere delay has strong correlation with the integer ambiguities for each satellite pair and triple-frequency RTK prevails dual-frequency RTK in terms of a higher ILS success-rate. Recall the relation among the integer ambiguities at each carrier, it is possible to formulate new integer ambiguities which can be resolved with higher success-rate and then apply these resolved integer ambiguities as constraints to fix the original integer ambiguities at each carrier. Therefore, we should achieve the goal of reducing the impacts of ionosphere delay on IAR first and then construct integer ambiguities with a longer wavelength.

In [20], the TCAR and CIR models have been compared with the LAMBDA method and also a partial fixing strategy is recommended in case the ILS success-rate of fixing the whole triple-frequency ambiguity vector is not high enough. In fact, the partial fixing has been further studied in Verhagen [22] and applying a partial ambiguity fixing achieves sufficiently higher success-rate than the original ambiguity vector. Therefore, we resort to partial fixing to fix the EWL and WL integer ambiguities first and then resolve the integer ambiguity vector at each carrier.

19.3.1 Partial EWL and WL Ambiguity Fixing

Following the Z-transformation discussion as specified in [20], the Z-transformation matrix for BDS triple-frequency can be constructed:

$$\begin{bmatrix} \Delta\nabla N_1 \\ \Delta\nabla N_{w12} \\ \Delta\nabla N_{ew} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \Delta\nabla N_1 \\ \Delta\nabla N_2 \\ \Delta\nabla N_3 \end{bmatrix} \quad (19.7)$$

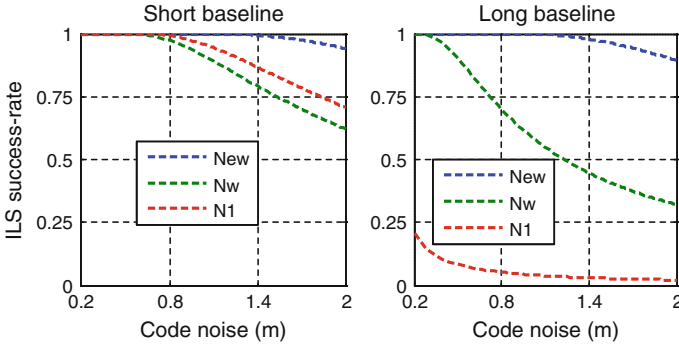


Fig. 19.2 ILS success-rates for N_{ew} (dash blue), N_w conditioned on N_{ew} (dash green), N_1 conditioned on N_w (dash red)

where $\Delta\nabla N_{w12} = \Delta\nabla N_1 - \Delta\nabla N_2$ and $\Delta\nabla N_{ew} = \Delta\nabla N_2 - \Delta\nabla N_3$. Correspondingly, the B1 and B3 ambiguity vector can be derived as $\Delta\nabla N_{w13} = \Delta\nabla N_{w12} + \Delta\nabla N_{ew}$. Since among these three ambiguity vector on the left side of (19.7), $\Delta\nabla N_{ew}$ suffers from the least ionosphere delay, the partial fixing strategy can be carried out from here and then the ambiguity vector of $[\Delta\nabla N_1 \quad \Delta\nabla N_w]^T$ and its associated variance-covariance matrix can be updated conditioned on the resolved $\Delta\nabla N_{ew}$ vector using the conditional least-squares. Similarly, the $\Delta\nabla N_1$ can be attempted to resolve when the $\Delta\nabla N_w$ is resolved.

In Fig. 19.2, the ILS success-rates of partial ambiguity fixing using triple-frequency observation are depicted. The left figure plots the ILS success-rates of N_{ew}, N_w conditioned on N_{ew} and the N_1 AR conditioned on the resolved N_w integer ambiguity vector. It is observed that the N_{ew} integer ambiguity is rather easy to resolve with sufficient high success-rate and the N_w IAR conditioned on the resolved N_{ew} can attain a higher success-rate than the success-rate of the whole original AR specified in the left figure of Fig. 19.1. Therefore, despite that N1 is very difficult to resolve when the measurement noise increases, the resolved N_w integer ambiguity can ensure a decimetre to even centimetre level positioning accuracy. Providing the N_w integer ambiguities, the N1 success-rate can be determined sequentially. An interesting phenomenon is that the N1 success-rate is even higher than that of N_w IAR in the left figure of Fig. 19.2, which could be understood by the absence of the ionosphere delay and the constraint from the resolved N_w integer ambiguity vector and its variance-covariance matrix.

The right figure presents the performance of partial ambiguity fixing for long baseline. Obviously, the ionosphere-delay affects ILS success-rate of N_1 more than that of N_w and N_{ew} . The success-rate of N_{ew} AR is still quite high even in long baseline case, which is useful for constraining N_w IAR. A decrease of the success-rate for N_w can be seen by comparing with the left figure. However, if the observation noise is small, the success-rate of N_w IAR is still extremely high, which means that decimetre or even centimetre level accuracy is pretty achievable even for long baseline RTK. In contrast with the left figure, the success-rate of N_1 drops

dramatically with the increased code noise, which implies that the consideration of the ionosphere delay has severe impact on the N_1 IAR. The success-rate of N_1 using the constraint from N_w does not show too much improvement comparing with the original whole IAR in Fig. 19.1. An explanation is that the resolved N_w integer ambiguities cannot separate the ionosphere delay from the N_1 ambiguity with enough accuracy.

It is worth mentioning that in practical applications, the user can estimate the float ambiguities from the geometry-based model because of the geometry constraint advantage [20]. Therefore, with the original float ambiguity vector and its variance, LAMBDA [18] or MLAMBDA [3] can be used to conduct the IAR. If the resolved integer ambiguities pass the validation threshold, no partial fixing is required. For more details on integer ambiguity validation, the user could refer to [7, 8, 15, 21, 22, 24]. If the ambiguity validation fails, the above-mentioned Z-transformation and partial fixing can be used. In our analysis, the R-ratio test [6] and the ILS success-rate are utilized as measures to study the AR performance. Note that the user can also compare the resolved N_{ew} and N_w ambiguity vector with the Hatch-Melbourne-Wenbunna (HMW) model [9] with satellites having high elevation to ensure the reliability of partial fixing.

19.3.2 The Double IF Linear Combinations

To fully eliminate the first-order ionosphere delay in long baseline RTK, the dual-frequency IF linear combination has been applied in GNSS positioning. When triple-frequency observation available, the IF linear combinations for the phase can be defined as:

$$\Delta\nabla L_{1k} = \Delta\nabla\rho + \Delta\nabla T + \lambda_{w1k}\Delta\nabla N_{w1k} + \lambda_{n1k}\Delta\nabla N_1 + \Delta\nabla\varepsilon_{1k} \quad (19.8)$$

with $\lambda_{w1k} = \frac{\lambda_1^2\lambda_k}{\lambda_k^2 - \lambda_1^2}$, $\Delta\nabla N_{w1k} = \Delta\nabla N_1 - \Delta\nabla N_k$ and $\lambda_{n1k} = \frac{\lambda_1\lambda_k}{\lambda_1 + \lambda_k}$. In Eq. (19.8), $\Delta\nabla N_{w1k}$ are obtained from the previous partial fixing and then only the $\Delta\nabla N_1$ is left to be resolved. If the observation quality is good, this integer ambiguity vector can be hopefully resolved epoch-wisely.

19.4 Numerical Analysis

In order to investigate the actual AR performance of the proposed processing scheme, two static data sets with BDS triple-frequency observations were collected using UB370 products (which are capable of tracking GPS L1 and L2, GLONASS L1, L2 and BDS B1, B2 and B3 signals) from Unicore Communications, Inc, Beijing. Data set 1 was 2 h long collected in Beijing, with a baseline length of 10.5 km. Data set 2 was nearly 1.5 h collected in Guangdong, with the baseline

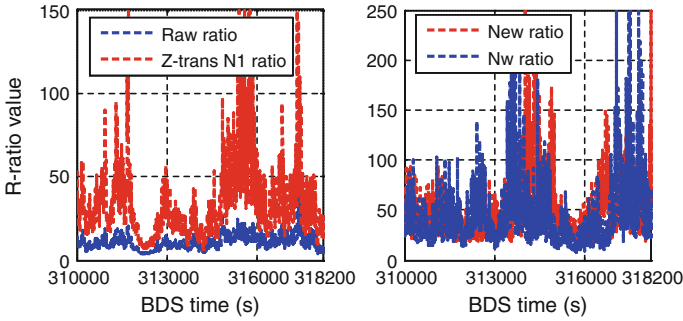


Fig. 19.3 Epoch by epoch ratio values for data set 1. Raw ratio indicates using the original ambiguity vector and Z-trans N_1 ratio means the ratio was obtained conditioned on the partial fixing of N_{ew} and N_w , the ratios of which are plotted in the *right* figure

length of 50 km. The data were processed with the original ambiguity vector first, and then processed with the proposed scheme for short and long baselines, respectively. The ratio values were analyzed on a epoch by epoch basis.

The left figure of Fig. 19.3 shows the differences of the Z-transformed N_1 AR and the original AR using triple-frequency observations. By partial fixing of N_{ew} and N_w integer ambiguity, the N_1 integer ambiguity was resolved with an apparent higher ratio than the original ratio. The right figure demonstrates that the N_{ew} and N_w ratios are high enough so that when an empirical ratio of 3 is applied, all the integer ambiguity vector of N_{ew} and N_w can be successfully validated. As a matter of fact, the minimum ratio for the N_1 , N_w and N_{ew} are 5.4, 7.5 and 14.6, while for the original ambiguity vector, the minimum ratio is 2.5. Therefore, the partial ambiguity fixing certainly ensures the reliability of IAR in a short baseline case, and when using 3.0 as the ambiguity validation criteria, all epochs can be successfully validated for partial fixing. However, when using the original ambiguity vector, 96.2 % can be validated.

Figure 19.4 plots the IAR performances for original IAR and partial IAR for a long baseline with strong ionosphere disturbance. Even though there are improvements for the N_1 ratio using the partial fixing strategy, unfortunately, the IAR performance is not good at all, which is as expected due to the inseparable ionosphere-delay. The right figure implies that because of the ionosphere-delay, even success-rates of resolving the N_{ew} and N_w fail occasionally, which implies that in case of strong ionosphere disturbance, even for short baseline, IAR becomes extremely troublesome with all current methods. However, the ratio value of N_w is clearly higher than that of both the raw ratio and the partial N_1 ratio. As a consequence, before successfully resolving the integer ambiguity vector at each carrier, the resolved N_w can be used to achieve decimeter to even centimeter-level accuracy for long baseline RTK.

To eliminate the first-order ionosphere-delay, the double IF linear combinations are adopted. The true AR performance (compared with the correct integer ambiguity vector) is summarized in Table 19.1.

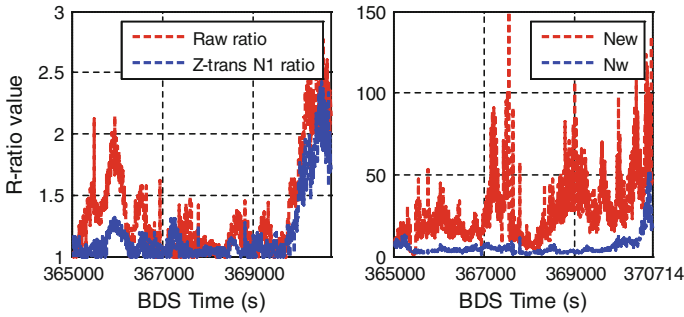


Fig. 19.4 Epoch by epoch ratio values for data set 2. Raw ratio indicates using the original ambiguity vector and Z-trans N_1 ratio means the ratio was obtained conditioned on the partial fixing of N_{ew} and N_w , the ratios of which are plotted in the *right* figure

Table 19.1 True IAR success-rates of using original AR, partial AR and Partial AR with IF

	1 epoch	20 epochs	50 epochs
Original AR (%)	5.9	10.3	14.2
Partial AR (%)	7.4	15.0	20.4
Partial AR with IF (%)	7.8	19.6	25.9

As shown in Table 19.1, the original AR has true success-rates of 5.9, 10.3 and 14.2 % for single epoch solution, 20 epochs batch solution and 50 epochs batch solution. The success-rates increase to 7.4, 15.0 and 20.4 % when the partial fixing is applied. By canceling the first-order ionosphere-delay, the partial AR with IF can yield success-rate as high as 7.8, 19.6 and 25.9 %. Comparing with the performance of partial AR only, the single epoch success-rate doesn't improvement significantly, which could be understood from the fact that the partial AR with IF has a shorter wavelength than the original wavelength on the B1 carrier, and the model strength is not strong enough. However, with the increasing of processed epochs, the partial AR with IF passes the partial AR only.

19.5 Conclusions

In this contribution, we have investigated the benefits of using BDS triple-frequency IAR in the cases of short baselines and long baselines. It turns out triple-frequency observations can lead to a higher ILS success-rate than the dual-frequency observations in case of short and long baselines. To further improve the ILS success-rate, we have proposed an enhanced triple-frequency AR scheme by a partial fixing of the EWL and WL integer ambiguities first, and then the integer ambiguity vector is resolved at each carrier with constraint of the resolved WL integer ambiguity in the case of short baselines or using the IF combinations to resolve the integer ambiguity

at each carrier for long baselines. Numerical analysis have indicated that such scheme results in a higher ratio and success-rate than the original triple-frequency IAR. Even though the proposed scheme enhances the reliability of triple-frequency IAR, there are still scenarios that IAR fails, which might be caused by the remaining unmodeled IF delay. Therefore, more research is required to compensate such bias with triple-frequency observations in the future.

References

1. BeiDou navigation satellite system signal in space interface control document, open service signal Version 2.0, 2013
2. Bernese GPS Software Version 5.0 (2007)
3. Chang XW, Yang X, Zhou T et al (2005) MLAMBDA: A modified LAMBDA method for integer least-squares estimation. *J Geodesy* 79:552–565
4. Feng Y (2008) GNSS three carrier ambiguity resolution using ionosphere-reduced virtual signals. *J Geodesy* 82:847–862
5. Forsell B, Martin-Neira M, Harris R (1997) Carrier phase ambiguity resolution in GNSS-2. In: *Proceedings ION GPS-97*:1727–1736
6. Frei E, Beutler G (1990) Rapid static positioning based on the fast ambiguity resolution approach FARA: theory and first results. *Manuscripta Geodaetica* 15:325–356
7. Han S (1997) Quality-control issues relating to instantaneous ambiguity resolution for real-time GPS kinematic positioning. *J Geodesy* 71:351–361
8. Hassibi A, Boyd S (1998) Integer parameter estimation in linear models with applications to GPS. *IEEE Trans Signal Process* 46(11):2938–2952
9. Hatch R (1983) The synergism of GPS code and carrier measurements. In: *International geodetic symposium on satellite Doppler positioning*, vol 2, pp 1213–1231
10. Hatch R, Jung J, Enge P, Pervan B (2000) Civilian GPS: the benefit of three frequencies. *GPS Solut* 3(4):1–9
11. Hofmann-Wellenhof B, Lichtenegger H, Collins J (2001) *Global positioning system: theory and practice*, 5th revised edn
12. Jung J, Enge P, Pervan B (2000) Optimization of cascade integer resolution with three civil GPS frequencies. *Proceedings of ION-GPS 2000*. Salt Lake City, UT, USA, pp 2191–2200
13. Li B, Feng Y, Shen Y (2010) Three carrier ambiguity resolution: distance-independent performance demonstrated using semi-generated triple frequency GPS signals. *GPS Solutions* 14:177–184
14. Li T, Wang J (2012) Some remarks on GNSS integer ambiguity validation methods. *Surv Rev* 44:230–238
15. Li T (2013) *Integer ambiguity estimation and validation for precise GNSS positioning*. Ph.D. dissertation
16. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Teunissen P, Nakamura S (2013) Initial assessment of the COMPASS/BeiDou-2 regional navigation satellite system. *GPS Solut* 17:211–222
17. Tang W, Deng C, Shi C, Liu J (2013) Triple-frequency carrier ambiguity resolution for BeiDou navigation satellite system. *GPS Solut* 18(3):335–344
18. Teunissen P (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *J Geodesy* 70:65–82
19. Teunissen P (1999) An optimality property of the integer least-squares estimator. *J Geodesy* 76:587–593

20. Teunissen P, Joosten P, Tiberius C (2002) A comparison of TCAR, CIR and LAMBDA GNSS ambiguity resolution. In: Proceedings ION GPS 2002, Portland, OR, 24–27 Sept 2002
21. Teunissen P, Verhagen S (2009) The GNSS ambiguity ratio-test revisited: a better way of using it. *Surv Rev* 41:138–151
22. Verhagen S (2005) The GNSS integer ambiguities: estimation and validation. Ph.D. dissertation
23. Vollath U, Birnbach S, Landau H (1998) Analysis of three-carrier ambiguity resolution (TCAR) technique for precise relative positioning in GNSS-2. In: Proceedings ION GPS-98, pp 417–426
24. Wang J, Stewart MP, Tsakiri M (1998) A discrimination test procedure for ambiguity resolution on-the-fly. *J Geodesy* 72:644–653
25. Yang YX, Li JL, Xu JY, Tang J, Guo HR, He HB (2011) Contribution of the compass satellite navigation system to global PNT users. *Chin Sci Bull* 56:2813–2819
26. Zhao QL, Dai ZQ, Hu ZG, Sun BZ, Shi C, Liu JN (2014) Three-carrier ambiguity resolution using the modified TCAR method. *GPS Solut*. doi:[10.1007/s10291-014-0421-5](https://doi.org/10.1007/s10291-014-0421-5)

Chapter 20

GPS/BDS One-Step Combined Precise Orbit Determination Based on Double-Differenced Mode

Yao Kong, Baoqi Sun, Xuhai Yang and Xiaozhen Zhang

Abstract In order to improve the precision of BDS precise ephemeris, GPS/BDS one-step combined precise orbit determination based on double-differenced mode was testified in this contribution. And double-differenced data was not formed between GPS and BDS, which eliminates the effect of inter system bias (ISB) on the precision of BeiDou precise orbit. Not only GPS and BDS orbit parameters are estimated, earth rotation parameter (ERP) and station coordinates are also estimated. Data of 114 tracking stations from MGEX are processed. The analysis shows that: Compared with IGS final orbit, the precision of GPS orbit is 5 cm, the precision of X_p , Y_p , length of day (LOD) is 0.06, 0.10 mas, 21.6 μ s. Validating the BDS orbit using SLR observation shows that the radial precision of BDS IGSO and MEO satellites is better than 6 cm, while that of GEO satellite is about 0.37 m.

Keywords BDS · Combined · Orbit determination · Double-differenced

20.1 Introduction

Since December 2012, BeiDou Satellite Navigation System (BDS), which is being developed by China, began to provide official navigation, positioning and timing service for Asia-Pacific area [1]. In order to monitor the signal of BDS and evaluate the performance of BDS, many BeiDou tracking networks have been set up, including BeiDou Experiment Tracking Stations (BETS), international GNSS Monitoring and Assessment System (iGMAS) and Multi-GNSS Experiment (MGEX). MGEX was established by International GNSS Service (IGS) in 2012 to enable an early familiarization with the new signals and systems, such as BeiDou,

Y. Kong (✉) · B. Sun · X. Yang · X. Zhang
National Time Service Center, Chinese Academy of Sciences, Xi'an 710600, China
e-mail: michale08@163.com

Y. Kong · X. Zhang
University of Chinese Academy of Sciences, Beijing 100039, China

Galileo, and QZSS. At present, more than 100 tracking stations have been set up around the world.

With the improvement of space segment of BDS and the increasing numbers of tracking stations, widely research on Precise Orbit Determination (POD) of BeiDou satellites have been carried out. In 2012, based on data from BETS, Shi, Zhao, Ge and He carried out POD of BeiDou satellites and the radial precision of 10 cm was achieved [2–5]. Steigenberger analysed the strategy of POD of BeiDou satellites with bernese software, which focused on arc lengths and solar radiation pressure (SRP) parameters [6]. In 2013, Based on data from MGEX and iGMAS, Zeng did research on GPS/BDS indirect fusion POD, in which double-differenced data was utilized [7]. Besides, Lou carried out POD of BeiDou with data of MGEX and BETS and the results show that the ECOM 5 parameter model was best suitable for it [8]. Due to the limited knowledge on measurement model and force model of BeiDou satellites, the above all research, whether un-difference mode or double-difference mode, all use the two-step method. The two-step method can be described as follows: firstly public parameters, such as receiver error, coordinates, troposphere delay, are estimated with GPS data using precise product from IGS; secondly, BeiDou orbit determination was conducted with the above parameter being fixed. However, little research on GPS/BDS one-step fusion POD has been done. Based on double-differenced mode, GPS/BDS one-step combined POD was proposed in this paper. Compare with POD of BDS satellite with Beidou data only, GPS/BDS combined POD can further improve the precision of Beidou orbit, by precisely estimating the public system bias. Not only are orbit parameters of BDS and GPS estimated, but also station coordinates and ERP are estimated. Data from MGEX was utilized to testify the POD strategy.

The paper is organized as follows. Section 20.2 describes the dataset being used and the processing strategy of GPS/BDS combined POD in details; Sect. 20.3 analyses the precision of GPS orbit and BDS orbit, including internal consistency and external validation. Besides, the ERP and coordinate repeatability are also assessed. Section 20.4 summarizes the results and derives the necessary conclusions.

20.2 Data and Processing

20.2.1 Data

All stations of MGEX are equipped with the Multi-GNSS receivers, which supports tracking of GPS as well as one of the new BeiDou, Galileo, or QZSS constellations. Observation files of MGEX are stored in RINEX 3.02 and its sample interval is 30 s. In this paper, observation of 114 stations from MGEX are utilized for data processing, of which 75 stations can track GPS and BDS constellations simultaneously and the other 39 stations can only track GPS constellation. Figure 20.1 illustrated the distribution of 114 MGEX stations being used. From Fig. 20.1, it is

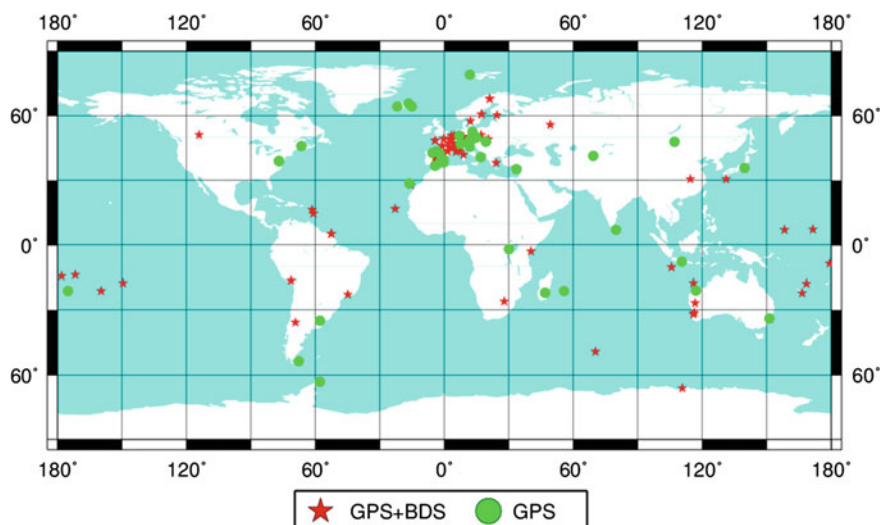


Fig. 20.1 Tracking station network of MGEX

clearly that there is only one station in china area. The data processed in this paper covers the period from 1 October 2014 until 31 October 2014.

20.2.2 Processing Strategy

The Bernese GNSS Software is a scientific, high-precision, multi-GNSS data processing software developed at the Astronomical Institute of the University of Bern (AIUB). However, the ability of data processing of BeiDou system is lacked in this software. Thus, a modified version of Bernese GNSS software was developed, which allows the process of BeiDou data.

The data process comprises of two parts: 1 day solution and 3 day solution. The dual-frequency data of GPS L1 L2 and BDS B1 B2 are employed and the relative weight between GPS and BDS data is 1:1. The process flow of 1 day solution is as follows. Firstly, with zero difference data, receiver clock jumps are detected and corrected and the code observation are smoothed with phase observation. After data importing, the receiver clock are synchronizing using zero difference code data and broadcast ephemeris of GPS and BDS. Then, Single-difference data is formed with a strategy selecting the baselines with the highest number of common-observations from all possible baselines. The pre-processing of phase observation, which runs in a baseline-wise mode, identifies and correct cycle slips using the triple-difference observations. In the next step, a first ambiguity-float network is computed based on the ionosphere-free LC of the double-difference observation. If not handled properly, ISB will bias the final orbit, which is formed when forming double-difference

Table 20.1 Summary of POD strategy

Measurement model	Description
Basic observables	DDs of ionosphere-free LC of Code and phase of GPS and BDS
Data sampling rate	300 s
Elevation angle cut-off	10°
Ionosphere	Ionosphere-free LC
Troposphere a prior model	GMF model
Satellite phase centre	Only nominal for BeiDou, igs08.atx for GPS
Station phase centre	PCO and PCV for GPS and BDS assumed the same
<i>Force model</i>	<i>Description</i>
Gravity field model	EGM2006_SMALL (12 × 12)
Third-body	Sun, Moon, Jupiter JPL DE405
SRP model	ECOM 5 model for GPS and BDS; No a prior model for BDS satellites
Tides and Relativistic effects	Solid earth tides, ocean tides, pole tides, IRES convention 2010
<i>Estimated parameters</i>	<i>Description</i>
Station coordinate	One set of coordinate per session
Phase ambiguities	Float solution
Troposphere	Zenith delay correction in interval of 2 h, mapped with GMF; One set of horizontal gradients per session
Orbit parameters	6 osculating orbital elements plus 5 solar radiation pressure coefficients(constant in D,Y,X periodic terms in X)
Earth rotation parameters	X and Y coordinates of Pole, UT1-UTC as piece-wise-linear function with 2 h spacing

data between GPS and BDS, So, double-difference data was not formed between systems. Based on this solution, the double-difference data was screened for outliers. Besides, bad stations and bad satellites are detected and automatically excluded from the solution. At last, the 1 day solution was finished and 1 day normal equation was generated, which including GPS/BDS six orbit element, ECOM 5 SRP parameters, coordinates, troposphere parameters and ERP. Important options of 1-day solution about measurement model, force model and parameterizations are given in Table 20.1.

The 3 day solution is generated by combining a set of three normal equations from consecutive 1-day solutions. The 3 day orbits are represented by one set of osculating elements, SRP parameters and pseudo-stochastic pulses at 12 h intervals. If satellite modelling problems or maneuver was detected, a 3-day arc is split up at the boundaries into two or three arcs. What is more, in the 3 day solution the ERP, which were set up at 2 h intervals in the 1 day solution, were transformed to daily parameters. When generating the 3 day solution, the ERP and troposphere

parameters are stacked at the boundary to form a continuous piecewise linear function over the whole 3-day interval.

20.3 Results and Discussion

In order to test the processing strategy above, not only GPS orbit and BDS orbit are evaluated, but also ERP and station coordinates are assessed. As far as BDS orbit, Satellite Laser Ranging (SLR) data was also used to validate the precision of BDS orbit.

20.3.1 Precision of GPS Orbit

In order to evaluate the precision of GPS orbit, IGS final orbit products were used as reference, whose precision is 2.5 cm. The monthly average of difference between IGS orbit and estimated orbit are given in Fig. 20.3. GPS PRN3, PRN13, PRN27 satellites are excluded from the data process because of satellite maneuver or modelling problems. From Fig. 20.2, it is clearly that the GPS orbit error (3DRMS) is better than 5 cm (Fig. 20.2).

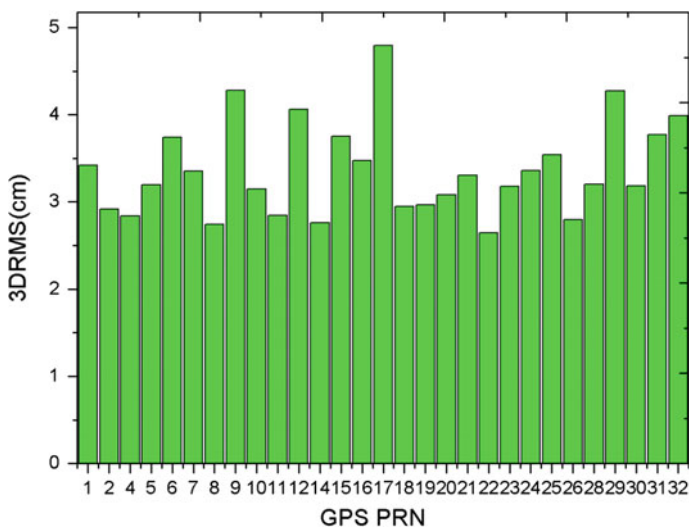


Fig. 20.2 GPS orbit error (compared with IGS final product)

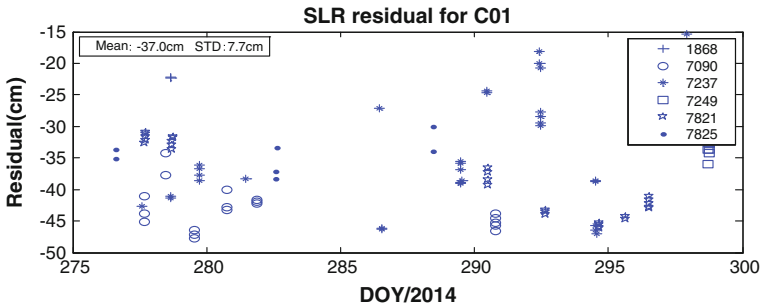


Fig. 20.3 SLR range residual derived from BDS precise ephemeris for C01

20.3.2 Precision of BDS Orbit

20.3.2.1 Internal Consistency

As far as internal consistency, three methods including overlap test, Day Boundary Discontinuities (DBD) and 2 day orbit fit RMS were used. The middle day of the 3 day arc is used for overlap test and computation of the day boundary discontinuities and 2 day fit RMS. The results are given in Table 20.2. The overlap test results and 2 day orbit fit RMS are monthly averages.

The overlap test results show that: for BDS GEO satellites, the radial error is better than 5 cm, the along error is on the several decimetres level and the normal error is better than 10 cm; for IGSO satellites and MEO satellites, the radial error is superior to 5 cm, the along error is better than 15 cm and that of normal is better

Table 20.2 Internal consistency of BDS precise orbit

Items		Overlap difference (cm)			DBD (cm)	2 day fit RMS (cm)
		R	T	N		
GEO	C01	5.8	30.0	8.8	22.7	6.1
	C02	3.4	60.1	14.0	47.2	7.2
	C03	4.9	36.7	8.3	43.8	5.0
	C04	11.1	126.5	9.0	89.6	13.9
	C05	4.5	68.5	13.0	76.5	7.9
IGSO	C06	5.7	13.6	12.6	13.1	3.9
	C07	3.7	13.7	9.6	16.2	2.8
	C08	3.9	11.5	8.2	12.8	2.6
	C09	4.9	12.5	11.5	14.4	4.4
	C10	3.1	12.1	8.3	14.4	2.4
MEO	C11	2.7	8.5	6.3	6.9	3.6
	C12	3.0	10.5	5.9	8.1	4.8
	C14	3.8	10.4	8.2	10.0	8.8

than 10 cm. It is clearly that the precision of C04 is significantly worse than other satellite, which may be caused by the poor geometry condition of it.

By computation the DBD values, the DBD of MEO is better than 10 cm, the DBD of IGSO is about 10 cm and that of GEO is from 20 cm to 90 cm. By analysis the 2 day fit RMS, it is shown that the fit RMS for IGSO and MEO is better than 5 cm, and that of GEO is also better than 15 cm, which is very surprising. It can also be found out that: the 2 day fit RMS values are smaller by a factor of 2–10 factors than the DBD values. The possible reasons are as follows: the DBD values are 3D position difference of two consecutive orbits at midnights, whereas the 2 day fit RMS values are some kind of average over 2 days.

Although the MGEX network is getting larger and larger, it is clearly that the precision of GEO satellites are significant worse than that of IGSO and MEO, especially in along direction.

20.3.2.2 External Validation

SLR, which is completely independent of microwave measurement, is well suited to validate GNSS (Global Navigation Satellite System) orbit derived from microwave measurement. Satellites of BDS are all equipped with laser retro-reflectors arrays. Since 2012 April, C01, C08, C10, C11 satellites of BDS are routinely being observed by the stations coordinated by the International Laser Range Service (ILRS) [9]. Based on SLR data of October 2014, precise orbit of BeiDou are validated. The 1 day orbit is the middle day of the 3 day orbit. The Marini-Murray model is employed for correcting the range measurement for troposphere refraction and the coordinate of SLR station are fixed to SLRF2008. Outliers in the SLR range residual larger than 100 cm were excluded. Moreover, outliers are also excluded with the rule of 3-sigma.

From Fig. 20.3, it can be seen that C01 satellite was observed by 6 SLR stations. Four stations, including 1868, 7821, 7237, 7249, are in the northern part of the globe, while the other two are in the southern part. It is clearly in Fig. 20.3 that the residual of C01 is between -50 cm and 15 cm and the average of it is -37 cm and its standard variation is 7.7 cm. The minus of the average value indicates that the SLR measurement is less than the distance computed with the coordinates and the orbit. Compared with C01, more stations are involved with the tracking of C08, C10 and C11 satellites. 9 stations observed the C08 and C10 satellites and 12 SLR stations observed the C11 satellites. From Figs. 20.4, 20.5 and 20.6, it is shown that the residual of these three satellites are between -10 and 10 cm. And the mean values are -3.5, 1.1 and -1.5 cm and the standard variation is 5.8, 4.8 and 4.8 cm. Because the SLR residual mainly reflects the radial precision of BeiDou orbit, the result of SLR validation is in agreement with the RMS value of overlap test. Being different from BDS IGSO and MEO, the average of residuals of C01 is much larger, which needs further investigation.

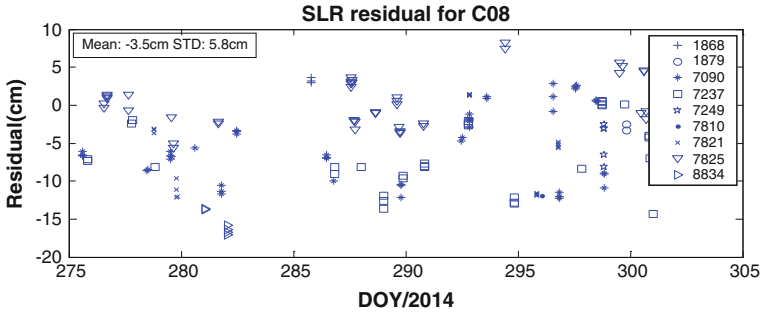


Fig. 20.4 SLR range residual derived from precise ephemeris for C08

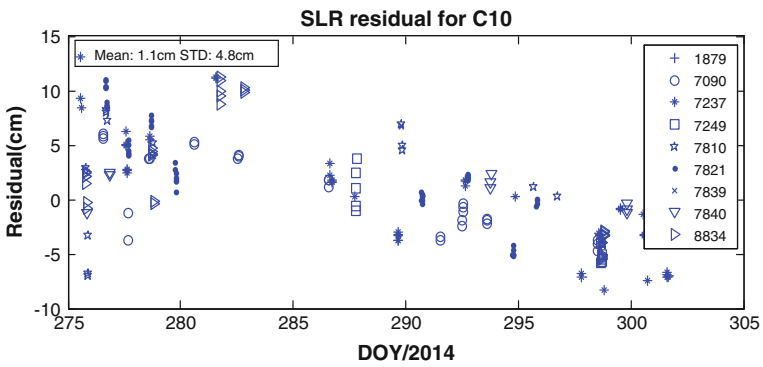


Fig. 20.5 SLR range residual derived from BDS precise ephemeris for C10

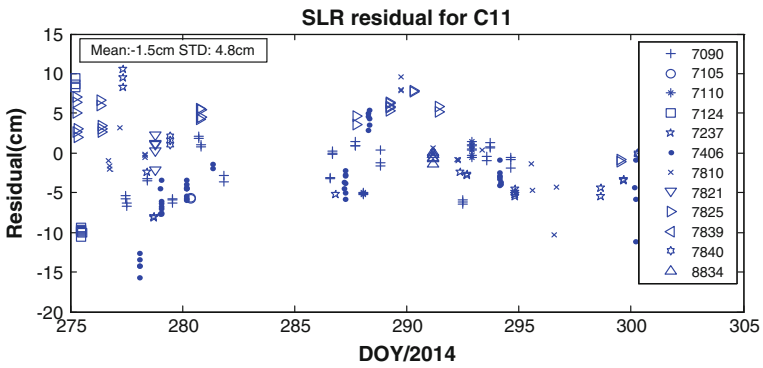


Fig. 20.6 SLR range residual derived from BDS precise ephemeris for C11

Table 20.3 Statistical result of estimated ERP

	Xp (μ as)		Yp (μ as)		LOD (μ s)	
	Mean	RMS	Mean	RMS	Mean	RMS
GPS + BDS	1.8	58.2	-70.2	100.7	10.7	21.6
GPS	-1.4	63.3	-80.8	104.1	6.7	9.6

Table 20.4 Statistical result of coordinate repeatability

Items	E (mm)	N (mm)	U (mm)
RMS	5.2	6.3	11.2

20.3.3 ERP and Coordinate Repeatability

To assess the quality of ERP parameters, including Xp and Yp and LOD, CODE final ERP products are used as reference. It contains ERP parameters of 3 days and time resolution is 12 h. The mean and RMS of three ERP parameters are given in Table 20.3. The precision of Xp, Yp, and LOD is 0.58, 0.63 mas, 21.6 μ s. Besides, the result of ERP, which is estimated only with GPS data, is also given in Table 20.3. Compared with GPS single system, the precision of LOD degrades from 9.6 to 21.6 μ s. The reason behind it should be further investigated. What is more, the precision of Yp is worse by 2 factors of that of Xp. This may be caused by the asymmetry of the tracking station of MGEX.

The coordinate repeatability of solution is an indicator of the coordinate and can also serve as an indicator for the quality of the orbit parameterization. If the processing strategy is not appropriate, the coordinate estimated will be biased, which will worsen the coordinate repeatabilities. The RMS of the residual of daily coordinate with respect to the mean solution for the east, north and height components are given in Table 20.4. With the processing strategy proposed in this paper, the repeatabilities in east, north and height component are 5.2, 6.3, 11.2 mm.

20.4 Summary and Conclusion

Based on data of MGEX, combined GPS and BDS for POD of Beidou satellites were carried out in the contribution, which used the double difference data. And double difference data was not formed between GPS and BDS data, which eliminate the effect of ISB. The analysis shows that the precision of GPS orbit is about 5 cm (3DRMS), which is better than that of BDS satellite. The precision of BDS MEO is better than that of IGSO and GEO satellites and that of GEO is the worst. SLR check shows that the radial error of BDS IGSO and MEO is about 5 cm and system bias of 0.37 m exists in the SLR residual of C01 satellite. Moreover, the precision of Xp, Yp and LOD is respectively 0.058, 0.10 mas, 21.6 μ s. The coordinate repeatabilities in E, N and U directions are 5.2, 6.3 and 11.2 mm.

With the increasing numbers of MGEX tracking stations, especially in china, the orbit precision of BDS satellites can be furtherly improved. Because the solution obtained in the contribution is just ambiguity-float solution, further research should be focused on ambiguity resolution.

Acknowledgments This work is co-supported by international GNSS Monitoring and Assessment Service (iGMAS), National Natural Science Foundation of China (41104021).

References

1. China Satellite Navigation Office (2012) BeiDou Navigation satellite system signal. In: Space interface control document 2.0. Technical report
2. Shi C, Zhao QL, Min Li et al (2012) Precise orbit determination of Beidou Satellites with precise positioning. *Sci China Earth Sci* 55:1079–1086
3. Zhao Q et al (2013) Initial results of precise orbit and clock determination for COMPASS navigation satellite system. *J Geodesy* 87:475–486
4. Ge M, Zhang HP, Jia XL, Song SL, Wickert J (2012) What is achievable with the current compass constellation. *GPS World* 1:29–34
5. He L, Ge M, Wang J, Wickert J, Schuh H (2013) Experimental study on the precise orbit determination of the BeiDou navigation satellite system. *Sensors* 13:2911–2928
6. Steigenberger P, Hugentobler U, Hauschild A, Montenbruck O (2013) Orbit and clock analysis of compass IGSO and GEO satellites. *J Geodesy* 87:515–525
7. Zeng G, Gong B, Wang JS, Li J, Zhu J (2014) BeiDou/GPS indirect fusion precision orbit determination. *Proc Chin Navig Conf* 3:55–63
8. Lou Y, Liu Y, Shi C, Yao X, Zheng F (2014) Precise orbit determination of BeiDou constellation based on BETS and MGEX network. *Sci Rep* 1–10
9. Pearlman MR, Degnan JJ, Bosworth JM (2002) The International Laser Ranging Service. *Adv Space Res* 30:135–143

Chapter 21

Research on Ranging Algorithm Based on Combined-Process Method Using Frequency-Difference, Time-Difference and Relative-Velocity

Rui-Qiang Yang, Peng Zhang, Xing-Wang Zhong, Deng-Feng Wang and Nian-Ke Zong

Abstract The Dual One Way Ranging (DOWR) technique with PRN-code, which can remove the effect of time-asynchronous and improve ranging precision, is widely used in Inter-Satellite-Ranging. The ranging value based on DOWR is delayed, thus cannot match the actual distance at the time the value got within a dynamic environment. In this paper, a synchronous ranging algorithm based on combined-process method using Frequency-difference, Time-difference and relative velocity is proposed, which can realize real-time measurement in high precision by the calibration of ranging value. A dynamic model based on DOWR is introduced first. then, a general analytic expression of distance measurement is constructed, which involves the frequency-difference, the time-difference and the velocity between two aircraft in different receiving patterns. Thus, the distance measurement can be deduced to a certain time, and then, a current-time ranging value would be extrapolated with local pseudo-speed information. The experiments results of certain product demonstrate the effectiveness of the proposed algorithm, the precision of synchronous distance measurement is better than 0.1 m in a dynamic environment.

Keywords DOWR · PRN-code ranging · Time-difference · Frequency-difference · Relative-velocity · Synchronous ranging

R.-Q. Yang (✉) · P. Zhang · X.-W. Zhong · D.-F. Wang · N.-K. Zong
Academy of Space Electronic Information Technology, Xi'an 710100, China
e-mail: yattas@126.com

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_21

249

21.1 Introduction

The inter-satellites-link is widely used in aircraft system such as navigation and positioning satellites, formation flying satellites. Satellites constellation is constructed by the establishment of inter-satellites measurement link and inter-satellites communication link in an inter-satellites-link. One of the important function of inter-satellites-link is inter-satellites ranging. The Dual One Way Ranging (DOWR) technique with PRN-code, which can remove the effect of time-asynchronous and improve ranging precision, is widely used in inter-satellite-ranging.

The DOWR technique needs the pseudo-distance of each satellites at one time. The transmission-delay and clock-diff in local pseudo-distance measurement may not be matched under dynamic environment, while the local time of each satellite will drift because of the frequency of crystal oscillator. The traditional DOWR technique do not pay attention to this situation, some methods using time-difference and frequency-difference have been done to modify the ranging value in dynamic environment. But they do not figure out the method to deal with the mis-match between transmission-delay and clock-diff. On the other hand, the ranging value of DOWR is time-laged. The processing in satellite A needs the pseudo-distance of satellite B, thus the ranging value of satellite A is the distance before current time.

The BD2 navigation satellites constellation will consist of many satellites, with MEO orbit, IGSO orbit and GEO orbit. The velocity between satellites in different orbits is comparatively large, for example, the relative velocity of two satellites in MEO orbit can reach 7.6 km/s. A combined-processing method using time-difference, frequency difference and relative speed is proposed in this paper in order to get high precision ranging value in dynamic environment. The algorithm model and processing procedure are listed. The hardware-in-loop simulation result shows that the precision of real-time ranging using this algorithm is better than 0.1 m in dynamic environment.

21.2 Principle of DOWR Technique with PRN-Code

The DOWR technique with PRN-code uses PRN-code' phase to express transmission delay. The two satellites extract local pseudo-distance in local transmit-frame-time separately, by the measurement of NCO accumulation value in local code tracking loop. The period of local-pseudo depends on the length of local data-frame. And the carrier-doppler-integral is calculated to get pseudo-speed, the pseudo-speed can be derived at any time. These local information are transmitted to the other satellites by inter-satellite-frame. The two satellites receive the signals, and demodulate the measurement information, then get inter-satellite time-difference, frequency-difference, distance and distance change rate based on the combined-processing method.

Fig. 21.1 Dual one way ranging measurement principle diagram

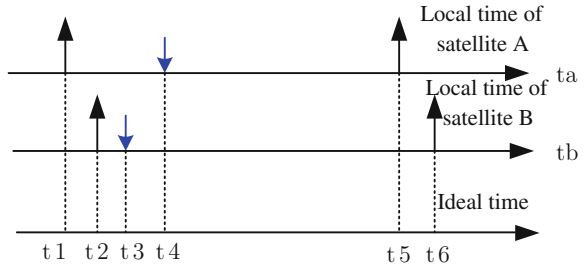


Figure 21.1 shows the principle of DOWR technique. Satellite A and satellite B send timing-signals using their own equipment, and receive timing-signal from peer satellite. The difference of timing-signal in satellite A and timing-signal from satellite B is defined as pseudo-distance of satellite A, which is mentioned as “pd_A (ta(t1), ta(t4))”, while “ta(t1)” means the local begin time of pd_A, “ta(t4)” means the local end time of pd_A. The symbol “t1” means the ideal begin time, the symbol “t4” means the ideal end time. The pseudo-distance of satellite B, which is mentioned as “pd_B(tb(t2), tb(t3))”, is defined the same way as pd_A. The pseudo-distance can be mentioned as “pd_A” or “pd_B” for short.

The value that local time of satellite A exceed local time of satellite B in ideal time “t”, is define as the clock_diff, mentioned as “clk_dt(t)”. The transmission-time in the ideal time “t” is define as “τ(t)”, which presents the distance between two satellite in ideal time “t”. Ignore the system zero-value of distance, the pseudo-distance can be expressed as follows:

$$\begin{cases} pd_A = clk_dt(t1) + \tau(t1) \\ pd_B = -clk_dt(t2) + \tau(t2) \end{cases} \quad (21.1)$$

Apparently, ignore the difference of clock-diff in time t1 and time t2, ignore the difference of transmission-delay in time t1 and time t2, we can evaluate the distance and clock-diff:

$$\begin{cases} \tau(t1) = \frac{pd_A + pd_B}{2} \\ clk_dt(t1) = \frac{pd_A - pd_B}{2} \end{cases} \quad (21.2)$$

21.3 Combined-Process Ranging Method

Several symbols are defined in order to describe the situation more conveniently:

- γ_a the frequency accuracy of satellite A, equals $(f_a - f_0)/f_0$, where f_a means the oscillating frequency of crystal oscillator in satellite A, f_0 means the ideal oscillating frequency of the crystal oscillator.

- γ_b the frequency accuracy of satellite A, equals $(f_b - f_0)/f_0$, where f_b means the oscillating frequency of crystal oscillator in satellite B, f_0 means the ideal oscillating frequency of the crystal oscillator.
- $t_a(t)$ the local time of satellite A, equals $(1 + \gamma_a) * t$, where t means the ideal time.
- $t_b(t)$ the local time of satellite B, equals $(1 + \gamma_b) * t$, where t means the ideal time.
- $clk_dt(t)$ clock-diff, stand for the value that local time of satellite A exceed local time of satellite B in ideal time “ t ”, equals $(\gamma_a - \gamma_b) * t$.
- $v(t)$ the relative velocity between satellite A and satellite B, equals $dR(t)/dt$. The value of $v(t)$ is negative when the two satellites flying towards each other.
- $f_{dopl_A}(t)$ the Doppler frequency of received signal in satellite A, equals $-\frac{v(t)}{c} * f_{RF_B}$, where f_{RF_B} means the carrier frequency of received signal.
- $f_{dopl_B}(t)$ the Doppler frequency of received signal in satellite B, equals $-\frac{v(t)}{c} * f_{RF_A}$, where f_{RF_A} means the carrier frequency of received signal.

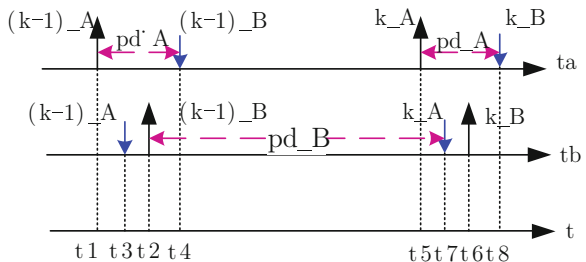
21.3.1 Dynamic Receiving Pattern

According to the definition of clock_diff, the clock_diff will get larger along with the working time of satellite. Thus a receiving pattern may take on with the change of inter-satellite-distance, as is shown in Fig. 21.2: the clock_diff get larger and is larger than the transmission-delay.

The local pseudo-distance are:

$$\begin{cases} pd_A(t1, t4) = clk_dt(t1) + \tau(t2) \\ pd_A(t5, t8) = clk_dt(t5) + \tau(t6) \\ pd_B(t2, t7) = -clk_dt(t2) + \tau(t5) + (t5 - t1) \end{cases} \quad (21.3)$$

Fig. 21.2 Dynamic receiving pattern of dual one way ranging



The time of satellite A and satellite B is separated into several scatter time by the local timing signal. The local timing signal is identified by sequence number. the period of the timing signal in satellite A is T_a , thus the local time of satellite A can be expressed as a time serial $\{ \dots, (k-1)_A, k_A, \dots \} \times T_a$. The local timing of satellite B can be expressed as a time serial $\{ \dots, (k-1)_B, k_B, \dots \} \times T_b$ too. So, the local pseudo distance can be wrote as follows:

$$\begin{cases} pd_A((k-1)_A, (k-1)_B) = clk_dt((k-1)_A) + \tau((k-1)_B) \\ pd_A(k_A, k_B) = clk_dt(k_A) + \tau(k_B) \\ pd_B((k-1)_B, k_A) = -clk_dt((k-1)_B) + \tau(k_A) + T_a \end{cases} \quad (21.4)$$

We can see from formula (21.4), in the expression of pd_B , the sequence number of clock-diff is “ $k-1$ ”, which is not the same as the sequence number “ k ” of transmission delay. This is the mis-match problem in dynamic environment.

The receiving pattern in not only one kind that is described in Fig. 21.2. Figure 21.3 shows another receiving pattern, where the local timing signal of satellite B exceed the local timing signal of satellite A, and the clock-diff is larger than the transmission-delay.

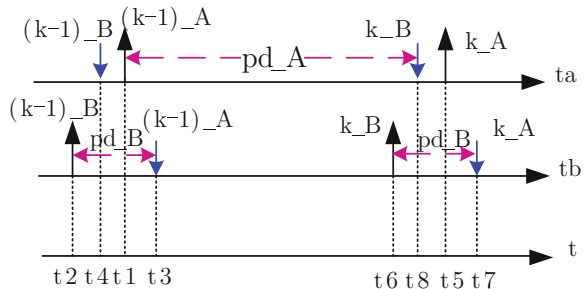
The local pseudo-distance is shown in formula (21.5):

$$\begin{cases} pd_B((k-1)_B, (k-1)_A) = -clk_dt((k-1)_B) + \tau((k-1)_A) \\ pd_B(k_B, k_A) = -clk_dt(k_B) + \tau(k_A) \\ pd_A((k-1)_A, k_B) = clk_dt((k-1)_A) + \tau(k_B) + T_b \end{cases} \quad (21.5)$$

21.3.2 Combined-Process Using Time-Diff, Clock-Diff, Relative Velocity

The transmission-delay and the clock-diff are mis-matched in dynamic environment, which result in the error in distance measurement. Here we re-get the DOWR calculation of Fig. 21.2 based on the definitions in Sect. 1.3, shown as follows:

Fig. 21.3 Dynamic receiving pattern-II of dual one way ranging



$$\begin{aligned}
 \widehat{\tau} &= \frac{pd_A(t1, t4) + pd_B(t2, t7)}{2} \\
 &= \frac{clk_dt(t1) - clk_dt(t2)}{2} + \frac{\tau(t2) + \tau(t5)}{2} \\
 &= \tau(t2) + \frac{(\gamma_a - \gamma_b) * (t2 - t1)}{2} + \frac{1}{2c} \int_{t2}^{t5} v(t) dt \\
 &= \tau(t2) + \frac{1}{2c} \int_{t1}^{t5} v(t) dt + \left(\frac{(\gamma_a - \gamma_b) * (t2 - t1) - \int_{t1}^{t2} \frac{v(t)}{c} dt}{2} \right)
 \end{aligned} \tag{21.6*}$$

$$\begin{aligned}
 \widehat{clk_dt} &= \frac{pd_A(t1, t4) - pd_B(t2, t7)}{2} \\
 &= \frac{clk_dt(t1) + clk_dt(t2)}{2} + \frac{\tau(t2) - \tau(t5)}{2} \\
 &= clk_dt(t1) + \frac{(\gamma_a - \gamma_b)(t2 - t1)}{2} - \frac{1}{2c} \int_{t2}^{t5} v(t) dt \\
 &= clk_dt(t1) - \frac{1}{2c} \int_{t1}^{t5} v(t) dt + \left(\frac{(\gamma_a - \gamma_b)(t2 - t1) + \int_{t1}^{t2} \frac{v(t)}{c} dt}{2} \right)
 \end{aligned} \tag{21.7*}$$

We can see clearly in formula (21.6*) that: ① the measuring-time of DOWR is the same as the begin time of peer-pseudo-distance; ② the DOWR distance measurements include errors caused by time-difference, frequency-difference and relative velocity; ③ the DOWR clock-diff measurements include errors caused by time-difference, frequency-difference and relative velocity too.

The carrier frequency of satellite A and satellite B are f_{RF_A} and f_{RF_B} separately. The carrier-doppler-integral, named ps_A in satellite A and ps_B in satellite B separately, is also extracted:

$$\begin{cases} ps_A(\Delta t) = (\gamma_a - \gamma_b) + (1 + \gamma_b) \frac{\bar{V}}{c} \\ ps_B(\Delta t) = (\gamma_a - \gamma_b) - (1 + \gamma_a) \frac{\bar{V}}{c} \end{cases} \tag{21.8}$$

The \bar{V} in formula (21.8) means the average velocity during Δt . The relative velocity and relative frequency-difference can be evaluated from formula (21.8):

$$\begin{cases} \bar{V} = \frac{c * (ps_A(\Delta t) - ps_B(\Delta t))}{2 + \gamma_a + \gamma_b} \\ (\gamma_a - \gamma_b) = \frac{ps_A(\Delta t) + ps_B(\Delta t)}{2 - (\bar{V}/c)} \end{cases} \tag{21.9}$$

The BD navigation satellites use atomic clock. Take Rb atomic clock for example, the frequency accuracy is better than $1 \times 10e^{-9}$, thus the relative velocity can be expressed approximately as follows:

$$v(t) = \frac{c * (ps_A(t) - ps_B(t))}{2} \quad (21.10)$$

The error of approximate expression in formula (21.10) is $d\bar{V} = \frac{-(\gamma_a + \gamma_b)}{2} * \bar{V}$. Take a MEO satellite for example, which get a relative velocity as 7.6 km/s, the measure error of the relative velocity is less than $7.6 \times 10e^{-6}$ m/s. So, a relative velocity of high precision are given by formula (21.10) under dynamic environment.

Using formula (21.9), the frequency-difference can be evaluated by the relative velocity from formula (21.10). Then the clock-diff can be evaluated using formula (21.7*):

$$clk_dt(t1) = \frac{\widehat{clk_dt} + (\bar{V}_2 * T_a)/(2c)}{1 + ((\gamma_a - \gamma_b) + (\bar{V}_1/c))/2} \quad (21.11)$$

The symbol \bar{V}_1 present the average velocity between ideal time t1 and t2, \bar{V}_2 present the average velocity between ideal time t1 and t5.

Now we got relative velocity, time-difference and frequency difference, then the distance in the begin time of local pseudo-distance in satellite A can be evaluated:

$$\tau(t1) = \hat{\tau} - (\bar{V}_2 * T_a)/(2c) - \left(\frac{(\gamma_a - \gamma_b) + \bar{V}_1/c}{2} \right) * clk_dt(t1) \quad (21.12)$$

The symbol \bar{V}_1 present the average velocity between ideal time t1 and t2, \bar{V}_2 present the average velocity between ideal time t1 and t5.

So, a distance measurement can be evaluated in dynamic environment by formula (21.9) to formula (21.12), using relative velocity, frequency-difference $((\gamma_a - \gamma_b))$ and time-difference $clk_dt(t1)$.

21.4 Real-Time Distance Measurement

The ranging value of DOWR is time-laged, while the processing in satellite A needs the pseudo-distance of satellite B, thus the ranging value is the distance before current time. Assume current time is t0, the time of distance measurement is t1, the distance can be expressed as follows:

$$R(t0) = R(t1) + (t1 - t0) * \bar{V} \quad (21.13)$$

The peer measurement has not reached at current time, the DOWR way can not be used to get relative velocity by only local pseudo-speed information. But the frequency-difference is almost the same as $(\gamma_a - \gamma_b)$, because of the stability of atomic clock. The relative velocity also can be evaluated using local pseudo-speed information according formula (21.8):

$$\bar{V} = c * (ps_A(\Delta t) - (\gamma_a - \gamma_b)) \quad (21.14)$$

Then the distance can be extrapolated to current time, which means the real-time measurement of distance.

21.5 Hardware-in-Loop Simulation

The validity and precision of this algorithm have been verified in certain product. The simulation system using m-serial to generate PRN-code, with the code-rate of 10.23 Mchips/s. the carrier is in Ka band, and the IF frequency is 140 MHz. In order to get high resolution of PRN-code's phase, a sample frequency of 82 MHz based on non-equal method is adopted. The carrier tracking loop uses 2-order FLL along with 3-order PLL, a classical Costas loop are used in the PLL, and the code tracking loop uses 2-order loop.

A hardware-in-loop simulation system is established using channel-simulation-equipment. The equipment has two independency channel, to simulate forward-link and backward-link separately. The dynamic environment can be constructed by the control of time-delay, Doppler-frequency and the amplitude of signal. The parameters in the experiment are: initial distance equals 700 km, max speed equals 240 m/s, max acceleration equals 1.5 g. The track is a sine-curve, and the max displacement equals about 4 km. The C/N0 is about 50 dB-Hz in both satellite A and satellite B. The simulation curve is shown in Fig. 21.4.

The precision of ranging is shown in Fig. 21.5, the standard deviation of ranging error is 0.047 m.

21.6 The End

The Dual One Way Ranging (DOWR) technique with PRN-code can remove the effect of time-asynchronous and improve ranging precision. But in dynamic environment, the transmission-delay and clock-diff may not be matched in local pseudo-distance measurement, while the local time of each satellite will drift because of the frequency of crystal oscillator. So there are errors in the ranging value of DOWR. On the other hand, the distance of DOWR is time-laged, the present ranging value doesn't match current time. A combined-process method using time-difference,

Fig. 21.4 Dynamic environment curve

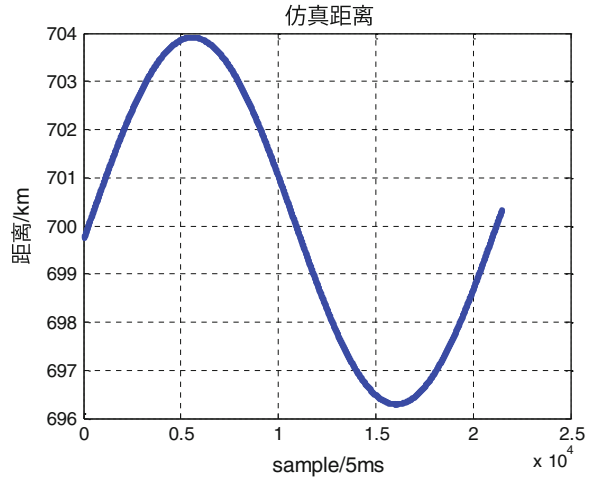
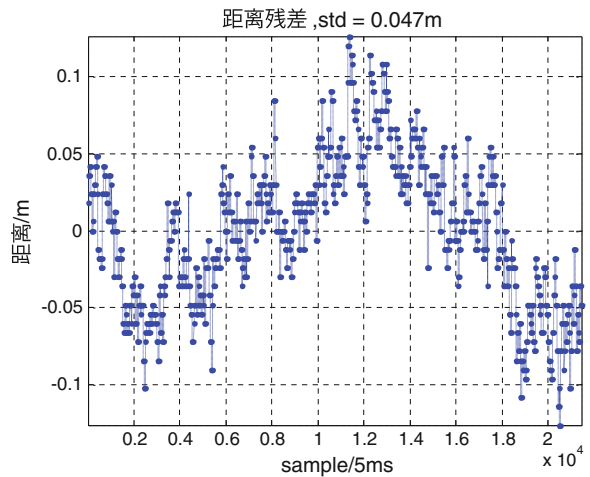


Fig. 21.5 Ranging precision in dynamic environment



frequency-difference and relative velocity are proposed in this paper. a general analytic expression of distance and clock-diff is constructed in different receiving patterns, thus the distance of certain time can be evaluated. And then the distance can be extrapolated to current time using local pseudo-speed information, which means the real-time measurement of distance in dynamic environment.

There are some approximate operations in the algorithm, especially the equalization of ideal time and the local time when the time-difference is not large. The approximate operations can bring errors in the situation which requires better precision.

References

1. Grelier T, Ghion A, Dantepal J et al (2007) Compass signal structure and first measurements. In: ION 20th International technical meeting of the satellite division, Fort Worth, Texas, USA, 25–28 Sept 2007
2. Hanson DW (1989) Fundamentals of two-way time transfers by satellite. In: The 43rd annual symposium on frequency and control, Denver, Colorado, USA, 31 May–3 June 1989
3. Huang Z, Lu JH (2003) Space-based passive positioning and modern small satellite technology. *J Inst Command Technol* 4(3):24–29
4. Chen YG, Li CJ, Li XH (2004) A precision analyzing & reckoning model in tri-station TDOA location. *ACTA Electron Sin* 2(9):1453–1455
5. Guo FC, Fan Y (2008) A method of dual-satellites geolocation using TDOA and FDOA and its precision analysis. *J Astronaut* 29(04):1381–1386
6. Zhong XW, Chen H (2010) Technology research on timing and frequency measurement system in satellite constellations. *J Astronaut* 31(04):1110–1116
7. Li X, Zhang QS, Xi QL, Zhong XW, Xiong ZF (2008) Intra-satellite baseline measurement via asynchronous communication link of autonomous formation flyer. *J Astronaut* (29)4: 1369–1374
8. Zhong XW, Chen H (2007) The analysis and correction techniques of movement influence on inter-satellite two-way time transfer. *Chin Space Sci Technol* 34(2):182–187

Chapter 22

Multipath Effect on Phase Center Calibration of GNSS Antenna

Lixun Li, Baiyu Li, Huaming Chen and Feixue Wang

Abstract Conventional phase center calibration is carried out in anechoic chamber, when in the outfield, the calibration values may be not useful because of the phase errors arose by multipath. In the paper, we present a model of antenna receiving signal; the model is improved by adding antenna model and the characterizations of right hand circular polarization and left hand circular polarization. Based on the model, we analysis the impact of up-to-down ratio on phase errors. As an example, we use a choke antenna to assess the performance of the simulation results. When the phase center variations tolerance is 2 mm, in order to mitigate the impact of phase center variations on high precision GNSS by using calibration values, the up-to-down ratios must be all higher than 15 dB.

Keywords GNSS · Carrier multipath · U/D ratio · Phase errors

22.1 Introduction

Antenna phase center offsets and variations play an important role in high precision GNSS applications, in order to mitigate the impact of phase center variations on high precision GNSS, a careful determination of antenna phase center offsets and variations is required. Conventional phase center calibration is carried out in anechoic chamber, when the antenna is used in outfield, the calibration values may be not useful because of the phase errors arose by multipath.

In work [1–3], an analytical expression is presented for calculating the effects of multipath on carrier phase. It considers that the antenna may be modeled as point source, the characters of polarization and radiation are ignored, while the receiving signal phase changes because of varies polarization and radiation. Work [3–5] study

L. Li (✉) · B. Li · H. Chen · F. Wang

College of Electronic Science and Engineering, National University of Defense Technology,
Changsha 410073, Peoples Republic of China
e-mail: lilixun1985@163.com

© Springer-Verlag Berlin Heidelberg 2015

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015*

Proceedings: Volume III, Lecture Notes in Electrical Engineering 342,

DOI 10.1007/978-3-662-46632-2_22

the characterizations of reflected signal based on the up-to-down (U/D) ratio. But the U/D only includes right hand circular polarization (RHCP) gain, in practical terms, the reflected signal includes RHCP and left hand circular polarization (LHCP) signal. The U/D must include LHCP gain too.

In the paper, we present a novel model of antenna receiving signal, the model is further improved from [1–5] by adding antenna model and the characters of RHCP and LHCP. Based on the model, two U/D ratios are present. We also analysis the impact on phase errors arose by multipath of U/D. As an example, we use a choke antenna to assess the performance of the simulation results. We found the multipath affect the phase evidently.

22.2 Signal Model

22.2.1 Antenna Model

In general, the antenna polarization includes RHCP, LHCP and linear polarization (LP), and the LP can be converted to linear combinations of RHCP and LHCP, so the universal antenna model is written by:

$$E_{antenna}(\theta, \varphi) = \begin{bmatrix} |E_L(\theta, \varphi)|e^{j\theta_L(\theta, \varphi)} \\ |E_R(\theta, \varphi)|e^{j\theta_R(\theta, \varphi)} \end{bmatrix} \quad (22.1)$$

where $|E_L(\theta, \varphi)|$, $\theta_L(\theta, \varphi)$ are the gain and phase of the LHCP output in (θ, φ) direction, $|E_R(\theta, \varphi)|$, $\theta_R(\theta, \varphi)$ are the gain and phase of the RHCP output in (θ, φ) direction, θ is the elevation angle, φ is the azimuth angle. In the paper, the horizontal plane is $\theta = 0^\circ$.

22.2.2 Model of Receiving Signal

22.2.2.1 Model of Line-of-Sight Signal

The line-of-sight (LOS) signal can be written by the following formula.

$$s_d(t) = Ae^{j(\omega t + \theta_0)}, P_d \quad (22.2)$$

Where A is Signal amplitude, ω is radian frequency, θ_0 is Initial signal phase, P_d is the polarization of the LOS signal.

$$P_d = \begin{bmatrix} \sin \gamma e^{j\varphi} \\ \cos \gamma e^{-j\varphi} \end{bmatrix} \quad (22.3)$$

Where $e^{j\varphi}$, $e^{-j\varphi}$ are the phase accumulate arose by signal polarization, which is the typical character of circular polarization wave.

When $\gamma = 45^\circ$: LP. $\gamma = 90^\circ$: LHCP. $\gamma = 0^\circ$: RHCP. The satellite signal is RHCP in GNSS, so in the paper:

$$P_d = \begin{bmatrix} 0 \\ e^{-j\varphi} \end{bmatrix} \quad (22.4)$$

22.2.2.2 Model of Multipath Signal

Multipath signal is introduced by reflection from the ground around the antenna. Compared to the LOS signal, the amplitude, phase and polarization of the multipath signal change. The multipath signal is written by:

$$s_m(t) = \alpha A e^{j(\omega t + \theta_0 + \theta_m)}, P_m \quad (22.5)$$

where α is the attenuation factor which varies between 0 and 1, P_m is the polarization of the multipath signal, θ_m is the phase change due to the reflector.

22.2.3 Ground Reflection Model

For a linear system, the Fresnel Reflection factors are written by the following formula.

$$r_{\perp} = \frac{\sin \theta - \sqrt{\varepsilon_1 - j \frac{\delta_1}{\omega \varepsilon_0}} \sqrt{1 - \frac{1}{\varepsilon_1} \cos^2 \theta}}{\sin \theta + \sqrt{\varepsilon_1 - j \frac{\delta_1}{\omega \varepsilon_0}} \sqrt{1 - \frac{1}{\varepsilon_1} \cos^2 \theta}} \quad (22.6)$$

$$r_{//} = \frac{\sqrt{\varepsilon_1 - j \frac{\delta_1}{\omega \varepsilon_0}} \sin \theta - \sqrt{1 - \frac{1}{\varepsilon_1} \cos^2 \theta}}{\sqrt{\varepsilon_1 - j \frac{\delta_1}{\omega \varepsilon_0}} \sin \theta + \sqrt{1 - \frac{1}{\varepsilon_1} \cos^2 \theta}} \quad (22.7)$$

where r_{\perp} is the Fresnel reflection factor for vertical polarization, $r_{//}$ is the Fresnel reflection factor for horizontal polarization, θ is the incident angle, ε_0 is the Free space dielectric, ε_1 is the Relative dielectric constant, σ is the Conductivity.

The electric signal can be written by:

$$E = \begin{bmatrix} E_L \\ E_R \end{bmatrix} \quad (22.8)$$

where the E_L , E_R are the complex phasor LHCP and RHCP components of the electric field.

Convert the Eq. (22.8) to following formula.

$$\begin{bmatrix} E_H \\ E_V \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{vmatrix} 1 & 1 \\ i & -i \end{vmatrix} \begin{bmatrix} E_L \\ E_R \end{bmatrix} \quad (22.9)$$

where the E_H , E_V are the complex phasor horizontal and vertical components of the signal.

The reflected signal can be written by the following formula.

$$\begin{bmatrix} E_{H_R} \\ E_{V_R} \end{bmatrix} = \begin{bmatrix} r_{//} & 0 \\ 0 & r_{\perp} \end{bmatrix} \begin{bmatrix} E_H \\ E_V \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} r_{//} & 0 \\ 0 & r_{\perp} \end{bmatrix} \begin{vmatrix} 1 & 1 \\ i & -i \end{vmatrix} \begin{bmatrix} E_L \\ E_R \end{bmatrix} \quad (22.10)$$

where the E_{H_R} , E_{V_R} are the complex phasor horizontal and vertical components of the reflected wave.

$$\begin{bmatrix} E_{L_R} \\ E_{R_R} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{vmatrix} 1 & -i \\ 1 & i \end{vmatrix} \begin{bmatrix} E_{H_R} \\ E_{V_R} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} r_{//} + r_{\perp} & r_{//} - r_{\perp} \\ r_{//} - r_{\perp} & r_{//} + r_{\perp} \end{bmatrix} \begin{bmatrix} E_L \\ E_R \end{bmatrix} \quad (22.11)$$

where the E_{L_R} , E_{R_R} are the complex phasor LHCP and RHCP components of the reflected wave.

Adding the accumulated phase arose by circular polarization to Eq. (22.11)

$$\begin{bmatrix} E_{L_R} \\ E_{R_R} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^{j\varphi} & 0 \\ 0 & e^{-j\varphi} \end{bmatrix} \begin{bmatrix} r_{//} + r_{\perp} & r_{//} - r_{\perp} \\ r_{//} - r_{\perp} & r_{//} + r_{\perp} \end{bmatrix} \begin{bmatrix} E_L \\ E_R \end{bmatrix} \quad (22.12)$$

For a circular polarization system, the Fresnel Reflection factor are written by the following formula

$$R = \frac{1}{2} \begin{bmatrix} (r_{//} + r_{\perp})e^{j\varphi} & (r_{//} - r_{\perp})e^{j\varphi} \\ (r_{//} - r_{\perp})e^{-j\varphi} & (r_{//} + r_{\perp})e^{-j\varphi} \end{bmatrix} \quad (22.13)$$

22.2.4 Antenna Receiving Signal

The LOS signal received by the antenna can be written by the following formula.

$$s_{id}(t) = E_{antenna}^*(\theta, \varphi) s_d(t) = A e^{j(\omega t + \theta_0)} E_{antenna}^*(\theta, \varphi) \cdot P_d \quad (22.14)$$

Compared to the LOS signal, the multipath signal received by the antenna can be written by:

$$s_{im}(t) = E_{antenna}^*(-\theta, \varphi) \cdot s_m(t) = A e^{j(\omega t + \theta_0 + \theta_m)} E_{antenna}^*(-\theta, \varphi) \cdot R^* \cdot P_d \quad (22.15)$$

22.2.5 Phase Errors Arose by Multipath

There are two important multipath sources; signal reflected by a building and off the ground, in high precision GNSS applications, the buildings are beneath the antenna, the signal reflected off the ground is the key multipath sources. So in Eq. (22.5), $\theta_m = \frac{4\pi h \sin \theta}{\lambda}$, where θ is the incident angle, h is the height of the antenna above the reflected plane.

The received signal is given as the sum of the LOS and multipath signals:

$$\begin{aligned} s(t) &= s_{id}(t) + s_{im}(t) = s_{id}(t) \left(1 + \frac{e^{j\theta_m} E_{antenna}^*(-\theta, \varphi) \cdot R^* \cdot P_d}{E_{antenna}^*(\theta, \varphi) \cdot P_d} \right) \\ &= s_{id}(t) (1 + \alpha e^{j\phi}) \end{aligned} \quad (22.16)$$

where

$$\begin{aligned} \phi &= \angle \frac{e^{j\theta_m} E_{antenna}^*(-\theta, \varphi) \cdot R^* \cdot P_d}{E_{antenna}^*(\theta, \varphi) \cdot P_d} \\ \alpha &= \left| \frac{e^{j\theta_m} E_{antenna}^*(-\theta, \varphi) \cdot R^* \cdot P_d}{E_{antenna}^*(\theta, \varphi) \cdot P_d} \right| \end{aligned} \quad (22.17)$$

So the phase errors arose by multipath can be given by

$$\Delta\phi = \angle s(t) - \angle s_{id}(t) = ar \sin \left(\frac{\sin \phi}{1 + \alpha^2 + 2\alpha \cos \phi} \right) \quad (22.18)$$

22.3 Simulation Results

The typically high-precision GNSS antenna has stable phase center, so in the paper, we assume the phase center variation of antenna is zero and the $\varphi = 0^\circ$, the conclusion is the same with others φ values.

The wet ground and dry ground are two typical reflected materials, and the approximate values for dielectric constant and conductivity are given in Table 22.1:

In our simulation, we choose the frequency $f = 1575.42$ MHz ($\omega = 2\pi f$), and the height of the antenna $h = 2$ m.

Equations (22.4) and (22.17) can be arranged the following formula.

$$\begin{aligned}\phi &= \angle \frac{|E_L(-\theta, \varphi)|(r_{//} - r_{\perp}) + |E_R(-\theta, \varphi)|(r_{//} + r_{\perp})}{2|E_R(\theta, \varphi)|} \\ \alpha &= \left| \frac{|E_L(-\theta, \varphi)|(r_{//} - r_{\perp}) + |E_R(-\theta, \varphi)|(r_{//} + r_{\perp})}{2|E_R(\theta, \varphi)|} \right|\end{aligned}\quad (22.19)$$

As Eq. (22.19) predicts, we will see that the two major factors affecting the phase errors are the relative amplitude between the up RHCP gain and down RHCP gain and the relative amplitude between the up RHCP gain and down LHCP gain. Two formulas for U/D ratio are given by:

$$\begin{aligned}U/D_1(\theta, \varphi) &= \frac{|E_R(\theta, \varphi)|}{|E_R(-\theta, \varphi)|} \\ U/D_2(\theta, \varphi) &= \frac{|E_R(\theta, \varphi)|}{|E_L(-\theta, \varphi)|}\end{aligned}\quad (22.20)$$

By rearranging Eq. (22.20), we can get the following relationship.

$$\begin{aligned}\phi &= \angle \left(\frac{(r_{//} + r_{\perp})}{2U/D_1} + \frac{(r_{//} - r_{\perp})}{2U/D_2} \right) e^{j\theta_m} \\ \alpha &= \left| \left(\frac{(r_{//} + r_{\perp})}{2U/D_1} + \frac{(r_{//} - r_{\perp})}{2U/D_2} \right) e^{j\theta_m} \right|\end{aligned}\quad (22.21)$$

Table 22.1 Characters of wet and dry ground

Ground type	Relative dielectric constant	Conductivity
Wet ground	16	0.002
Dry ground	4	0.001

22.3.1 Dry Ground

Figure 22.1 shows the phase errors arose by dry ground reflected multipath changing against elevation for several values of U/D_1 and U/D_2 .

As shown in Fig. 22.1. We note that as U/D_1 and U/D_2 ratios increase, the phase errors decrease, when U/D_1 and U/D_2 are all higher than 15 dB, the max phase error is 0.68° (0.36 mm), when whichever of U/D_1 and U/D_2 is lower than 15 dB, the min phase error is 4.56° (2.41 mm).

22.3.2 Wet Ground

Figure 22.2 shows the phase errors arose by wet ground reflected multipath changing against elevation for several values of U/D_1 and U/D_2 .

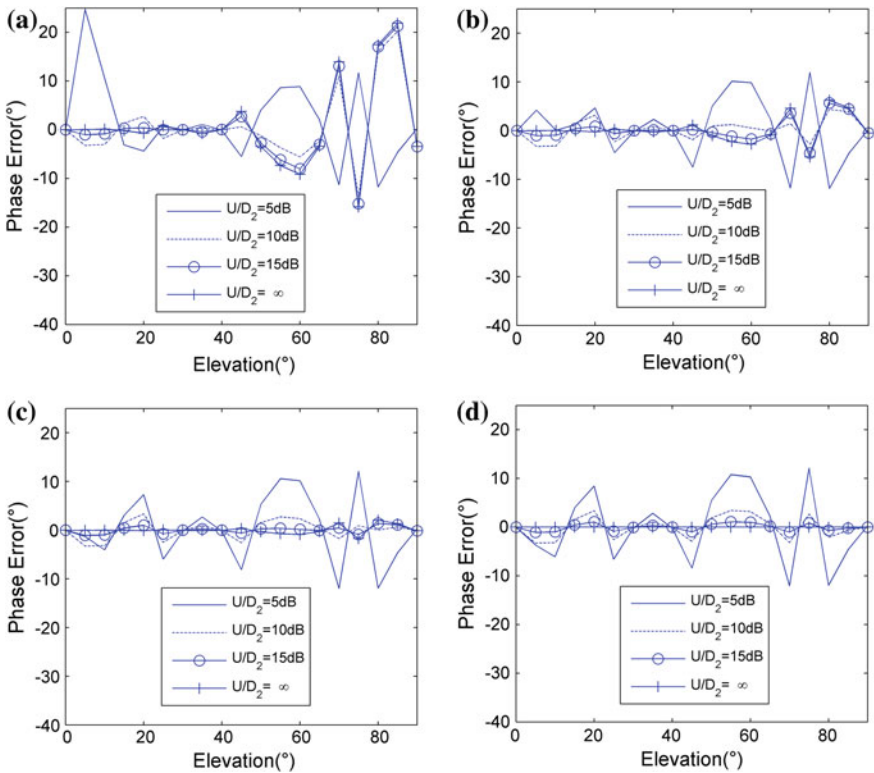


Fig. 22.1 Effect of the multipath around dry ground on the carrier phase against elevation angle. **a** $U/D_1 = 5$ dB **b** $U/D_1 = 10$ dB **c** $U/D_1 = 15$ dB **d** $U/D_1 = \infty$

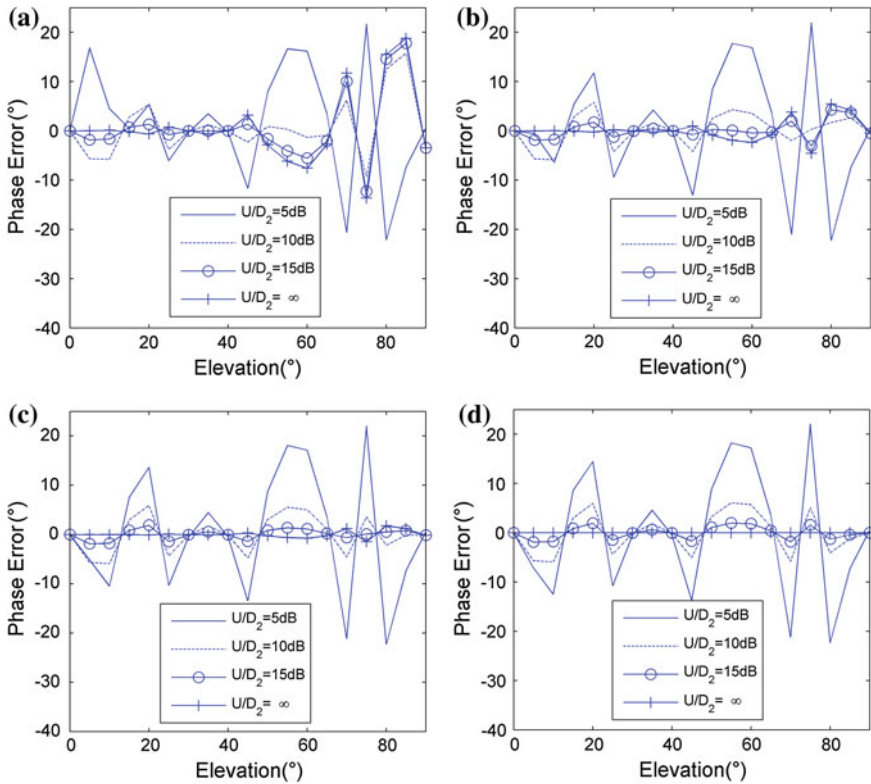


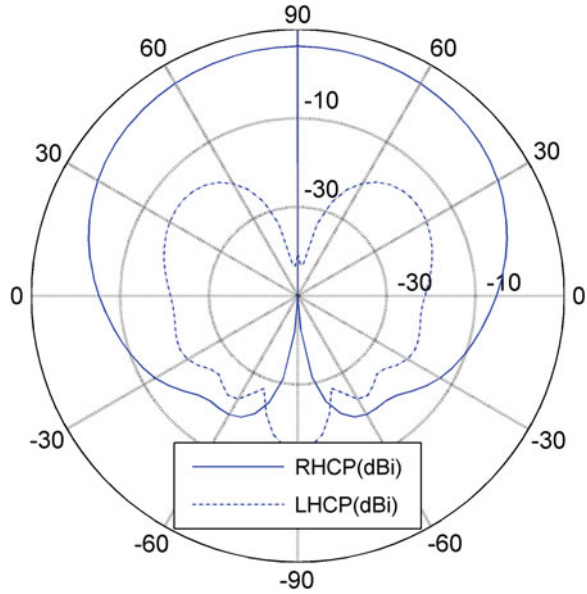
Fig. 22.2 Effect of the multipath around wet ground on the carrier phase against elevation angle. **a** $U/D_1 = 5$ dB **b** $U/D_1 = 10$ dB **c** $U/D_1 = 15$ dB **d** $U/D_1 = \infty$

As shown in Fig. 22.2, We note that as U/D_1 and U/D_2 ratio increase, the phase errors decrease, when U/D_1 and U/D_2 are all higher than 15 dB, the max phase error is 1.48° (0.78 mm), when whichever of U/D_1 and U/D_2 is lower than 15 dB, the min phase error is 5.53° (2.93 mm).

22.3.3 Summary

When the phase center variations tolerance is 2 mm (typically), in order to mitigate the impact of phase center variations on high precision GNSS by using calibration values, the U/D_1 and U/D_2 ratios must be all higher than 15 dB.

Fig. 22.3 The RH and LHCP gain patterns for the choke antenna at L1 ($\phi = 0$)



22.4 Examples

In order to assess the performance of the simulation results, a choke antenna is analyzed; Fig. 22.3 shows the RH and LHCP gain pattern of the antenna.

As presented in Fig. 22.4, the U/D_2 of the antenna is higher than 15 dB above horizontal, while the U/D_1 is less than 15 dB at low-elevation angles ($\theta \leq 25^\circ$).

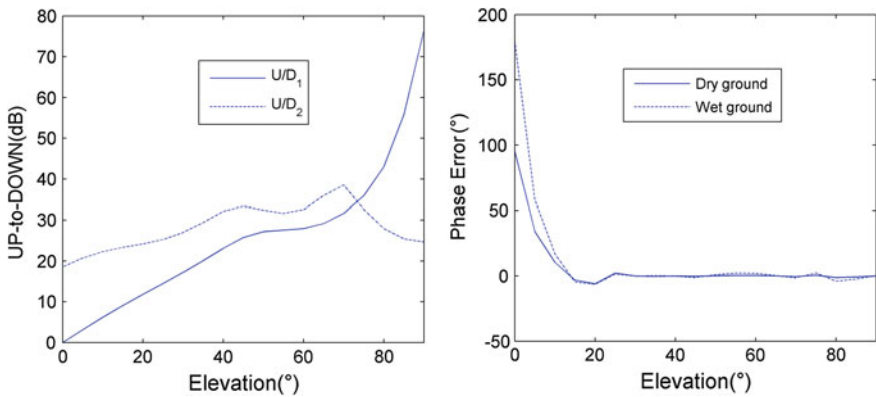


Fig. 22.4 The U/D ratios in dB (left plot) and the effect of the multipath on the carrier phase (right plot) against elevation angle for choke antenna

We note that the min phase error are 1.65° (0.87 mm) for wet ground reflected and 4.47° (2.37 mm) for dry ground reflected at low-elevation angles ($\theta \leq 25^\circ$) because that the U/D_1 is less than 15 dB at high-elevation angles ($\theta > 25^\circ$), the U/D_1 and U/D_2 are all higher than 15 dB, the max phase error are 2.32° (1.23 mm) for wet ground reflected and 2.17° (1.15 mm) for dry ground reflected.

When the phase center variations tolerance is 2 mm, in order to mitigate the impact of phase center variations on high precision GNSS application by using calibration values, the elevation cut-off angle will be set at 25° or higher.

22.5 Conclusions

In the paper, we have applied the new antenna and multipath model to illustrate the phase errors arose by multipath against U/D_1 and U/D_2 , it was shown that the multipath affect the phase evidently, when whichever of U/D_1 and U/D_2 ratio is below 15 dB, the phase errors are higher than 2 mm, which exceeds the phase center variations tolerance. In high precision GNSS application, before using calibration values, following three steps are necessary:

1. Obtain the RHCP and LHCP gain patterns.
2. Calculate the U/D_1 and U/D_2 ratio of the antenna.
3. Set the elevation cutoff based on the U/D_1 and U/D_2 ratio (>15 dB).

In order to expand the coverage of antenna applied in high precision GNSS application, we must heighten the U/D_1 and U/D_2 ratio at low-elevation angle.

References

1. Godet J, Dantepal J, Grondin M, et al (1999) Multipath error analysis and calibration for GPS differential phase. In: The ION GPS National technical meeting, 521–530
2. Wanninger L, May M (2001) Carrier phase multipath calibration of GPS reference stations. *Navigation* 48:113–124
3. Charles C (1999) Multipath-rejecting GPS antennas. *Proc IEEE* 87:86–91
4. Liu H (2010). Research on the multipath effects and mitigation techniques in satellite navigation systems. Graduate School of National University of Defense Technology
5. Sun L (2011). Research on anti-jamming and multipath mitigation by reduced distributed vector sensor in satellite navigation systems. Graduate School of National University of Defense Technology

Chapter 23

An Improved TCAR Based on the Optimal Combination of Carrier Phase and Pseudo-range Observations

Xing Wang, Wenxiang Liu, Yangbo Huang and Guangfu Sun

Abstract The core issue for high precision relative dynamic positioning application is how to fix integer carrier ambiguity fast and accurately. Traditional three carrier ambiguity resolution (TCAR) method find the suitable combination of carrier phase and pseudo-range observations using integer search traversal method, which make the lower success reliability for narrow-lane ambiguity resolution. In the analysis of the error characteristics of the combined observations, we introduce a new procedure to select the optimal combination of carrier phase and pseudo-range observations based on the minimum of total noise level, with real-time estimating accurately and mitigating even eliminating the impact of ionospheric delay, successfully fixed three linearly independent combinations of the integer ambiguity, then determine the basic integer ambiguities. Finally, the performance of the improved TCAR in short-baseline condition is showed by processing the observations collected by two BeiDou triple-frequency receivers, the results show that, this method fix the integer ambiguities of double differenced extra-wide lane and wide line from single epoch, total noise level was decreased almost 15 %, with several epochs smoothing, the success reliability of basic ambiguities achieve 95 %, which is improved at least 30 % compare with traditional TCAR method.

Keywords GNSS · Three carrier ambiguity resolution · Optimal combination · Ionospheric delay estimation

X. Wang (✉) · W. Liu · Y. Huang · G. Sun
College of Electronic Science and Engineering, National University of Defense Technology,
Changsha, Hunan, People's Republic of China
e-mail: wangxing-1010@163.com

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_23

269

23.1 Introduction

Currently, the new generation of global navigation satellite systems (GNSS) transmit three or more frequency signals, such as GPS, Galileo, Chinese BeiDou Navigation Satellite System, etc. Multi-frequency signals not only mean more observations of pseudo-range and carrier phase, but also provide various types of linear combination of observations to meet different needs, which is benefit for ambiguity resolution. A method was proposed to cascading ambiguity resolution which rounds the float value of ambiguity to the nearest integer without complex searching calculation. The TCAR [1, 2] and cascading integer resolution (CIR) [3] are significantly researched as this kind of method. The premise of these methods achieving higher success reliability for ambiguity resolution is that the noise of combined observations is far less than the combined wavelength. It is usually comforted for extra-wide line (EWL: $\lambda \geq 2.93$ m) and wide-line (WL: $2.93 > \lambda \geq 0.75$ m) ambiguity resolution, but for medium-line (ML: $0.75 > \lambda \geq 0.19$ m) and narrow-line (NL: $\lambda < 0.19$ m) ambiguity resolution, this condition is hard to satisfied. Especially in the long-baseline conditions, double differenced (DD) of the code and carrier phase measurements are difficult to eliminate the ionospheric delay, tropospheric delay and other errors completely. The means of these errors are not zero, which can not be eliminated or minished through multi-epoch smoothing, make it hard to fix the integer ambiguity and influent the accuracy of positioning.

Generally, various TCAR methods can be classified into two kinds of models: the geometry-free and geometry-based integer determination models. The geometry-free model refers to the observation model without the distance between the receiver and satellite and eliminates the geometry-based error, such as tropospheric delay and orbital error and so on, also with simple calculation for the ambiguity resolution, which still has a lot of researches. This paper focus on geometry-free model and presents an optimal combination of carrier phase and pseudo-range observations to improve TCAR algorithm, mainly has the following three improvements: firstly, by calculating and using the optimal combined coefficients, reduces the total noise level which contain both ionospheric delay and observation errors, makes it easily to fix the integer ambiguity; secondly, through the real-time estimation of ionospheric delay by using the Hatch filter, the estimation accuracy of ionospheric delay and success probability for ambiguity resolution are great improved; lastly, by using a linear transformation to calculate the basic integer ambiguities through three groups using independent linear combined integer ambiguities, avoiding the lower success probability for direct calculation of basic integer ambiguities.

23.2 Combination of Three Frequency GNSS Observations

For researching combination models of TCAR method, the observation equations of carrier phase and code measurements are first introduced. With any three GNSS signals which can be set as f_1, f_2 and f_3 , considering the impact of the satellite orbital error, the tropospheric delay and the first-order ionospheric delay which can not be fully eliminated by DD observations in the long-baseline, the fundamental DD code and phase observation equations can be expressed as:

$$\Delta P_i = \Delta\rho + \Delta\delta_{orb} + \Delta\delta_{tro} + \frac{f_1^2}{f_i^2} \Delta\delta I_1 + \varepsilon_{\Delta P_i} \quad (23.1)$$

$$\Delta\Phi_i = \Delta\rho + \Delta\delta_{orb} + \Delta\delta_{tro} - \frac{f_1^2}{f_i^2} \cdot \Delta\delta I_1 - \lambda_i \cdot \Delta N_i + \varepsilon_{\Delta\Phi_i} \quad (23.2)$$

where “ Δ ” is the DD operator product; ΔP_i and $\Delta\Phi_i$ are the DD code and phase measurements for frequency f_i ; $\Delta\rho$ is the DD geometric distance from satellite to receiver; $\Delta\delta_{orb}$ and $\Delta\delta_{tro}$ are the DD orbital error and tropospheric delay; $\Delta\delta I_1$ is the first-order ionospheric delay for frequency f_1 ; $\varepsilon_{\Delta P_i}$ and $\varepsilon_{\Delta\Phi_i}$ are the noise of code and carrier phase measurements.

23.2.1 Geometry-Based Models for Carrier Phase Combination

The geometry-based models combination of carrier phase can be generally formulated as [4]:

$$\begin{aligned} \Delta\Phi_{(i,j,k)} &= \frac{i \cdot f_1 \cdot \Delta\Phi_1 + j \cdot f_2 \cdot \Delta\Phi_2 + k \cdot f_3 \cdot \Delta\Phi_3}{i \cdot f_1 + j \cdot f_2 + k \cdot f_3} \\ &= \Delta\rho + \Delta\delta_{orb} + \Delta\delta_{tro} - \beta_{(i,j,k)} \cdot \Delta\delta I_1 - \lambda_{(i,j,k)} \cdot \Delta N_{(i,j,k)} + \varepsilon_{\Delta\Phi_{(i,j,k)}} \end{aligned} \quad (23.3)$$

where the coefficients i, j, k are arbitrary integer; the combined wavelength $\lambda_{(i,j,k)}$ and combined integer ambiguity $\Delta N_{(i,j,k)}$ are defined as:

$$\lambda_{(i,j,k)} = \frac{c}{i \cdot f_1 + j \cdot f_2 + k \cdot f_3} = \frac{\lambda_1 \cdot \lambda_2 \cdot \lambda_3}{i \cdot \lambda_2 \cdot \lambda_3 + j \cdot \lambda_1 \cdot \lambda_3 + k \cdot \lambda_1 \cdot \lambda_2} \quad (23.4)$$

$$\Delta N_{(i,j,k)} = i \cdot \Delta N_1 + j \cdot \Delta N_2 + k \cdot \Delta N_3 \quad (23.5)$$

The combined first-order ionospheric delay factor $\beta_{(i,j,k)}$ is defined as:

$$\beta_{(i,j,k)} = \frac{f_1^2(i/f_1 + j/f_2 + k/f_3)}{i \cdot f_1 + j \cdot f_2 + k \cdot f_3} \quad (23.6)$$

Assuming the noise levels for all three frequency carrier phase measurements are the same and unrelated, i.e. $\varepsilon_{\Delta\Phi 1}^2 = \varepsilon_{\Delta\Phi 2}^2 = \varepsilon_{\Delta\Phi 3}^2 \equiv \varepsilon_{\Delta\Phi}^2$, the variance of the combined DD carrier phase noise is:

$$\varepsilon_{\Delta\Phi(i,j,k)}^2 = \frac{(i \cdot f_1)^2 + (j \cdot f_2)^2 + (k \cdot f_3)^2}{(i \cdot f_1 + j \cdot f_2 + k \cdot f_3)^2} \cdot \varepsilon_{\Delta\Phi}^2 \equiv \mu_{(i,j,k)}^2 \cdot \varepsilon_{\Delta\Phi}^2 \quad (23.7)$$

23.2.2 Geometry-Based Model for Code Measurements Combination

For convenient analysis, we introduce the geometry-based model for combination of code measurements as follow with the condition: $l + m + n = 1$.

$$\begin{aligned} \Delta P_{(l,m,n)} &= l \cdot \Delta P_1 + m \cdot \Delta P_2 + n \cdot \Delta P_3 \\ &= \Delta\rho + \Delta\delta_{orb} + \Delta\delta_{tro} + \beta_{(l,m,n)} \cdot \Delta\delta I_1 + \varepsilon_{\Delta P(l,m,n)} \end{aligned} \quad (23.8)$$

The combined first-order ionospheric delay factor $\beta_{(l,m,n)}$ is defined as:

$$\beta_{(l,m,n)} = l + \frac{f_2^2}{f_1^2}m + \frac{f_3^2}{f_1^2}n = q_1 \cdot l + q_2 \cdot m + q_3 \cdot n \quad (23.9)$$

where $q_i = (f_1/f_i)^2$ is the first-order ionospheric delay factor of frequency f_i . Also assuming that $\varepsilon_{\Delta P 1}^2 = \varepsilon_{\Delta P 2}^2 = \varepsilon_{\Delta P 3}^2 \equiv \varepsilon_{\Delta P}^2$, the variance of the combined DD code noise is:

$$\varepsilon_{\Delta P(l,m,n)}^2 = (l^2 + m^2 + n^2)\varepsilon_{\Delta P}^2 \equiv \mu_{(l,m,n)}^2 \cdot \varepsilon_{\Delta P}^2 \quad (23.10)$$

23.2.3 Ambiguity Resolution Using the Combination of Carrier Phase and Code Measurements Based on Geometry-Free Model

The tradition TCAR method resolves the ambiguity by rounding the float ambiguity value to its nearest integer. The float ambiguity value is determined by geometry-free

model combination of carrier phase and code measurements to eliminate or reduce the geometry-related terms, such as the geometric distance from satellite to receiver, the orbital error and tropospheric delay [5]:

$$\Delta N_{(i,j,k)} = \frac{\Delta P_{(l,m,n)} - \Delta \Phi_{(i,j,k)}}{\lambda_{(i,j,k)}} - \frac{\beta_{(l,m,n)} + \beta_{(i,j,k)}}{\lambda_{(i,j,k)}} \Delta \delta I_1 - \frac{\varepsilon_{\Delta P_{(l,m,n)}} - \varepsilon_{\Delta \Phi_{(i,j,k)}}}{\lambda_{(i,j,k)}} \quad (23.11)$$

The total noise level in cycle for combined ambiguity resolution is:

$$\sigma_{TN} = \frac{\sqrt{\left(\beta_{(l,m,n)} + \beta_{(i,j,k)}\right)^2 \cdot (\Delta \delta I_1)^2 + \varepsilon_{\Delta P_{(l,m,n)}}^2 + \varepsilon_{\Delta \Phi_{(i,j,k)}}^2}}{\lambda_{(i,j,k)}} \quad (23.12)$$

It is clear that, the smaller value of σ_{TN} means the better performance for ambiguity resolution.

23.3 The Improved TCAR Based on the Optimal Combination of Carrier Phase and Pseudo-range Observations

Traditional TCAR method finds the suitable combined coefficients of carrier phase and pseudo-range using integer search traversal method, which usually does not achieve the minimal total noise level.

If the optimal coefficients (l, m, n) of pseudo-range for giving combined coefficients of carrier phase (i, j, k), which reach the minimal total noise level, can be found, we can achieve the highest success reliability for ambiguity resolution.

23.3.1 Optimal Combination of Code Measurements

For giving combined coefficients i, j, k of carrier phase, the combined wavelength $\lambda_{(i,j,k)}$ is a positive number, so the condition of $\sigma_{TN} = \min$ can be the same with:

$$\left(\beta_{(l,m,n)} + \beta_{(i,j,k)}\right)^2 \cdot (\Delta \delta I_1)^2 + \varepsilon_{\Delta P_{(l,m,n)}}^2 + \varepsilon_{\Delta \Phi_{(i,j,k)}}^2 = \min \quad (23.13)$$

Therefore, the problem is equivalent to finding the extreme value in the conditional function. Then, the Lagrange function $f(l, m, n, \theta)$ is established as follow:

$$f(l, m, n, \theta) = \left(\beta_{(l,m,n)} + \beta_{(i,j,k)} \right)^2 \cdot (\Delta\delta I_1)^2 + \mu_{(l,m,n)}^2 \cdot \varepsilon_{\Delta P}^2 + \mu_{(i,j,k)}^2 \cdot \varepsilon_{\Delta\Phi}^2 + \theta(l + m + n - 1) \quad (23.14)$$

Calculating the first derivative of the function mentioned above and setting the value equal to zero:

$$\begin{aligned} \frac{\partial f(l, m, n, \theta)}{\partial l} &= 2q_1 \cdot \left(q_1 \cdot l + q_2 \cdot m + q_3 \cdot n + \beta_{(i,j,k)} \right) \cdot (\Delta\delta I_1)^2 + 2l \cdot \varepsilon_{\Delta P}^2 + \theta = 0 \\ \frac{\partial f(l, m, n, \theta)}{\partial m} &= 2q_2 \cdot \left(q_1 \cdot l + q_2 \cdot m + q_3 \cdot n + \beta_{(i,j,k)} \right) \cdot (\Delta\delta I_1)^2 + 2m \cdot \varepsilon_{\Delta P}^2 + \theta = 0 \\ \frac{\partial f(l, m, n, \theta)}{\partial n} &= 2q_3 \cdot \left(q_1 \cdot l + q_2 \cdot m + q_3 \cdot n + \beta_{(i,j,k)} \right) \cdot (\Delta\delta I_1)^2 + 2n \cdot \varepsilon_{\Delta P}^2 + \theta = 0 \\ \frac{\partial f(l, m, n, \theta)}{\partial \theta} &= l + m + n - 1 = 0 \end{aligned} \quad (23.15)$$

Solving these equations, the solutions can be lent for:

$$\begin{aligned} l &= \frac{\left(q_2^2 + q_3^2 - q_1q_2 - q_1q_3 + (q_2 + q_3 - 2q_1) \cdot \beta_{(i,j,k)} \right) \cdot (\Delta\delta I_1)^2 + \varepsilon_{\Delta P}^2}{2(q_1^2 + q_2^2 + q_3^2 - q_1q_2 - q_1q_3 - q_2q_3) \cdot (\Delta\delta I_1)^2 + 3\varepsilon_{\Delta P}^2} \\ m &= \frac{\left(q_1^2 + q_3^2 - q_1q_2 - q_2q_3 + (q_1 + q_3 - 2q_2) \cdot \beta_{(i,j,k)} \right) \cdot (\Delta\delta I_1)^2 + \varepsilon_{\Delta P}^2}{2(q_1^2 + q_2^2 + q_3^2 - q_1q_2 - q_1q_3 - q_2q_3) \cdot (\Delta\delta I_1)^2 + 3\varepsilon_{\Delta P}^2} \\ n &= \frac{\left(q_1^2 + q_2^2 - q_1q_3 - q_2q_3 + (q_1 + q_2 - 2q_3) \cdot \beta_{(i,j,k)} \right) \cdot (\Delta\delta I_1)^2 + \varepsilon_{\Delta P}^2}{2(q_1^2 + q_2^2 + q_3^2 - q_1q_2 - q_1q_3 - q_2q_3) \cdot (\Delta\delta I_1)^2 + 3\varepsilon_{\Delta P}^2} \end{aligned} \quad (23.16)$$

The optimal combined coefficients of code measurements make the total noise level minimum. Taking BeiDou Navigation System for example, assuming the standard deviation of DD carrier phase noise $\varepsilon_{\Delta\Phi} = 5$ mm and the standard deviation of DD pseudo-range noise $\varepsilon_{\Delta P} = 0.3$ m. According to Ref. [6], setting the DD ionospheric Delay is 0.1 m for short-baseline condition and 1 m for long-baseline condition. Also choosing $\Delta P_{(0,0,1)}$ as combined code measurements for traditional TCAR method by comparison. Table 23.1 shows the total noise level for some optimal and traditional combinations.

23.3.2 Real-Time Estimation of Ionospheric Delay

Ionospheric delay is the key limitation for geometry-free models TCAR algorithm to fix the integer ambiguity, especially for long range application. With the spatial

Table 23.1 Total noise level for optimal and traditional combined coefficients

(i, j, k)	$\lambda_{(i,j,k)}$ (m)	Improved TCAR				Traditional TCAR	
		Short-baseline ($\Delta\delta I_1 = 0.1$ m)		Long-baseline ($\Delta\delta I_1 = 1$ m)		$\Delta\delta I_1 = 0.1$ m	$\Delta\delta I_1 = 1$ m
		Coefficients (l, m, n)	σ_{TN} (cycle)	Coefficients (l, m, n)	σ_{TN} (cycle)	σ_{TN} (cycle)	σ_{TN} (cycle)
(0, -1, 1)	4.884	(0.3250, 0.3392, 0.3358)	0.0461	(0.1036, 0.4940, 0.4024)	0.0504	0.0680	0.0698
(1, 0, -1)	1.025	(0.3404, 0.3284, 0.3312)	0.1731	(0.5269, 0.1979, 0.2752)	0.1913	0.2960	0.4045
(1,1, -2)	1.297	(0.3445, 0.3255, 0.3300)	0.1453	(0.6393, 0.1193, 0.2414)	0.1775	0.2393	0.3770
(1, 3, -4)	2.765	(0.3666, 0.3101, 0.3233)	0.1272	(1.2457, -0.3050, 0.0593)	0.1911	0.2393	0.3584
(1, -4, 3)	0.557	(0.3334, 0.3333, 0.3333)	0.3284	(0.3337, 0.3331, 0.3332)	0.3284	0.5492	0.5890

correlation reducing, the DD ionospheric delay increase significantly even up to several meters, which seriously restrict accuracy for the long baseline precision positioning. If we can real-time and accurately estimate the ionospheric delay and mitigate even eliminate the impact of ionospheric delay, the success probability for ambiguity resolution will be great improved.

The improved TCAR using optimal combined pseudo-range coefficients can fix the more than two combined EWL/WL integer ambiguities from single epoch with success probability above 99 %. Feng (2008) gave the method to estimate the DD ionospheric delay for frequency f_1 by using the two combined carrier phase observations discrepancy to eliminate the geometry-based factor [6]:

$$\Delta\hat{\delta I}_1 = \frac{\Delta\Phi_{(i_1,j_1,k_1)} - \Delta\Phi_{(i_2,j_2,k_2)} - (\lambda_{(i_1,j_1,k_1)} \cdot (i_1,j_1,k_1) - \lambda_{(i_2,j_2,k_2)} \cdot \Delta N_{(i_2,j_2,k_2)})}{\beta_{(i_2,j_2,k_2)} - \beta_{(i_1,j_1,k_1)}} + \frac{-\varepsilon_{\Delta\Phi(i_1,j_1,k_1)} + \varepsilon_{\Delta\Phi(i_2,j_2,k_2)}}{\beta_{(i_2,j_2,k_2)} - \beta_{(i_1,j_1,k_1)}} \tag{23.17}$$

The standard deviation of estimation is:

$$\sigma(\Delta\hat{\delta I}_1) = \sqrt{\frac{\varepsilon_{\Delta\Phi(i_1,j_1,k_1)}^2 + \varepsilon_{\Delta\Phi(i_2,j_2,k_2)}^2}{(\beta_{(i_2,j_2,k_2)} - \beta_{(i_1,j_1,k_1)})^2}} = \sqrt{\frac{\mu_{(i_1,j_1,k_1)}^2 + \mu_{(i_2,j_2,k_2)}^2}{(\beta_{(i_2,j_2,k_2)} - \beta_{(i_1,j_1,k_1)})^2}} \cdot \varepsilon_{\Delta\Phi} \tag{23.18}$$

Table 23.2 Estimation accuracy of ionospheric delay for some EWL/WL combinations

(i_1, j_1, k_1)	(i_2, j_2, k_2)	$\beta_{(i_1, j_1, k_1)}$	$\beta_{(i_2, j_2, k_2)}$	$\sigma(\Delta\delta\tilde{I}_1)$ (m)
(0, -1, 1)	(1, -6, 5)	-1.59150	-1.43174	2.01052
(0, -1, 1)	(1, 2, -3)	-1.59150	-0.96979	0.32197
(0, -1, 1)	(1, -4, 5)	-1.59150	0.65205	0.38991
(0, -1, 1)	(1, -4, 3)	-1.59150	-1.39529	0.78661
(1, 0, -1)	(1, -4, 5)	-1.23064	0.65205	0.45878
(3, 19, -22)	(4, 15, -19)	-2.68940	0.10036	0.31634

The estimation accuracy of ionospheric delay is only related to the observation noise and combined coefficients. With assuming the standard deviation of DD carrier phase noise $\varepsilon_{\Delta\Phi} = 5$ mm. Table 23.2 lists the estimation accuracy of several combinations:

Although the estimation accuracy of ionospheric delay can be improved by choosing the optimal two groups of carrier phase observations, the minimum value of standard deviation is still about 0.32 m, which is a large value compare with the wavelength of the carrier phase. So this paper introduces Hatch filter to smooth ionospheric delay estimate with variation between two epochs [7]:

$$\Delta\delta\tilde{I}_1(t) = \omega \cdot \Delta\delta\hat{I}_1(t) + (1 - \omega) \cdot \left(\Delta\delta\tilde{I}_1(t - 1) + \Delta\delta\tilde{I}_1(t, t - 1) \right) \quad (23.19)$$

Among them, the weighting factor ω is usually setting as $1/n$, where n is the epoch numbers which are needed to smooth. The variation of ionospheric delay $\Delta\delta\tilde{I}_1(t, t - 1)$ can be calculated by frequency f_1 and f_3 phase measurements between two closed epochs:

$$\Delta\delta\tilde{I}_1(t, t - 1) = \frac{\Delta\Phi_1(t) - \Delta\Phi_3(t) - \Delta\Phi_1(t - 1) + \Delta\Phi_3(t - 1) + \Delta\varepsilon_{\Phi_1}(t, t - 1) - \Delta\varepsilon_{\Phi_1}(t, t - 1)}{q_3 - 1} \quad (23.20)$$

Since the observation noise of carrier phase is Normal distribution, the accuracy of ionospheric delay's variation between two closed epochs is about 2 cm. Through the smoothing, the estimation accuracy of the ionospheric delay can up to 3 cm.

23.3.3 Fixing Integer Ambiguity After Eliminating the Ionospheric Delay

As the foregoing statement, the ionospheric delay restricts traditional TCAR algorithm for ML/NL ambiguity resolution. Above section proposed the method for

ionospheric delay estimating. After eliminating the influence of ionospheric delay, the float ambiguity value can be expressed as follows:

$$\Delta N_{(i,j,k)} = \frac{\Delta P_{(l,m,n)} - \Delta \Phi_{(i,j,k)} - \frac{\beta_{(l,m,n)} + \beta_{(i,j,k)}}{\lambda_{(i,j,k)}} (\Delta \delta I_1 - \Delta \tilde{\delta I}_1)}{\frac{\varepsilon_{\Delta P_{(l,m,n)}} - \varepsilon_{\Delta \Phi_{(i,j,k)}}}{\lambda_{(i,j,k)}}} \quad (23.21)$$

Then, the noise variance for ambiguity resolution is only related with the combined coefficients and carrier phase and pseudo-range observations noise, not associated with the ionospheric and tropospheric delay which evidently mitigates or even eliminates the impact of baseline length for ambiguity resolution. So this method can be used for long baseline ambiguity resolution.

23.3.4 The Basic Integer Ambiguity Resolution

In real-time precision relative positioning applications, in order to obtain high-precision observations, we care more about basic integer ambiguity. If three groups of combined integer ambiguities are known, then the combined integer ambiguities and the basic integer ambiguities can be expressed in the form of a linear equation:

$$\underbrace{\begin{bmatrix} i_1 & j_1 & k_1 \\ i_2 & j_2 & k_2 \\ i_3 & j_3 & k_3 \end{bmatrix}}_A \cdot \underbrace{\begin{pmatrix} \Delta N_1 \\ \Delta N_2 \\ \Delta N_3 \end{pmatrix}}_X = \underbrace{\begin{pmatrix} \Delta N_{(i_1,j_1,k_1)} \\ \Delta N_{(i_2,j_2,k_3)} \\ \Delta N_{(i_3,j_3,k_3)} \end{pmatrix}}_b \quad (23.22)$$

In order to get the basic integer ambiguities (ΔN_1 , ΔN_2 , ΔN_3), the combined coefficient matrix A is required invertible, which is equivalent to that the three groups of combined coefficients are linear independent. Using the geometry-free TCAR method, we can easily fix more than two groups of EWL/WL combined ambiguities with combined coefficients satisfy $i + j + k = 0$. So we only need to determine another combined integer ambiguity whose combined coefficients satisfy $i + j + k \neq 0$, then the basic integer ambiguities can be solved by the nonlinear equation $X = A^{-1} \cdot b$. In order to avoid the calculate NL combination, whose wavelength is short and vulnerable affected by observation noise, this paper use the optimal combined pseudo-range coefficient calculation method mentioned above to show several groups of carrier phase combination whose combined coefficients satisfy $i + j + k \neq 0$, in which the standard deviation of ionospheric delay for frequency f_1 after eliminating is setting as 3 cm (Table 23.3).

As eliminating the ionospheric delay, the float ambiguity is unbiased, so the success probability can be improved by smoothing multi-epochs filter.

Table 23.3 Single epoch success probability of improved TCAR for Optimal combination coefficients after eliminating ionospheric delay

(i, j, k)	$\lambda_{(i,j,k)}$ (m)	Coefficients (l, m, n)	$\beta_{(i,j,k)} + \beta_{(l,m,n)}$	σ_{TN} (cycle)	Success probability (%)
(-2, 2, 1)	0.5348	(0.35881, 0.31551, 0.32568)	6.43970	0.4904	69.20
(4, -3, -2)	3.4887	(0.17100, 0.44690, 0.38210)	-41.03020	0.3789	81.31
(6, 0, -7)	0.6157	(0.30576, 0.35262, 0.34162)	-6.96896	0.4910	69.15

23.4 Experiment and Analysis

As the foregoing statement, the accuracy for ambiguity resolution of improved TCAR is not associated with the ionospheric and tropospheric delay, which means with less influence by baseline length for ambiguity resolution. So we chose experimental data which were collected on October 7th, 2014 in Changsha by two BeiDou triple-frequency monitoring receivers with 10 km distances, and antennas were placed on the roof, without surrounded by trees and tall buildings, to mitigate the impact of multipath, set the PRN 6 as the reference satellite.

Figures 23.1 and 23.2 show the floating and integer ambiguity resolutions of PRN 4 for EWL $\Delta N_{(0,-1,1)}$ and WL $\Delta N_{(1,2,-3)}$, since the baseline is short, both the traditional method and the improved method have 100 % success probability for single epoch ambiguity resolution, but the total noise level is decreased by about 15 % with improved method.

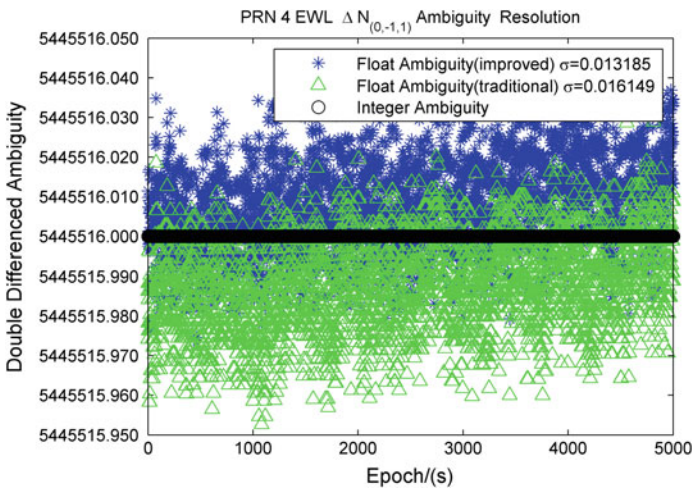


Fig. 23.1 EWL $\Delta N_{(0,-1,1)}$ float and integer ambiguity

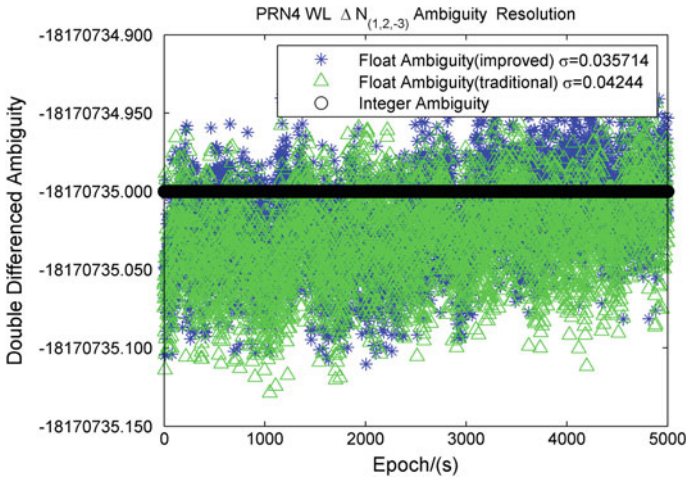


Fig. 23.2 WL $\Delta N_{(1,2,-3)}$ float and integer ambiguity

When the carrier phase combined coefficients are satisfied the conditions $i + j + k \neq 0$, the total noise level is usually higher than the needs for single epoch ambiguity resolution. This paper sets the smooth epochs are 30 s, the result of ML $\Delta N_{(2,-2,1)}$ ambiguity resolution is shown as below:

Figure 23.3 shows that by using optimal combined coefficients, eliminating the impact of ionospheric delay, and smoothing floating ambiguity, the improved TCAR can successfully fix the integer ambiguity after 60 s initial time.

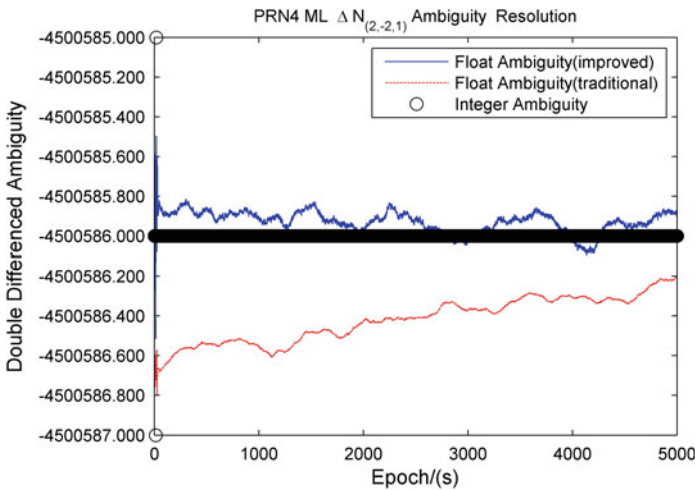


Fig. 23.3 ML $\Delta N_{(2,-2,1)}$ float and integer ambiguity

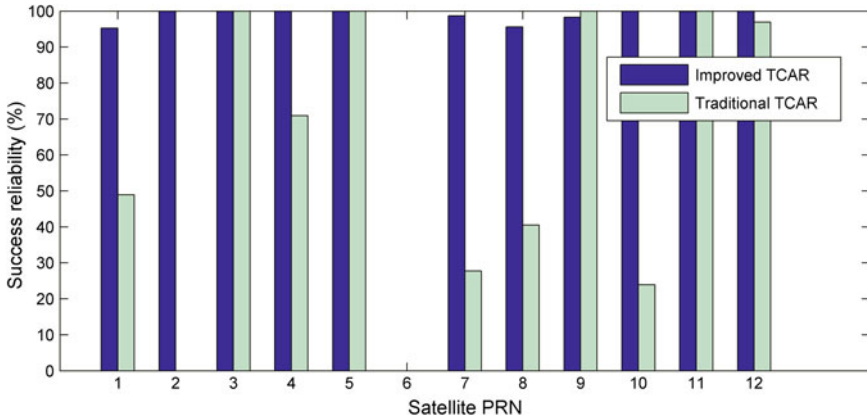


Fig. 23.4 Success reliability for $\Delta N_{(2,-2,1)}$ ambiguity resolution

Figure 23.4 shows the traditional TCAR and the improved TCAR success reliability for $\Delta N_{(2,-2,1)}$ ambiguity resolutions. The success reliability is calculated by statistical method, which is defined the percentage of correctly solved epoch number to the total epoch number after the 60 s initialization and 30 epochs smoothing. It can be seen that the improved TCAR has higher than 95 % success reliability for $\Delta N_{(2,-2,1)}$ combined ambiguity resolution for all satellites, equivalent to that the success reliability for basic ambiguity resolution is over 95 %. Compared with the traditional TCAR, the improved TCAR can improve success reliability over 30 % for partially visible satellites.

23.5 Concluding Remarks

This paper presents an optimal combination of carrier phase and Pseudo-range to improve TCAR algorithm, which base on the minimum of the total noise level, gives the calculation method of the optimal coefficients of pseudo-range, and uses hatch filter to smoothing the estimation of ionospheric delay to improve the accuracy for ambiguity resolution. Finally, BeiDou triple-frequency data in short-baseline were used to verify this algorithm. The results show that: this method can successfully fix the DD EWL/WL integer ambiguities from single epoch, and reduce the total noise level more than 15 %; by real-time estimating and mitigating even eliminating the impact of the ionospheric delay, the success reliability of basic ambiguities can achieve 95 %.

References

1. Forssell B, Martin-Neira M, Harris R (1997) Carrier phase ambiguity resolution in GNSS-2. In: Proceedings of ION GPS-97, Kansas City, pp 1727–1736, 16–19 Sept 1997
2. Vollath U, Birnbach S, Landau H (1998) Analysis of three carrier ambiguity resolution (TCAR) technique for precise relative positioning in GNSS-2. In: Proceedings of ION GPS-98, pp 417–426, 15–18 Sept 1998
3. Jung J, Enge P, Pervan B (2000) Optimization of cascade integer resolution with three civil GPS frequencies. In: Proceedings of ION GPS-2000, Salt Lake City, pp 2191–2200
4. Li B, Feng Y, Shen F (2010) Three carrier ambiguity resolution: distance-independent performance demonstrated using semi-generated triple frequency GPS signals. *GPS Solutions* 14:177–184
5. Feng Y, Rizos C, Higgins M (2007) Multiple carrier ambiguity resolution and performance benefits for RTK and PPP positioning services in regional areas. In: Proceedings of ION GNSS 20th international technical meeting of the satellite division, TX, USA, 2007, pp 668–678
6. Feng Y (2008) GNSS three carrier ambiguity resolution using ionosphere-reduced virtual signals. *J Geodesy* 82(12):847–862
7. Deng J, Pan S, Hong Z (2014) A resolution method for ionospheric delay with optimal combination of three-frequency data. *Geometrics and Information Science of Wuhan University, Wuhan*, pp 600–604

Chapter 24

A Hybrid Navigation Constellation Inter-satellite Link Assignment Algorithm for the Integrated Optimization of the Inter-satellite Observing and Communication Performance

Bo Xu, Donghui Wang, Wenxiang Liu and Guangfu Sun

Abstract The optimization of the inter-satellite link assignment is an important technical issue, especially, for the case that the number of inter-satellite spot-beam antennas is less than assignable links. Based on Minimum Spanning Tree (MST) theory in graph theory, this paper proposes a link assignment algorithm for the hybrid navigation constellation using GEO, IGSO and MEO satellites. Using GEO satellites as the core transfer node, this algorithm assigns inter-satellite links with lower communication costs firstly, and then assigns links at least once for all visible satellites by applying the iterative assignment. As a consequence, the integrated optimization of the inter-satellite observing and communication performance can be achieved. The simulated results show that, compared to the mesh-link assignment algorithm used for the Iridium system, the proposed link assignment algorithm can be able to improve the precision of orbit determination and reduce whole-network average communication delay.

Keywords Hybrid navigation constellation · Link assignment · Inter-satellite observing · Inter-satellite communication

24.1 Introduction

Hybrid constellation has been applied widely in major GNSS (Global Navigation Satellite Systems) to improve the ground coverage performance of navigation signals. Currently, the main navigation satellite types include GEO (Geostationary

B. Xu (✉) · D. Wang · W. Liu · G. Sun
School of Electronic Science and Engineering, National University of Defense Technology,
Changsha 410073, China
e-mail: xubo973109@sina.com

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015
Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_24

283

Orbit), IGSO (Inclined Geostationary Orbit) and MEO (Medium Earth Orbit). The basic constellation of the GPS satellite navigation system is the MEO constellation, and its enhanced system WAAS is equipped with 3 GEO satellites to realize the enhanced coverage of a fixed region [1]. The similar enhancement methods are used for the European GNSS and the Russia GNSS. In order to improve the coverage performance in China and the surrounding regions, the BDS (BeiDou Navigation Satellite System) employs the constellation form with 5GEO + (20–30)NGEO (IGSO and MEO) [2]. The coverage characteristics of the BDS hybrid constellation, and other three major GNSS (GPS, GALILEO and GLONASS) in Asia have been comparatively analyzed from the perspective of satellite visibility and DOP (Dilution of Precision) in [3]. The navigation service performance of the BDS has been detailed analyzed in [4].

The main function of the navigation inter-satellite link comprises of the inter-satellite observing and communication. The autonomous navigation of the whole constellation can be realized through the navigation inter-satellite link. The UHF-band wide beam antenna has already been used to establish inter-satellite links for the GPS Block IIR and Block IIF satellites [5]. The autonomous navigation could improve the wartime viability of the navigation constellation greatly [6]. Establishing spot beam inter-satellite links is a development trend of satellite navigation systems. At the GPSIII stage, inter-satellite links may use the Ka (23 GHz) or V (60 GHz) band, and the use of spot beams for inter-satellite link antennas can improve communication rate and anti-jamming ability [7]. The GALILEO system has also conducted extensive research on inter-satellite links. In this research, the inter-satellite observing and communication performance was analyzed [8–10].

Since the number of inter-satellite antennas spot beams is less than that of visible satellites, there is the issue of optimization in the inter-satellite link establishment. It is necessary to select suitable satellites from all visible satellites to establish links by utilizing the limited beam resources rationally. The link assignment method is the primary issue in the building of inter-satellite link systems, for it determines inter-satellite observing and communication performance directly. Some scholars have studied inter-satellite networking technologies for navigation constellations. William and Marc [11] studied the inter-satellite visibility and the required data volume for the inter-satellite communication. Shi et al. [12] proposed the analytical algorithm for the MEO constellation inter-satellite visibility, providing a basis for the research of the link assignment algorithm. Using the idea of limited state topology, [13] simplifies network designing problems of topological changes resulting from satellite motion, and proposed a link dispatching and assignment algorithm based on the link load equalization. Stephen et al. [14] discussed the mesh-link assignment method applied in the Iridium system. Shi et al. [15] studied the inter-satellite link assignment algorithm of the MEO navigation constellation, and conducted a comparative analysis with the mesh-link assignment method of the Iridium system. Han et al. [1] studied the routing algorithm for inter-satellite links of navigation constellations. However, the above mentioned algorithms were aimed at communication satellite

systems or MEO constellations. Thus, they cannot be completely applied in hybrid navigation constellations. Currently, no published research on the link assignment of hybrid navigation constellations is available.

24.2 Link Assignment Issue and Its Mathematical Model

24.2.1 Link Assignment Issue

The essence of the inter-satellite link assignment issue is to select suitable satellites to establish inter-satellite signal connection links. In order to ensure the quality of communication links, the established links are usually chosen from permanently visible satellites. Taking the Iridium communication system of the U.S. as an example, in which 4 permanent visible links are established for each satellite using the mesh-link assignment method. Figure 24.1 is a schematic diagram of 4 permanent links established in the Walker 24/3/2 constellation using the mesh-link assignment method. The constellation has 3 orbits, with 8 satellites each. For each satellite links are established with two adjacent satellites on the same orbit and a permanently visible satellite on each of the other two orbits.

For the satellite navigation inter-satellite link assignment algorithm, both the demand of the inter-satellite communication and inter-satellite observing should be considered. The navigation inter-satellite links operate with the TDMA (Time Division Multiple Access). Figure 24.2 is a schematic diagram of the TDMA structure. The timeline is divided into several cycles with equal intervals. In each cycle, the constellations run at the time sequence programmed in advance. Inter-satellite observing is conducted at the initial stage of each cycle, and may continue

Fig. 24.1 Schematic diagram of mesh-link establishment

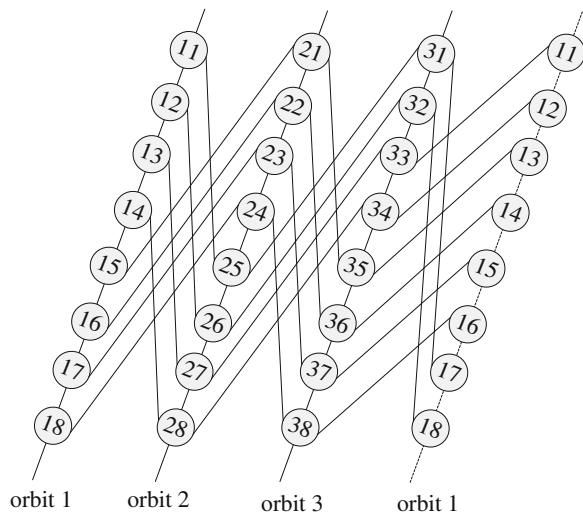
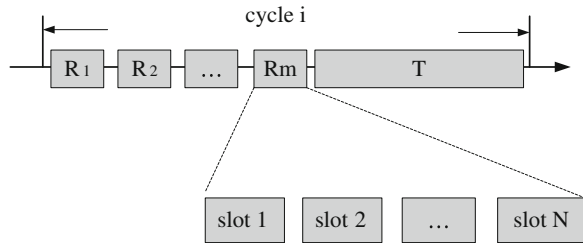


Fig. 24.2 TDMA structure of navigation inter-satellite link



for n sub-frames. Each observing sub-frame R may be divided into some slots based on the number of satellites, and each satellite occupies a slot and transmits the observing signal. The linked satellites receive the signal to complete inter-satellite observing. The inter-satellite communication can be conducted within the remaining time T of the inter-satellite link cycle.

In the inter-satellite observing, the attention should be paid to the number of established links of a satellite. The reason for this is that the more inter-satellite observed information is obtained, the higher orbit determination precision will be. Therefore, a variable link should be assigned in observing sub-frame R to maximize inter-satellite observed information. However, in the inter-satellite communication, the attention should be paid to the link quality, such as link transmission delays, bit error rates, etc. Thus the link with the lowest communication cost should be assigned in communication sub-frame T . The mesh-link assignment method is not applicable to satellite navigation inter-satellite links, since its poor performance on increasing the number of inter-satellite observing links. Besides, the link transmission delay of mesh-link assignment method is also not optimized. It is necessary to establish a novel topology structure with the best inter-satellite observing and communication performance.

24.2.2 Mathematical Model for Link Assignment

The following mathematical model is established for the link assignment based on the graph theory:

The whole constellation can be regarded as a weighted directed graph G . Its 2-tuple is defined as:

$$G = (V, E) \tag{24.1}$$

Where, V is a vertex set that meets the following conditions:

$$V = \{v_i | 1 \leq i \leq N, v_i \in S\} \tag{24.2}$$

where, S is the set of all satellites of the whole constellation, and the total number of satellites is N .

E is the edge set of Graph G, meeting the following conditions:

$$E = \{ \langle v_i, v_j \rangle \mid i \neq j, v_i, v_j \in V \} \quad (24.3)$$

The elements in E indicate the links established between two satellites.

Graph G uses an adjacency matrix to conduct storage, meeting the following relationship:

$$A[i, j] = \begin{cases} w_i^j, & \langle v_i, v_j \rangle \in E \\ 0, & \langle v_i, v_j \rangle \notin E \end{cases} \quad (24.4)$$

where, w_{ij} is the weight of the link and may be indicated by the link transmission cost; matrix A is a symmetric matrix. The meaning of A is that the link cost from satellite V_i to satellite V_j is equal to that from satellite V_j to satellite V_i . Thus, the diagonal elements of matrix A are 0.

The criterion for judging if $\langle v_i, v_j \rangle$ belongs to E is whether two satellites are visible. In other words, the view vectors of two satellites are not obstructed by the earth, and both satellites are within the beam coverage of the satellite's antennas.

According to the above mathematical model, the link assignment issue may be expressed as the link assignment matrix L and the cost matrix C within a certain period of time, based on the constellation state matrix A entered at a certain moment. The link assignment matrix L may be expressed as:

$$L = \{L_i \mid 0 \leq i \leq n\} \quad (24.5)$$

where, n is the number of link assignment matrixes to be generated and determined by system demand. Usually, in an inter-satellite link work cycle, the link assignment matrix sequence of all sub-frames in this cycle can be calculated, given the constellation state at the starting moment.

For a certain link assignment matrix L_t , the following conditions are met:

$$L_t = [l_{ij}], l_{ij} = \begin{cases} 1, & \text{link} \\ 0, & \text{unlink} \end{cases} \quad (24.6)$$

where, L is a symmetric matrix. Since the number of inter-satellite link antenna beams is limited, a beam can only be linked with one satellite. Therefore, the number of elements equal to 1 on any row or column of the matrix L_t should not exceed the number of satellite antenna beams.

The cost matrix C corresponds to the link assignment matrix L on a one-to-one basis. C may be expressed as follows:

$$C_i[c_{ij}], c_{ij} = \begin{cases} w_i^j, & l_{ij} = 1 \\ 0, & l_{ij} = 0 \end{cases} \quad (24.7)$$

The whole-network average link transmission cost may be used to measure the inter-satellite communication performance. The average transmission cost from a

certain satellite V_k to all other satellites in the whole network may be expressed as follows:

$$C_{avr} = \frac{1}{N-1} \sum_{i=1, i \neq k}^N r_{ki} \quad (24.8)$$

where, r_{ki} is the total cost of the link transmission path from V_k to V_i , namely the sum of all inter-node costs on the path.

24.3 Hybrid Constellation Link Assignment Algorithm for the Integrated Optimization of the Inter-satellite Observing and Communication Performance

The aim of this paper is to develop the link assignment algorithm for the hybrid constellation with three types of satellites, namely GEO, IGSO and MEO. In the satellite-ground operation mode, on the one hand, the ephemeris parameters have to be uploaded into the satellites by the ground station, on the other hand, the inter-satellite observing data has to be downloaded to the ground control system. The data transmission between the ground station and the whole constellation is the main communication demand of inter-satellite links, and the satellite-ground connection transfer node must be stable and reliable. Among the three types of satellites, GEO satellites remain immobile relative to the earth and may realize fixed observation with the ground control system. Undoubtedly, since this characteristic, the GEO satellite is the best choice of the connecting node between the ground and the whole constellation. Therefore, for a hybrid navigation constellation, the GEO satellite should be the satellite-ground connecting node and the core of link assignment of the whole satellite network.

The optimization objective of the hybrid constellation link assignment algorithm is to minimize the average communication cost of all satellites from the ground control system to the whole constellation with a GEO satellite being the core node, given that the inter-satellite observed information is maximized.

In the navigation constellation, provided that the number of connected nodes is fixed, the lower the assigned link cost is, the lower the whole-network average transmission cost will be. In order to maximize the number of observing links, as many as different links should be established. Thus, it is necessary to record the link assignment state and to assign unassigned low-cost links first through the iterative process. In order to realize the integrated optimization of the inter-satellite observing and communication performance, in some necessary cases, the high-cost links should also be assigned appropriately to meet the observing requirement, with the cost of partial communication performance loss.

The input of the algorithm is the inter-satellite visibility and link cost at a certain moment, and its output is a series of link assignment matrixes $\{L1, L2, \dots, Ln\}$. The specific process of the algorithm is as follows:

- Step 1: Calculate the state matrix A of the whole network based on the inter-satellite visibility and the link cost at a certain moment;
- Step 2: Establish the MST: Based on A , the MST L_{min} of the whole constellation can be established with all GEO satellites as the core. Consequently, a strongly connected graph can be formed for the whole constellation. According to the MST concept in the graph theory, the MST is the connected network with the smallest number of edges and the lowest whole-network cost. In this paper, the MST generation algorithm is improved in adaptability to keep GEO satellites as the core nodes of the MST. All GEO satellites are directly connected with each other. Although the resulting tree cost is not the lowest, it does ensure that GEO satellites are located at the core of the network;
- Step 3: Assign links for each GEO satellite on the basis of MST L_{min} as follows:
- (1) Firstly, assign an IGSO satellite link for each GEO satellite and make sure that different GEO satellites are linked to different IGSO satellites. If link $\langle v_i, v_j \rangle$ is assigned here, the elements at Row i and Column j , and the Row j and Column i of the assignment matrix are set to 1 according to Eq. (24.9) below. Meanwhile, the corresponding elements of the satellite state matrix A are set to infinity, and this link will not be assigned any more.

$$L_{min}[i, j] = 1, L_{min}[j, i] = 1, A[i, j] = \infty, A[j, i] = \infty \quad (24.9)$$

- (2) If the number of elements equal to 1 on the row or column of a GEO satellite in the link assignment matrix L_{min} does not exceed the number of satellite antennas k , links from GEO satellites to MEO satellites will be further assigned. The link with the lowest cost should be chosen from all visible satellite links for the assignment, and a GEO satellite should not be linked to two or more MEO satellites on the same orbit.
- (3) Update link assignment matrixes and constellation state matrixes, and the output link assignment matrix L_{min}^{GEO} according to Eq. (24.9). After above 3 steps, the link assignment for GEO satellites has been completed. Figure 24.3 is a schematic diagram of the link establishment thereof (assuming that the number of inter-satellite antenna beams is 4):

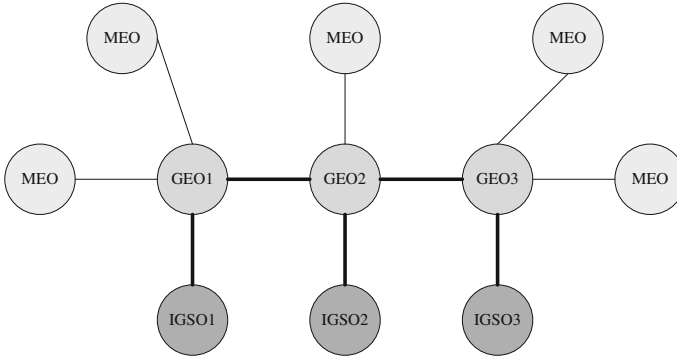


Fig. 24.3 Schematic diagram of GEO satellite link assignment

Step 4: Assign links to IGSO and MEO satellites. For any satellite V_i , the link assignment method is as follows:

- (1) Arrange all links connected to satellite V_i in the ascending order of cost;
- (2) Judge if satellite V_i needs link assignments on the condition that the number of elements equal to 1 on Row i of L_{\min}^{GEO} does not exceed that of satellite antenna beams k . If satellite V_i needs link assignments, all links of satellite V_i will be searched in the ascending order of cost. If any link $\langle v_i, v_j \rangle$ meets the condition of Eq. (24.10) below, this link will be assigned to satellite V_i .

$$\begin{aligned} A[i, j] &\neq 0 \\ L_{\min}^{GEO}[i, j] &= 0 \end{aligned} \tag{24.10}$$

- (3) Update link state matrixes and link assignment matrixes according to Eq. (24.9), and output the first link assignment matrix L_1 .

Step 5: Iterate the link establishment: On the basis of the state matrix A outputted by Step 4, repeat Steps 3, 4 and 5 to generate link assignment matrixes $\{L_2, \dots, L_n\}$. If the inter-satellite link of the whole network is already maximized for a certain link assignment matrix L_i , it is only necessary to consider the requirement of minimum communication cost in the subsequent generation of link establishment matrixes $\{L_{i+1}, \dots, L_n\}$. The state matrix A should be reset as the one outputted in Step 1, and any assigned link may be reassigned.

In sum, the process of the algorithm is as Fig. 24.4.

It can be seen from the above algorithm process that a series of link assignment matrixes $\{L_1, L_2, \dots, L_n\}$ can be generated based on the constellation state matrix A at a certain moment. The number of link assignment matrixes can be configured.

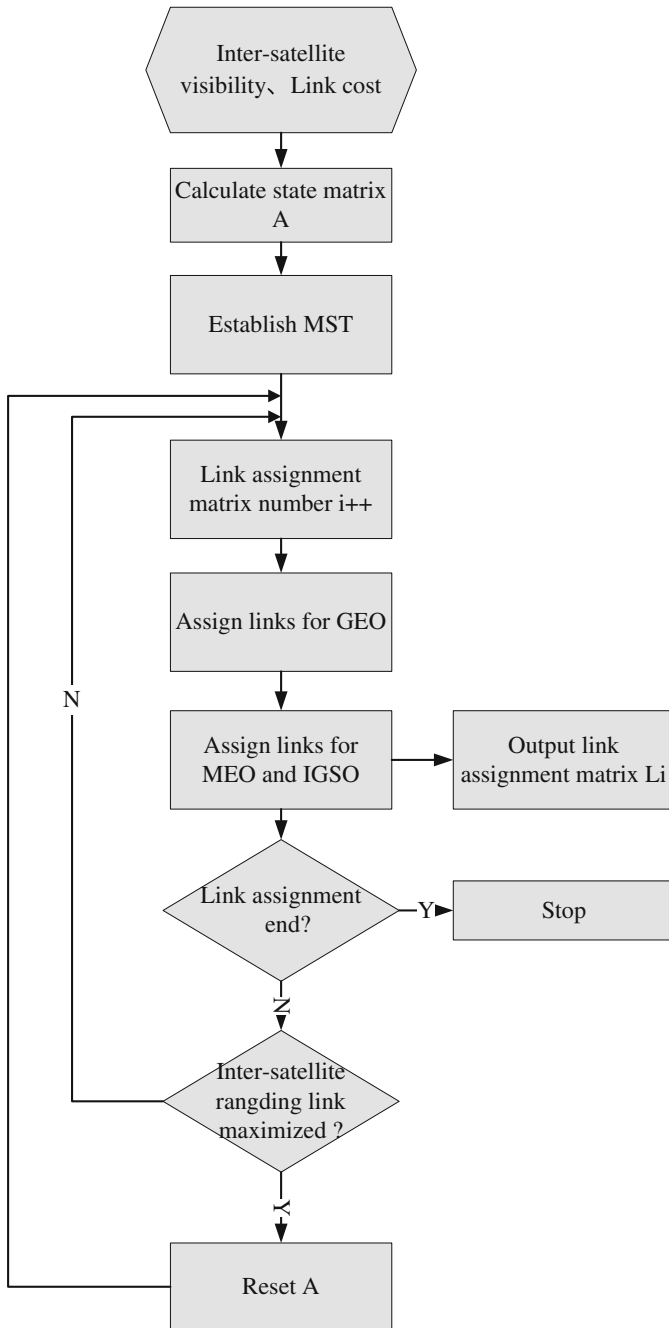


Fig. 24.4 Flowchart of the hybrid constellation link assignment algorithm

To maximize the inter-satellite observing, the unassigned links are always assigned firstly during the link assignment iteration in step 5 of the algorithm. Since lower-cost links are assigned firstly, the whole-network transmission cost can be minimized.

24.4 Simulation Analysis

In this section, a 3GEO + 3IGSO + 24MEO hybrid constellation is taken as an example for the simulation analysis of the proposed link assignment algorithm. The longitudes of the GEO satellites are 60° , 110° and 140° , respectively. The ascending node of the IGSO satellites is 118° . The MEO satellites are in Walker 24/3/2 distribution, with an orbit inclination of 55° and an orbit height of 26,559.8 km [1]. The STK software of AGI Inc. is used to simulate the constellation scenario, with the starting time of simulation being July 1, 2008 12:00. It is assumed that each satellite has 4 inter-satellite link spot beam antennas and at most 4 links are established. It is assumed that the inter-satellite link cycle is 15 min. The N bodies gravity model is DE405, the Earth gravity model is EGM96 (12*12), the Solar Radiation Pressure model is classic model with 20 % model error.

24.4.1 Simulation Analysis of the Orbit Determination Performance

It is assumed that the inter-satellite observed arc is 1 day, and the integration step of the orbit is 5 min. Adopting 7 IGS monitoring stations distributed in local area, namely bjfs, chan, kunm, lhaz, shao, tnml and urum. In the perturbation force model, N-body gravity, earth gravity and solar radiation pressure are considered. The inter-satellite and satellite-to-station observing error model is subject to normal distribution $N(0.6, 0.3)$.

The visible satellite number of MEO11 is as Fig. 24.5.

The visible satellite number of MEO11 reaches 14–17.

The results of orbit determination based on combined inter-satellite and satellite-to-station ranges using the proposed algorithm and the mesh-link establishment method are listed in Table 24.1, respectively.

Taking satellite MEO11 as an example, Fig. 24.6 shows the URE of the orbit determination under two algorithms:

It can be seen from the simulation result that the orbit determination precision of the proposed algorithm is much better than that of the mesh-link assignment method. The reason for this is that the algorithm proposed in this paper ensures that all visible satellites are assigned to links at least once in a cycle, thereby maximizing the number

Fig. 24.5 The Visible satellite number of MEO11

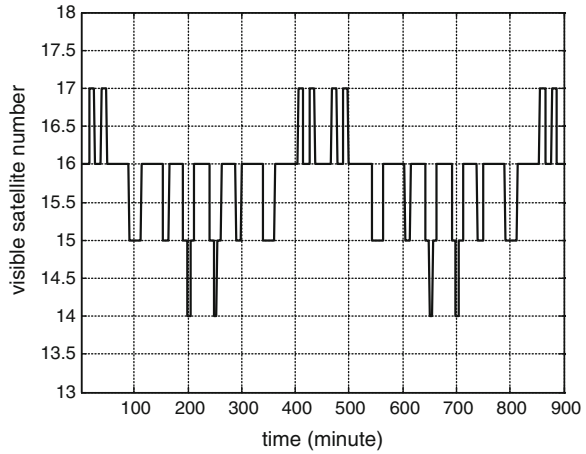
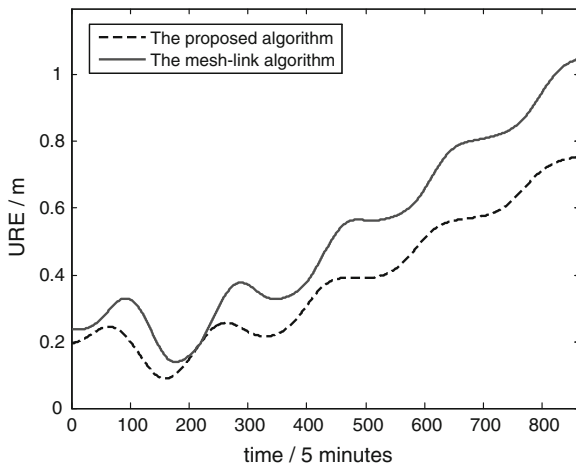


Table 24.1 Comparison of orbit determination performance between the proposed algorithm and the mesh-link assignment method

	The proposed algorithm	The mesh-link algorithm
Average 3-dimensional residual error in orbit determination, RMS/m	3.11	4.23
Forecast 2-h URE/m	0.18	0.25

Fig. 24.6 The URE for orbit determination of satellite MEO11



of inter-satellite observing links. However, for the mesh-link assignment method, each satellite is assigned to 4 links regularly in a cycle, so that the number of observing links is limited.

24.4.2 Simulation Analysis of the Inter-satellite Data Transmission Performance

As the information starting node, the GEO3 satellite can be chosen to send data packages to all satellites in the whole network. The average link transmission delay from GEO3 to all satellites in the whole network could be calculated. The simulation duration was 15 h (a constellation operation cycle), for 60 inter-satellite link cycles in total. Based on the calculation of 1,000,000 simulations, the link transmission performance of the algorithm proposed in this paper is better than that of the mesh-link assignment method for over 97.0 % of the time. The average statistics of link transmission delay of the three algorithms are given in the Table 24.2. (Averages of all simulation moments of all satellites):

It can be seen from the above statistics that the link transmission delay of the algorithm proposed in this paper is lower than that of the mesh-link assignment method by 11.6 %.

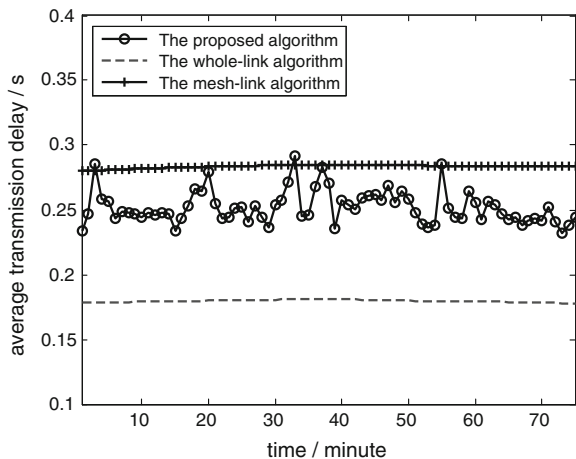
Figure 24.7 below shows the simulation results of link transmission delay of the first 5 inter-satellite link cycles.

It can be seen from the above result that the overall link transmission performance of the algorithm proposed in this paper is better than that of the mesh-link establishment method. However, since consideration is also given to the inter-satellite observing performance in the proposed algorithm, its link transmission

Table 24.2 Comparison of the three algorithms in terms of the average link transmission delay

Link assignment method	Average link transmission performance/s
The mesh-link algorithm	0.284
The proposed algorithm	0.251
The whole-link algorithm	0.179

Fig. 24.7 Comparison of the three algorithms in terms of the link transmission delay performance



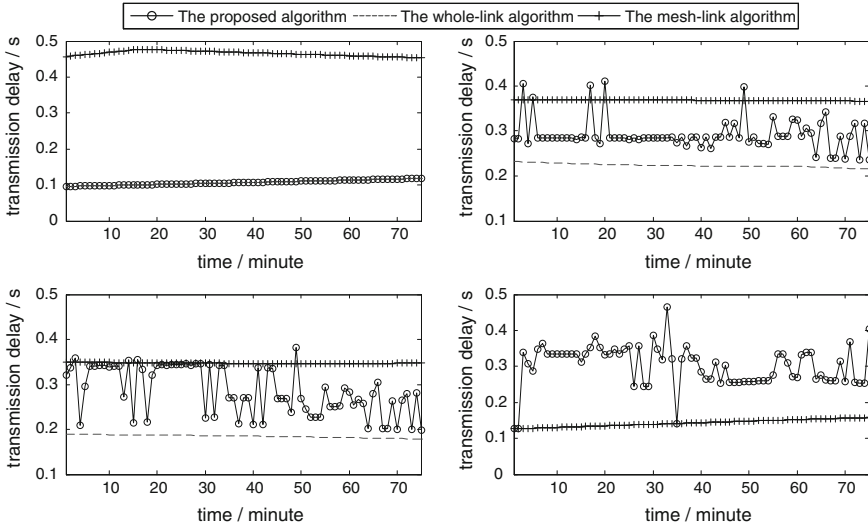


Fig. 24.8 Comparison of the three algorithms in terms of the average transmission delay performance of point-to-point links of hybrid constellations

performance is slightly lower than that of the mesh-link establishment method at some particular moments.

The simulation analyzes the point-to-point link transmission delay results from GEO3 to IGSO1, MEO11, MEO21 and MEO31, covering the high-high links between GEO and IGSO, and high-low links between GEO and the three MEO orbital planes. Since the constellation is symmetric, this simulation conditions can represent the state of the whole constellation. The simulation results of the first 5 cycles are shown in Fig. 24.8.

It can be seen from Fig. 24.6 that the point-to-point link transmission performance of the proposed algorithm is better than that of the mesh-link assignment method and closer to the theoretical value. In the proposed algorithm, GEO3 is linked directly to IGSO1, and then the point-to-point link transmission delay from GEO3 to IGSO1 is the same as that of the whole-link assignment method. In the mesh-link, GEO3 is linked directly to MEO31, so its point-to-point transmission delay is equal to the delay in case of whole-link. In other cases, the performance of the proposed algorithm is better than that of the mesh-link assignment method.

24.5 Conclusion

This paper proposes a link assignment algorithm for the integrated optimization of the inter-satellite observing and communication performance for hybrid navigation constellation. Based on the MST theory in the graph theory, the algorithm reduces whole-network communication cost by assigning low-cost links firstly. Then, all

visible satellites are assigned links at least once by the iterative assignment in a cycle. Thus, the inter-satellite observed information can be maximized. Compared to the mesh-link assignment algorithm used in the Iridium system, the optimized algorithm proposed in this paper can improve the orbit determination precision based on combined inter-satellite and satellite-to-station ranges and reduce the whole-network communication cost effectively.

Although the link assignment algorithm proposed in this paper is developed for hybrid constellations, it may be extended to ordinary network link assignment issues with core nodes. The proposed algorithm is feasible in engineering, and may provide a reference for the inter-satellite link networking of the BDS satellite navigation system.

References

1. Han SH, Gui QM, Li JW (2013) Establishment criteria, routing algorithms and probability of use of inter-satellite links in mixed navigation constellations. *Adv Space Res* 51:2084–2092
2. China Satellite Navigation Office (2012) Development of BeiDou navigation satellite system. In: Proceedings of the munich satellite navigation summit, Munich, Germany, 17–24 March 2012
3. Chiang KW, Huang YS, Tsai ML, Chen KH (2010) The perspective from Asia concerning the impact of COMPASS/BEIDOU-2 on future GNSS. *Surv Rev* 42(315):3–19
4. Chen HC, Huang YS, Chiang KW, Yang M, Rau JR (2009) The performance comparison between GPS and BeiDou-2/COMPASS: a perspective from Asia. *J Chin Inst Eng* 32(5):679–689
5. Fisher SC, Ghassemi K (2002) GPS IIF—the next generation. In: Proceedings of the IEEE, vol 87, issue 1, pp 24–47
6. Rajan JA (2002) Highlights of GPS II-R autonomous navigation. In: ION 58th annual meeting of the institute of navigation and CIGTF 21st guidance test symposium, Albuquerque, NM, pp 354–363
7. Maine K, Anderson P, Bayuk F (2004) Communication architecture for GPSIII. In: IEEE aerospace conference, Aerospace Corporation, Los Angeles, CA, pp 124–129
8. Francisco AF (2011) Inter-satellite ranging and inter-satellite communication links for enhancing GNSS satellite broadcast navigation data. *Adv Space Res* 47:786–801
9. Sanchez M, Pulido JA (2008) The ESA “GNSS+” project inter-satellite ranging and communication links in the frame of the GNSS infrastructure evolutions. In: The 21st international technical meeting of the satellite division of the institute of navigation. The Institute of Navigation, Georgia, pp 2538–2546
10. Irma RP, Cristina GS (2011) Inter-satellite links for satellite autonomous integrity monitoring. *Adv Space Res* 47(2):197–212
11. William GC, Marc PL (2009) Analysis of uncoordinated ISL networks. In: Proceedings of the global telecommunications conference, GLOBECOM, vol 1–6, Honolulu, Hawaii, USA
12. Shi LY, Ou G, Gu QT et al (2011) An analytical algorithm for inter-satellite visibility calculation with restriction of satellite antenna’s elevation. *J Natl Univ Defense Technol* 33(4): pp 97–101
13. Hong S, Chang BWK, Chang GL (1998) FSA-Based link assignment and routing in low-earth orbit satellite networks. *IEEE Trans Veh Technol* 47(3):1037–1048
14. Stephen R, Partt RAR, Carl E et al (1999) An operational and performance overview of the Iridium low earth orbit satellite system. In: Communications surveys and tutorials, IEEE
15. Shi LY, Xiang W, Tang XM (2011) A link assignment algorithm applicable to crosslink ranging and data exchange for satellite navigation system. *J Astronaut* 32(9):1971–1977

Chapter 25

Distributed Orbit Determination Based on Increased Measurement Covariance EKF for Global Navigation Satellite System with Inter-satellite Link

Kai Xue, Yuanlan Wen, Ying Liao, Yisheng Song,
Tianxiang Su and Zhi Zhang

Abstract It is a tendency for Global Navigation Satellite System to finish Autonomous Distributed Orbit Determination with in-satellite link measurement. Based on analysis of Whole-Constellation Centralized EKF and Iterative Cascade EKF, a new Increased Measurement Covariance EKF is proposed. Three algorithms are compared in performance. The new algorithm shows its advantages in Autonomous Distributed Orbit Determination.

Keywords Whole-Constellation centralized EKF · Iterative cascade EKF · Increased measurement covariance EKF

25.1 Introduction

Navigation constellation is a large distributed satellite system. With support of GAS and inter-satellite links, the entire system will constitute a complex network link, and each satellite will fulfil inter-satellite measurement, inter-satellite communications and on-board data processing. Although satellites are physically disconnected, they together achieve the same space mission. In recent years, the significant progress has

K. Xue (✉) · Y. Wen · Y. Liao · Z. Zhang
College of Aerospace Science and Engineering, National University of Defense Technology,
Changsha 410073, China
e-mail: xuekai_39153130@163.com

Y. Song
College of Defense Engineering, University of Science and Technology,
Nanjing 210007, China

T. Su
Troop 61892, Shantou, Guangdong 515071, China

been made in distributed satellite system. However, it is still a lack of proficient distributed algorithms of autonomous orbit determination for navigation constellation in engineering. This paper mainly focuses on distributed algorithms of autonomous orbit determination, but not on time synchronization.

25.2 Equations of Measurement and Motion

25.2.1 Equations of Measurement

Given $\vec{r}^i = [x^i \ y^i \ z^i]^T$ and $\vec{v}^i = [v_x^i \ v_y^i \ v_z^i]^T$ are the 3×1 position vector and velocity vector of i -th satellite, respectively, P_K^i is vector of the kinematical parameters, P_D^i is vector of dynamic parameters. State vector $\mathbf{X}^i = [\vec{r}^{iT} \ V^{iT} \ P_K^{iT} \ P_D^{iT}]^T$ to be estimated contains the above parameters. Their reference values are \mathbf{X}^{i*} , and their improved values are $\delta \mathbf{x}^i = \mathbf{X}^i - \mathbf{X}^{i*}$.

After the hardware delay, plasmaspheric delay, relativistic effect, multi-path effect and antenna phase center offset are corrected, two One-way inter-satellite PRs between i -th satellite and j -th satellite need to be transmit to the same measurement epoch (e.g. Ranging frame epoch t), PR equations are as follows:

$$\rho^{i \rightarrow j}(t) = c \cdot [\delta t^j(t) - \delta t^i(t)] + d(\mathbf{X}^j(t), \mathbf{X}^i(t)) + \varepsilon^{i \rightarrow j}(t) \quad (25.1)$$

$$\rho^{j \rightarrow i}(t) = c \cdot [\delta t^i(t) - \delta t^j(t)] + d(\mathbf{X}^j(t), \mathbf{X}^i(t)) + \varepsilon^{j \rightarrow i}(t) \quad (25.2)$$

where, c is the speed of light, $\delta t^i(t)$ and $\delta t^j(t)$ are clock error of i -th satellite and the j -th satellite, respectively, $\varepsilon^{i \rightarrow j}(t)$ and $\varepsilon^{j \rightarrow i}(t)$ are the measurement error, respectively, $d(\mathbf{X}^j(t), \mathbf{X}^i(t)) = \sqrt{(x^i - x^j)^2 + (y^i - y^j)^2 + (z^i - z^j)^2}$ is the geometric distance between the two satellites.

With Eqs. (25.1) and (25.2) added, Equations of distance measurement that contains only orbit parameter is derived.

$$\rho^{ij}(t) = [\rho^{j \rightarrow i}(t) + \rho^{i \rightarrow j}(t)]/2 = d(\mathbf{X}^j(t), \mathbf{X}^i(t)) + \varepsilon^{ij}(t) \quad (25.3)$$

where, $\varepsilon^{ij}(t_k) = [\varepsilon^{i \rightarrow j}(t_k) + \varepsilon^{j \rightarrow i}(t_k)]/2$. With Eq. (25.2) subtracted from Eq. (25.1), Equations of time measurement that contains only clock error parameter is derived.

$$\rho_{clock}^{ij}(t) = [\rho^{j \rightarrow i}(t) - \rho^{i \rightarrow j}(t)]/2 = c \cdot [\delta t^j(t) - \delta t^i(t)] + [\varepsilon^{j \rightarrow i}(t) - \varepsilon^{i \rightarrow j}(t)]/2 \quad (25.4)$$

Through the combination of one-way PR, distance measurements and clock bias measurements are decoupled so that the ephemeris and clock are estimated, respectively.

Linearizing Eq. (25.3) with Taylor expansion method at the reference state \mathbf{X}^{i*} and \mathbf{X}^{j*} leads that

$$\rho^{ij}(t) = d(\mathbf{X}^{i*}, \mathbf{X}^{j*}) + \left. \frac{\rho^{ij}(t)}{\partial \mathbf{X}^i} \right|_{\mathbf{X}^{i*}} \delta x^i + \left. \frac{\rho^{ij}(t)}{\partial \mathbf{X}^j} \right|_{\mathbf{X}^{j*}} \delta x^j + \cdots + \varepsilon^{ij}(t) \quad (25.5)$$

Then Eq. (25.5) is converted into the linearized measurement equation.

$$z^{ij}(t) = \mathbf{H}^i \delta \mathbf{x}^i(t) + \mathbf{H}^j \delta \mathbf{x}^j(t) + \varepsilon^{ij}(t) \quad (25.6)$$

where $z^{ij}(t) = \rho^{ij}(t) - d(\mathbf{X}^{i*}, \mathbf{X}^{j*})$; Measurement matrix is

$$\mathbf{H}^i = -\mathbf{H}^j = \left. \frac{\rho^{ij}(t)}{\partial \mathbf{X}^i} \right|_{\mathbf{X}^{i*}} = -\left. \frac{\rho^{ij}(t)}{\partial \mathbf{X}^j} \right|_{\mathbf{X}^{j*}} = \left[\frac{x^i - x^j}{d(\mathbf{X}^{i*}, \mathbf{X}^{j*})} \quad \frac{y^i - y^j}{d(\mathbf{X}^{i*}, \mathbf{X}^{j*})} \quad \frac{z^i - z^j}{d(\mathbf{X}^{i*}, \mathbf{X}^{j*})} \right] \quad (25.7)$$

Similarly, taking GAS as a pseudo-satellite (PR need to be corrected with tropospheric delay furtherly), equations of distance measurement that contains only orbit parameter between g -th GAS and i -th satellite is derived.

$$\rho^{ig}(t) = [\rho^{g-i}(t) + \rho^{i-g}(t)]/2 = d(\mathbf{X}^i(t), \mathbf{X}^g(t)) + \varepsilon^{ig}(t) \quad (25.8)$$

In Eq. (25.8), only state vector of i -th satellite is unknown, the state parameters of GAS $\mathbf{X}^g(t)$ are known. The ground position reference is thereby introduced into the satellite state by Eq. (25.8), which overcomes the deficiency problem of the datum mark in data processing only using inter-satellite link measurements. The linearized measurement equation is

$$z^{ig}(t) = \mathbf{H}^i \delta \mathbf{x}^i(t) + \varepsilon^{ig}(t) \quad (25.9)$$

25.2.2 Equations of Motion

Satellites will be affected by a variety of forces when they run around the Earth. For navigation satellites, only the Earth's gravity, sun, moon, planet's gravity, solar radiation pressure and relativistic effects is considered. Assuming the satellite is in a free state of motion, the equation of the i -th satellite can be written as

$$\dot{\mathbf{X}}^i(t) = f^c(\mathbf{X}^i, \mathbf{w}^i) \quad (25.10)$$

where, \mathbf{w}^i is vector of Gaussian white noise, i.e.

$$E[\mathbf{w}^i(t)] = 0, \quad E[\mathbf{w}^i(t)(\mathbf{w}^i(\tau))^T] = \mathbf{Q}^i(t)\delta(t - \tau) \quad (25.11)$$

where, $\mathbf{Q}(t)$ is the system process noise covariance matrix. It is symmetric non-negative definite matrix; $\delta(t - \tau)$ is the Dirac function, which satisfies

$$\delta(t - \tau) = \begin{cases} 0 & t \neq \tau \\ \infty & t = \tau \end{cases}, \quad \int \delta(\tau) d\tau = 1 \quad (25.12)$$

Noting that continuous function f^c has a linear relationship with the system disturbance \mathbf{w}^i , Eq. (25.10) can be written as follows

$$\dot{\mathbf{X}}^i(t) = \mathcal{F}^c(\mathbf{X}^i(t)) + \mathbf{G}\mathbf{w}^i \quad (25.13)$$

where, $\mathbf{G} = [0_{3 \times 3} \quad \mathbf{I}_{3 \times 3} \quad 0 \quad 0]^T$; Eq. (25.13) is non-linear equation, it is linearized by Taylor expansion in the reference state of \mathbf{X}^{i*} .

$$\dot{\mathbf{X}}^i(t) = \mathcal{F}^c(\mathbf{X}^{i*}(t)) + \left. \frac{\partial \mathcal{F}^c(\mathbf{X}^i)}{\partial \mathbf{X}^i} \right|_{\mathbf{X}^{i*}} (\mathbf{X}^i(t) - \mathbf{X}^{i*}(t)) + \dots + \mathbf{G}\mathbf{w}^i \quad (25.14)$$

Equation (25.14) is written compactly as:

$$\delta \dot{\mathbf{x}}^i(t) = \mathbf{F}(t) \delta \mathbf{x}^i(t) + \mathbf{G}\mathbf{w}^i \quad (25.15)$$

where $\delta \dot{\mathbf{x}} = \dot{\mathbf{X}} - \mathcal{F}^c(\mathbf{X}^*)$, $\mathbf{F}(t)$ is the dynamics matrix (Wen 2009).

$$\mathbf{F}(t) = \left. \frac{\partial \mathcal{F}^c(\mathbf{X})}{\partial \mathbf{X}} \right|_{\mathbf{X}^*} = \begin{bmatrix} 0 & \mathbf{I}_{3 \times 3} & 0 & 0 \\ \frac{\partial \mathcal{F}^c}{\partial \mathbf{r}} & \frac{\partial \mathcal{F}^c}{\partial \mathbf{r}} & \frac{\partial \mathcal{F}^c}{\partial \mathbf{p}_D} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Big|_{\mathbf{X}^*} \quad (25.16)$$

Equation (25.15) gives the state equation of stochastic linear continuous systems in form of an state increment, and its general solution is

$$\delta \mathbf{x}^i(t) = \mathbf{\Phi}^i(t, t_0) \delta \mathbf{x}_0^i(t) + \mathbf{G} \int_{t_0}^t \mathbf{\Phi}^i(t, \tau) \mathbf{w}^i(\tau) d\tau \quad (25.17)$$

where, $\mathbf{\Phi}(t, t_0)$ is system state transition matrix, it is the solution of the following matrix equations.

$$\dot{\mathbf{\Phi}}^i(t, t_0) = \mathbf{F}(t) \mathbf{\Phi}^i(t, t_0), \quad \mathbf{\Phi}^i(t_0, t_0) = \mathbf{I} \quad (25.18)$$

where, \mathbf{I} is unit matrix with the same dimension as $\mathbf{F}(t)$. $\mathbf{\Phi}^i(t, t_0)$ has the following properties further.

$$\Phi^i(t, \tau)\Phi^i(\tau, t_0) = \Phi^i(t, t_0), \quad [\Phi^i(t, \tau)]^{-1} = \Phi^i(\tau, t) \quad (25.19)$$

In the actual numerical solution process, Eq. (25.17) also need to be discretized from t_{k-1} to t_k .

$$\delta \mathbf{x}_k^i = \Phi^i(t_k, t_{k-1})\delta \mathbf{x}_{k-1}^i + \mathbf{G} \int_{t_{k-1}}^{t_k} \Phi^i(t_k, \tau) \mathbf{w}_{k-1}^i(\tau) d\tau \quad (25.20)$$

In sampling interval from t_{k-1} to t_k , $\mathbf{w}_{k-1}^i(\tau)$ can be considered as a constant, let it be denoted as \mathbf{w}_{k-1}^i , and set

$$\mathbf{G} \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) = \mathbf{G}_k \quad (25.21)$$

Then the discrete state equation is

$$\delta \mathbf{x}_k^i = \Phi_{k-1}^i \delta \mathbf{x}_{k-1}^i + \mathbf{G}_k^i \mathbf{w}_{k-1}^i \quad (25.22)$$

where, Φ_{k-1}^i denotes the state transition matrix from t_{k-1} to t_k . The statistical properties of \mathbf{w}_{k-1}^i is

$$E[\mathbf{w}_{k-1}^i] = 0, \quad E[\mathbf{w}_{k-1}^i (\mathbf{w}_{k-1}^i)^T] = \mathbf{Q}_{k-1}^i \quad (25.23)$$

25.3 Orbit Determination with Whole-Constellation Centralized EKF

The state vector \mathbf{X} and their improved values $\delta \mathbf{x}$ of whole-constellation centralized extended Kalman filter (WCCEKF) contains the entire state of the all satellites in the constellation.

$$\mathbf{X} = \left[(\mathbf{X}^1)^T \quad (\mathbf{X}^2)^T \quad \cdots \quad (\mathbf{X}^i)^T \quad \cdots \quad (\mathbf{X}^n)^T \right]^T$$

$$\delta \mathbf{x}_k = \left[(\delta \mathbf{x}_k^1)^T \quad (\delta \mathbf{x}_k^2)^T \quad \cdots \quad (\delta \mathbf{x}_k^i)^T \quad \cdots \quad (\delta \mathbf{x}_k^n)^T \right]^T$$

where, n is the sum of navigation satellites. The state equation of all satellites in constellation can be obtained through Eq. (25.22).

$$\delta \mathbf{x}_k = \Phi_k \delta \mathbf{x}_{k-1} + \mathbf{G}_k \mathbf{w}_{k-1} \quad (25.24)$$

where, both Φ_k and \mathbf{G}_k are diagonal matrix.

$$\begin{aligned}\Phi_{k-1} &= \text{diagonal} [(\Phi_{k-1}^1) \quad (\Phi_{k-1}^2) \quad \cdots \quad (\Phi_{k-1}^i) \quad \cdots \quad (\Phi_{k-1}^n)] \\ \mathbf{G}_k &= \text{diagonal} [(\mathbf{G}_k^1) \quad (\mathbf{G}_k^2) \quad \cdots \quad (\mathbf{G}_k^i) \quad \cdots \quad (\mathbf{G}_k^n)^T]\end{aligned}$$

w_{k-1} denotes noise vector, $w_{k-1} = [(w_{k-1}^1) \quad (w_{k-1}^2) \quad \cdots \quad (w_{k-1}^i) \quad \cdots \quad (w_{k-1}^n)]$, its statistical characteristics are as follows:

$$E[\mathbf{w}_{k-1}] = 0, \quad E[\mathbf{w}_{k-1}^i (\mathbf{w}_{k-1}^j)^T] = \mathbf{Q}_{k-1}^i \delta_{ij} \quad (25.25)$$

$$E[\mathbf{w}_{k-1}, \mathbf{w}_{k-1}] = \mathbf{Q}_{k-1} = \text{diagonal} [\mathbf{Q}_{k-1}^1 \quad \mathbf{Q}_{k-1}^2 \quad \cdots \quad \mathbf{Q}_{k-1}^i \quad \cdots \quad \mathbf{Q}_{k-1}^n] \quad (25.26)$$

where, δ_{ij} is function of Kronecker- δ .

Equations (25.6) and (25.9) can be unified as

$$z_k = \mathbf{H}_k \delta \mathbf{x} + \varepsilon_k \quad (25.27)$$

where, for inter-satellites measurement, $\mathbf{H}_k = [\mathbf{0} \cdots (\mathbf{H}^i)^T \cdots \mathbf{0} \cdots (\mathbf{H}^j)^T \cdots \mathbf{0}]^T$, $z_k = z_k^{ij}(t_k)$, $\varepsilon_k = \varepsilon^{ij}(t_k)$; for GAS measurement, $\mathbf{H}_k = [\mathbf{0} \cdots (\mathbf{H}^i)^T \cdots \mathbf{0}]^T$, $z_k = z_k^{ig}(t_k)$, $\varepsilon_k = \varepsilon^{ig}(t_k)$. Assumed that $\mathbf{R}_k = \text{Cov}(\varepsilon_k, \varepsilon_k)$ is measurement error covariance matrix, given the initial values $\bar{\mathbf{X}}_0 = E(\mathbf{X}_0^*)$ and $\bar{\mathbf{P}}_0 = \text{Cov}(\mathbf{X}_0^*, \mathbf{X}_0^*)$. The centralized extended Kalman filter equations are:

$$\bar{\mathbf{P}}_k = \Phi_{k-1} \hat{\mathbf{P}}_{k-1} \Phi_{k-1}^T + \mathbf{G}_{k-1} \mathbf{Q}_{k-1} \mathbf{G}_{k-1}^T \quad (25.28)$$

$$\mathbf{K}_k = \bar{\mathbf{P}}_k \mathbf{H}_k^T (\mathbf{H}_k^T \bar{\mathbf{P}}_k \mathbf{H}_k + \mathbf{R}_k)^{-1} \quad (25.29)$$

$$\begin{aligned}\delta \hat{\mathbf{x}}_k &= \mathbf{K}_k z_k \\ \hat{\mathbf{X}}_k &= \bar{\mathbf{X}}_k + \delta \hat{\mathbf{x}}_k\end{aligned} \quad (25.30)$$

$$\hat{\mathbf{P}}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \bar{\mathbf{P}}_k \quad (25.31)$$

In WCCEKF algorithm, inter-satellite measurement data is combined with satellite-to-GAS measurement data. The state vector to be estimated takes orbit parameters of all satellites in constellation as an array \mathbf{X} , which dimension is

$$N = 6n + \sum_{i=1}^n D_i$$

D_i is the sum of systematic error parameters and kinetic the parameters that are need to be estimated for i -th satellite. State vector of each satellites is associated to that of other satellites through the state vector covariance matrix, which dimension is $N \times N$. WCCEKF algorithms required to calculate matrix $(\mathbf{H}_k^T \bar{\mathbf{P}}_k \mathbf{H}_k + \mathbf{R}_{kk})$ with

Table 25.1 The calculation amount of the filtering

Process	The calculation amount of the traditional EKF
The complexity of the process of time updated	$4N^3 + N^2$
The complexity of the process of measurement updated	$N^3 + 3MN^2 + NM^2 + 2MN + 2M^3$
The complexity of the whole algorithm	$5N^3 + 3MN^2 + NM^2 + N^2 + 2MN + 2M^3$

dimension $M \times M$ (M is the dimensions of the measurement vector), and calculated amount of its inverse matrix is proportional to M^3 . Table 25.1 shows the calculated amount of filtering algorithm.

If data processing is performed as WCCEKF mode, mainframe computer in MCS will integrate inter-satellite measurement data and satellite-to-GAS measurement data to finish orbit determination, then produces navigation message and integrity of information. The process is complex and the calculated amount is huge, so centralized EKF cannot be implemented in on-board computer of satellite which has limited computing power.

In addition, the risk of this centralized EKF is also big, once the satellite and the backup satellite that implement autonomous navigation data processing fail, the entire navigation constellation will not work. Therefore the algorithm of the on-board distributed data processing is needed to study. Namely, on-board computer of each satellite only deals with measurement data related to itself.

25.4 Distributed EKF Orbit Determination

25.4.1 The Reduced Order EKF Based on Iterative Cascade

For the mode of autonomous navigation based on inter-satellite links, when the measurement equations of one satellite are linearized, the reference state of the other satellites is needed. Taking into account that the autonomous navigation reference ephemeris is provided early by MCS, they have enough precision. So it can be assumed that state vector $(\hat{X}^j, \{j = 1, 2, \dots, n | j \neq i\})$ of other satellites is known. A unified satellite measurement equation can be derived from Eqs. (25.6) and (25.9) for the i -th satellite.

$$z_k^i = H_k^i \delta x_k^i + v_k^i \tag{25.32}$$

where, for inter-satellite measurement, $z_k^i = z_k^{ij}(t_k)$, $v_k^i = H_k^j \delta x_k^j + \varepsilon^{ij}(t_k)$; for GAS measurement, $z_k^i = z_k^{ig}(t_k)$, $v_k^i = \varepsilon^{ig}(t_k)$, and measurement error covariance matrix can be assumed as $\text{Cov}(\varepsilon^{ig}(t_k), \varepsilon^{ig}(t_k)) = \text{Cov}(\varepsilon^{ij}(t_k), \varepsilon^{ij}(t_k)) = R_k^i$, initial state is given as $\bar{X}_0^i = E(X_0^{i*})$, $\bar{P}_0^i = \text{Cov}(X_0^{i*}, X_0^{i*})$. Then Extend Kalman Filter (EKF) equation is:

$$\bar{P}_k^i = \Phi_{k-1}^i \hat{P}_{k-1}^i \Phi_{k-1}^{iT} + G_{k-1}^i Q_{k-1}^i G_{k-1}^{iT} \quad (25.33)$$

$$K_k = \bar{P}_k^i H_k^{iT} (H_k^{iT} \bar{P}_k^i H_k^i + R_k^i)^{-1} \quad (25.34)$$

$$\begin{aligned} \delta \hat{X}_k^i &= K_k^i z_k^i \\ \hat{X}_k^i &= \bar{X}_k^i + \delta \hat{X}_k^i \end{aligned} \quad (25.35)$$

$$\hat{P}_k^i = (I - K_k^i H_k^i) \bar{P}_k^i \quad (25.36)$$

In such a distributed mode, only the measurement data related to satellite itself is included in the measurement equation onboard. The dimension of state X^i is $6 + D_i$. It greatly reduces the computational complexity. But obviously, due to that the state of each satellite needs to be estimated, the above process is approximated, so it is a reduced-order suboptimal filter. In order to improve filtering accuracy, a common approach is to iterate the above process. That is, within a data frame, the satellites' state vector and their covariance matrix are sent out as reference state for other satellites to linearize the measurement equation; Once one satellite receives other satellites' state vector and covariance matrix, it again linearize measurement equation and calculate new measurement matrix H^i , the filter gain, and covariance updated. When the difference of estimated state before and after iteration of the all satellites is less than the convergence threshold, the entire state estimation of the constellation is converged, and it goes into new measurement cycle. So it is called Iterative Cascade EKF (ICEKF)

When ICEKF is adopted in autonomous navigation for the low orbit formation flying satellite [18], it handles large number of space-borne GPS measurement data and a small quantity of inter-satellite measurement data, the estimation process could achieve convergence after 3–4 iterations, and its performance is good in one measurement cycle.

However, when ICEKF is used in autonomous navigation for the high orbit formation flying satellite and a large number of inter-satellite ranging are handled, it needs 4 to 5 iterations before the solution convergence, resulting in the complicated process and huge communication. Sometime unstable solutions are yielded. It shows poor performance [1, 2].

25.4.2 The Reduced Order EKF Based on the Increased Measurement Covariance

The primary problem of ICEKF is that: when the state of the i -th satellite is estimated, it assumes that the states of other satellites are known. In fact, they have error in the process of being estimated. So ICEKF needs multiple iterations to approach the optimal solution. For the autonomous navigation constellation based

on inter-satellite link, due to large number of satellites, ICEKF requires not only a heavy communication load, but also more time-consuming. The process is complicated, and it is difficult to realize.

One of solutions is to let the measurement covariance matrix R_k^i absorb the error of the states of other satellites. Namely, it is to increase R_k^i and it is called increased measurement covariance R EKF (IREKF). R_k^i can be added constant ΔR_k^i , but ΔR_k^i is difficult to accurately reflect all introduced error. So it is necessary to consider a dynamic method to increase R_k^i . Let us re-examine the error of satellite measurement Eq. (25.32) between i -th satellite and j -th satellite, the corresponding measurement covariance matrix is following:

$$\begin{aligned}
 E[v_k^i v_k^{iT}] &= E[(\mathbf{H}_k^j \delta \mathbf{x}_k^j + \varepsilon^{ij}(t_k))(\mathbf{H}_k^j \delta \mathbf{x}_k^j + \varepsilon^{ijT}(t_k))] \\
 &= E[\mathbf{H}_k^j \mathbf{x}_k^j \mathbf{x}_k^{jT} \mathbf{H}_k^{jT} + \mathbf{H}_k^j \mathbf{x}_k^j \varepsilon^{ijT}(t_k) + v_{ij} \mathbf{x}_k^{iT} \mathbf{H}_k^{jT} + \varepsilon^{ij}(t_k) \varepsilon^{ijT}(t_k)] \\
 &= E[\mathbf{H}_k^j \mathbf{x}_k^j \mathbf{x}_k^{jT} \mathbf{H}_k^{jT}] + E[\varepsilon^{ij}(t_k) \varepsilon^{ijT}(t_k)] \\
 &= \mathbf{H}_k^j \bar{\mathbf{P}}_k^j \mathbf{H}_k^{jT} + R_k^i
 \end{aligned} \tag{25.37}$$

When there are multiple inter-satellite measurements, the measurement covariance matrix is increased to

$$R_{kBU}^i = \mathbf{H}_k^j \bar{\mathbf{P}}_k^j \mathbf{H}_k^{jT} + R_k^i \tag{25.38}$$

where, R_{kBU}^i is measurement covariance matrix that is increased; $\bar{\mathbf{P}}_k^j$ is the prediction state covariance of the satellite related to measurement.

With R_k^i in Eq. (25.34) replaced with R_{kBU}^i in Eq. (25.38), the execution from Eq. (25.33) to (25.36) is to perform IREKF, it is the reduced order EKF which increases the observation covariance, which means that measurements include not only measurement errors, but also related satellite state errors. It is expected to reduce the number of iterations of estimation, even without iteration.

25.5 Simulation and Analysis

The navigation constellation is designed as Walker 24/3/2:55°, 22,116 km. The chosen orbit dynamical models are: the earth's gravitational effects of (70 × 70); the lunar, solar and planetary perturbations; the solar radiation pressure and general relativistic forces, and RK8 variable step integrator is used. IERS96 model is adopted for Earth Orientation Parameters.

It is assumed that each satellite establishes links with 4 satellites. They are 2 satellites in front and in the rear in orbit plane, and 2 adjacent satellites in right orbit and left orbit. To avoid ground disturbance, satellite links above a height of 1000 km are considered. TDMA mode is adopted for inter-satellite measurement and data transmitting. The total error of inter-satellite PR with UHF program is 0.5 m (1σ).

8 monitoring stations in China are set up in the simulation, they are Xiamen, Kashi, Bjing, Lhasa, Sanya, Urumqi, Jiamusi and Xi'an. With minimum elevation 10° , The Hopfield/Marini model is used in tropospheric delay correction for the satellite-GAS PR. The total error of ranging is 3 m (1σ).

In order to compare the basic performance of distributed filtering algorithm, simulation does not consider the impact of complex factors, that is, the PR measurement noise of inter-satellites is assumed to be the standard normal distribution without the impact of pollution distribution. Under such conditions, it is conducive to analyze and compare the accuracy and performance of distributed filtering algorithm.

Simulation and orbit determination process are as follows:

- (1) Satellite orbit simulation. The 'theory orbit' of satellites, PRs of the inter-satellite links and satellite-GAS links are produced;
- (2) Satellite orbit estimation. The satellite orbit is estimated and compared with the above-mentioned 'theory orbit' to analyze the accuracy of orbit estimated.

The errors of the orbit determination with WCCEKF, ICEKF and IREKF are shown from Figs. 25.1, 25.2 and 25.3.

Only the Orbit determination error of satellite SV-01 is showed in above diagrams because that other satellites are similar to satellite SV-01 in the Walker constellation. The followings should be noted:

- (1) After filtering results tend to steady, the average RMS of the position with WCCEKF is about 1.6 m, while that with IREKF is about 2.9 m; that with ICEKF is about 4.5 m. The average RMS of position error, radial error, along-track error and normal error with three algorithms are shown in Fig. 25.4.
- (2) WCCEKF is the optimal estimation with the highest precision and the bigger amount of calculation; While ICEKF and IREKF is distributed sub-optimal estimation with small amount of calculation. This is due to an assumption that state of other satellite states is known without any error. To some extent,

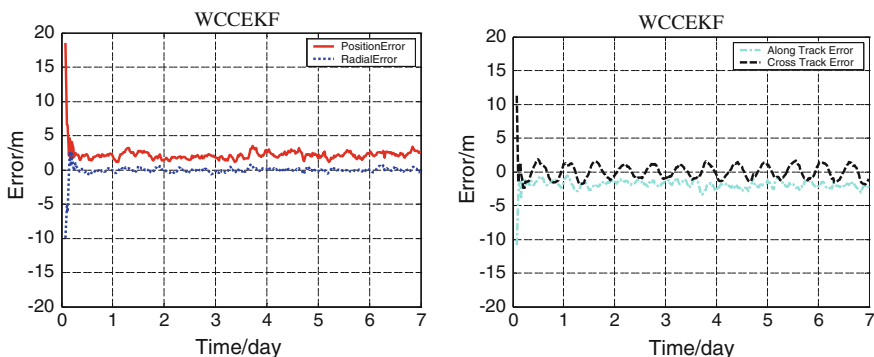


Fig. 25.1 Orbit determination error of SV-01 with WCCEKF

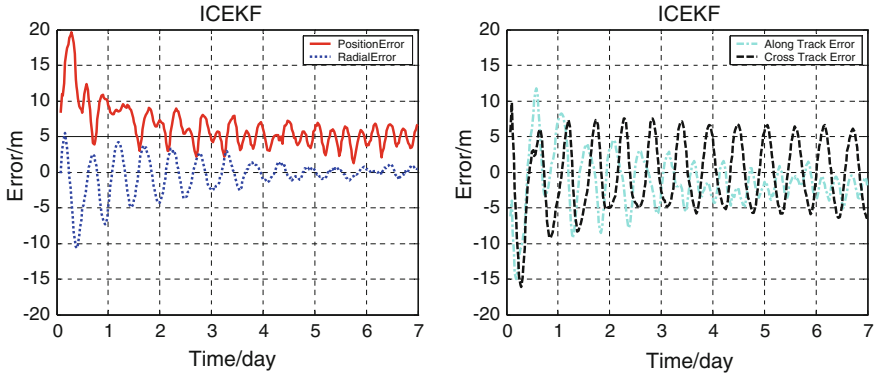


Fig. 25.2 Orbit determination error of SV-01 with ICEKF

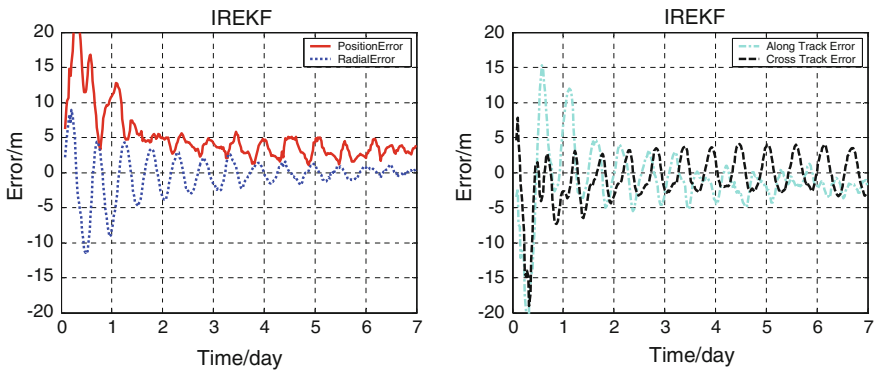


Fig. 25.3 Orbit determination error of SV-01 with IREKF

Fig. 25.4 The average RMS of Orbit determination error with three algorithms

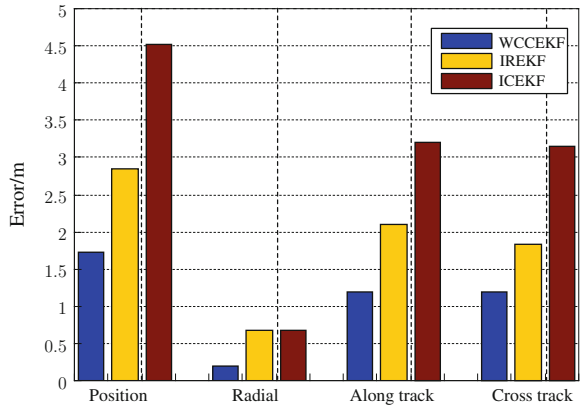


Table 25.2 Performance of WCCEKF, ICEKF and IREKF

Algorithm	Description	Amount of calculation	Amount of Communication	Navigation Accuracy with UHF link
IREKF	Increased measurement covariance R EKF	Minimum	Minimum	Better
ICEKF	Iterative Cascade EKF	Normal	Normal	Normal
WCCEKF	Whole-constellation centralized EKF	Maximum	Maximum	best

IREKF improves the estimate accuracy by increasing the measurement covariance without any iteration, resulting in better accuracy and less amount of calculation and communication than ICEKF.

25.6 Conclusion

The performances of IREKF, ICEKF and WCCEKF are shown in Table 25.2. Compared to others, IREKF is distributed sub-optimal estimation with small amount of calculation which is suitable to fulfill distributed data processing in on-board computer of satellites in GNSS.

References

1. Ferguson P (2003) Distributed estimation and control technologies for formation flying spacecraft. Department of Aero/Astro, Massachusetts Institute of Technology
2. Park C, How JP (2001) Precise relative navigation using augmented CDGPS. In: ION-GPS Conference
3. Rajan JA, Rawicz H BP et al (2003) Modernizing GPS autonomous navigation with anchor capability. GNSS 2003, Portland
4. Kaplan ED, Hegarty CJ (2006) Understanding GPS: principles and applications 2nd edn.. Artech House Inc, Norwood
5. Menn MD, Bernstein H (1994) Ephemeris observability issues in the global positioning system (GPS) autonomous navigation (AUTONAV) IEEE, pp 677–680
6. Ananda MP (1990) Global positioning system(GPS) autonomous navigation. In: IEEE position location and navigation symposium, Las Vegas, Nevada, pp 435–455
7. Rajan JA (2002) Highlights of GPS IIR autonomous navigation. In: Proceedings of the ION 58th annual meeting, Albuquerque, NM, pp 354–363
8. Rajan JA, Orr M (2003) On-orbit validation of GPS IIR autonomous navigation. In: Proceedings of the ION 59th annual meeting, Albuquerque, NM, pp 411–419
9. Fisher SC, Kamran G (1999) GPS IIF—the next generation. In: Proceedings of the IEEE
10. Podlesney D (2009) GPS III space segment

11. Ollie L, Lany B, Art G et al (2003) GPS III system operations concepts. In: Proceedings of the 16th international technology meeting of the satellite division of the institute of navigation. The Institute of Navigation, Portland, pp 380–388
12. Eissfeller B, Wolf TZR, Hammesfabr J, Hornbostel A, Halm JH, Tavella P (2000) Autonomous satellite state determination by use of two-directional links. *Int J Satell Commun* 18:325–346
13. <http://www.gpsdancer.org>
14. Wang F, Liu W, Ling X (2011) Distributed autonomous orbit determination of global navigation constellation via inter-satellite pseudo-ranging measurements. In: China satellite Navigation conference (CSNC), Beijing
15. Yang L, Zhou J, Liu J, Niu F (2011) Study of distributed onboard orbit determination method for global navigation satellite constellation. In: China satellite navigation conference (CSNC), Beijing
16. Song X, Mao Y, Jia X, Wu X (2010) The distributed processing algorithm for autonomously updating the ephemeris of navigation satellites by inter-satellite links. *Geom Inf Sci Wuhan Univ* 35(10):1161–1164
17. Yang Y, Gao W, Zhang X (2010) Robust Kalman filtering with constraints: a case study for integrated navigation. *J Geod* 84:373–381
18. Ferguson P, How JP (2003) Decentralized estimation algorithms for formation flying spacecraft. In: Proceedings of the AIAA guidance, navigation, and control conference, pp AIAA-2003–5442
19. Wen Y, Zhu J, Li Z, Liao Y (2009) Simulation and analysis of integrated orbit determination of satellites constellation. *J Astronaut* 30(1):155–163

Chapter 26

Strategy and Accuracy Analysis of Space-Borne GPS Single-Frequency Real-Time Orbit Determination

Fuhong Wang, Lei Guo and Xuewen Gong

Abstract The real-time orbit determination for Low Earth Orbiters (LEOs) is generally based on dual-frequency tracking data from on-board GPS receivers, and positioning accuracies for LEOs currently vary from 0.5 to 1.0 m with sophisticated reduced dynamic orbit determination techniques. In this paper, different strategies are designed and applied to real-time orbit determination only using single-frequency pseudo-range and carrier phase measurements to analyze their corresponding performances. Then some simulative tests are carried out to process the GRACE-A space-borne GPS data in the different solar sunspot activities. The test results demonstrate that the position and velocity accuracy (3DRMS) are up to 0.9 m and 0.9 mm/s respectively using only single-frequency GPS pseudo-range data, and their accuracies could be improved to 0.55 m and 0.55 mm/s using single-frequency combination of pseudo-range and carrier phase measurements, which is close to the dual-frequency real-time orbit determination. Therefore a low-cost single frequency space-borne GPS receiver can be used in real-time orbit determination for LEO missions when appropriate strategies and methods would be selected.

Keywords Single-frequency · Space-borne GPS · Real-time orbit determination · Strategy · Accuracy analysis

26.1 Introduction

Nowadays, on-board GPS receivers constitute a primary means of tracking and navigation spacecraft launched into LEO [1]. A high accuracy is required in many science missions and applications such as altimetry, gravimetry, or atmospheric sounding call for sub-decimeter position accuracy (see [2–4]). While, for some remote

F. Wang · L. Guo (✉) · X. Gong
School of Geodesy and Geomatics, Wuhan University, Wuhan, China
e-mail: 845755829@qq.com

sensing and resources satellites such as the remote sensing satellite KOMPSAT-1 of the Republic of Korea, the remote sensing satellite QuickBird of American and the resources satellite BIRD of Germany, they require relatively lower positioning accuracy. The single-frequency GPS receivers are more preferable for such missions. Using single-frequency GPS receivers can save costs and reducing energy consumption. What's more, it can also reduce the total quantity of data for storage and downlink. It has been widely used in home and abroad on the space missions [3].

With the implementation of the GPS satellite ephemeris improvement plan in 2006, the accuracy of the broadcast ephemeris have been improved greatly [4]. The ionospheric range delay is left as the dominant error source for the determination of real-time orbits for LEO satellites. While dual-frequency GPS observations can take advantage of dispersive nature of the ionosphere and remove its effect with ionospheric-free combination. As to single-frequency observations, there are two main ways to deal with the ionospheric delay, one is to relay on external model such as the Klobuchar model [5] or the global ionospheric maps (GIM) [6]. Based on the GIM from JPL, Peng Dongju used the scale factor of ionosphere method and demonstrated that the effect of eliminating ionospheric delay is remarkable [7]. But the theory is complex, and relay on the GIM from JPL. If it was used in real-time orbit determination, it will increase the data volume uploaded from the ground and the cost for daily maintenance, reduce the autonomic operation of satellite. Another method is taking advantage of the fact that the ionosphere delays of first order for phase and code have precisely the same value but opposite signs. Using GRAPHIC combination can eliminate the ionosphere delay effectively [4]. Precise orbit determination was processed with this method by Montenbruck etc. and a 3DRMS accuracy of 0.1 m was achieved [2–4]. But when this method is used in the real-time orbits determination for LEO satellites, it is quite different with the precise orbit determination in the strategies and accuracy. The data used in for this paper are GRACE-A GPS data in times of low and high solar activity, then design different real-time orbit determination strategies, use the self- developed real-time orbit determination software to simulate real time orbit determination. And compare the result with the precise orbit given by JPL, conclude our analysis by comparing the real-time orbit determination accuracy with different strategies.

26.2 Real-Time Orbit Determination Strategy

LEO satellites real-time orbit determination means that use GPS tracking data to estimate the parameters of the satellites like position and velocity on board. And the parameters are offered for attitude control and data collection to realize the autonomous operation for LEO satellites. Currently real-time orbit determination algorithm uses appropriate dynamic model and extended Kalman filter to estimate the parameters of satellites with space-born GPS tracking data. The detail description on the theory and algorithm of real-time orbit determination is available in reference [8].

Table 26.1 Same dynamical model and reference frame of three orbit determination methods

Dynamical model	
Gravity field model	EGM2008 70×70
N body gravity	Approximate formula to calculate the sun and moon position
Atmospheric drag	Improved Harris-priester density model
Solar radiation pressure	Cannonball model
Empirical acceleration	First-order Gauss-Markov process
Reference frame	
Earth centered Inertial (ECI)	J2000.0
Earth-centered, Earth-fixed (ECEF)	WGS84
Earth rotation parameter	IERS

The ionospheric range delay has great influence on the code pseudo-range and carrier phase observation of space-born GPS receivers, up to 10 m or more. Based on the different ways of processing the ionospheric delay with single-frequency GPS data, we design two kind of orbit determination strategies. Method 1: using improved Klobucahr model to correct ionospheric delay with C/A code; Method 2: using GRAPHIC combination to eliminate ionospheric delay. As a comparison, we also add dual-frequency pseudo-range ionospheric-free combination, called Method 3.

The dynamic model, reference frame and measurement model used in this paper are shown in Tables 26.1 and 26.2.

26.2.1 Measurement Model

The height of the Leo satellite is commonly at 160–2000 km, so the influence of troposphere delay can be ignored. The observed L1 carrier phase and C/A code range can be described as following expression:

Table 26.2 Measurement models of three orbit determination methods

	Method 1	Method 2	Method 3
GPS observations	C/A	C/A, L1	P1, P2
Ionospheric model	Improved Klobuchar model	GRAPHIC combination	Dual-frequency iono-free combination
Receiver clock	Stochastic model to simulate	Stochastic model to simulate	Stochastic model to simulate
GPS ephemeris	GPS broadcast ephemeris	GPS broadcast ephemeris	GPS broadcast ephemeris

$$\rho_{L1} = \varphi\lambda = \rho_{geom} + c(\delta t_r - \delta t^s) - V_{ion} - N\lambda + \varepsilon_{L1} \quad (26.1)$$

$$\rho_{C/A} = \rho_{geom} + c(\delta t_r - \delta t^s) + V_{ion} + \varepsilon_{C/A} \quad (26.2)$$

Here, ρ_{L1} , $\rho_{C/A}$ are respectively phase measurement and C/A code range observations on L1 carrier. N is ambiguity and λ is wavelength of L1. δt_r is receiver clock offset and δt^s satellite clock offset. V_{ion} is correction of ionospheric first-order delay on L1. ε_{L1} and $\varepsilon_{C/A}$ are respectively measurement noise of code and phase. ρ_{geom} is the geometric range between receiver and the GPS satellite.

The first-order ionospheric delay accounts for more than 99.9 % of the ionospheric delay effect. The correction of ionospheric delay will directly influence the accuracy of orbit determination when processing space-born GPS data from single-frequency receivers.

For method 1, we use the improved Klobuchar model to correct ionospheric delay. The main idea of this algorithm is that choose a suitable effective height of the residual ionosphere, assume a Chapman profile to describe the altitude variation of the electron density, calculate the height of equivalent compression monolayer above the receiver altitude, and estimate the fractional total electron content (TEC) above the receiver. Then with the position of LEO and GPS satellites, estimations of puncture point of GPS signals and its zenith distance on the equivalent compression monolayer is derived. Thirdly calculate total electron content on the vertical direction of puncture point. The ionospheric path delay is then obtained from the fractional TEC, zenith distance and etc. A 90 % correction of the total ionospheric effects has been demonstrated by Simsky [9].

For method 2, by adding code and carrier range, we obtain a combined expression:

$$\begin{aligned} \rho_{C_1L_1} &= (\rho_{L1} + \rho_{C/A}) / 2 \\ &= \rho_{geom} + c(\delta t_r - \delta t^s) - N\lambda/2 + \varepsilon \end{aligned} \quad (26.3)$$

$\rho_{C_1L_1}$ is the GRAPHIC combination measurement, ε is the measurement noise for the combination. By this way, the influence of ionospheric delay is eliminated effectively. What's more, it exhibits a noise with standard deviation

$$\sigma(\varepsilon) = \frac{1}{2} \sqrt{\sigma^2(\varepsilon_{C/A}) + \sigma^2(\varepsilon_{L1})} \quad (26.4)$$

For the fact that the standard deviation of measurement noise for carrier phase (about 1 mm) is much lower than that of C/A code. Then we obtain

$$\sigma(\varepsilon) \approx \frac{1}{2} \sigma(\varepsilon_{C/A}) \quad (26.5)$$

This roughly half the noise of C/A code.

26.2.2 Estimation

The method of using GPS tracking data as observation, combining with dynamic model and EKF, is implemented to estimate the parameters. The filter state for Method 1 and 3 is completely the same.

$$X_1 = [\mathbf{r}_{1 \times 3} \quad \mathbf{v}_{1 \times 3} \quad b_r \quad \dot{b}_r \quad C_D \quad C_R \quad \mathbf{a}_{1 \times 3}] \quad (26.6)$$

Here, $\mathbf{r}_{1 \times 3}$ and $\mathbf{v}_{1 \times 3}$ are the LEO satellite position and velocity vector in ECI. $b_r = c\delta t_r$, $\dot{b}_r = c\dot{\delta t}_r$, and δt_r are $\dot{\delta t}_r$ receiver clock offset and clock rate. The speed of light in vacuo is c . The atmosphere drag and radiation pressure coefficients C_D and C_R , and empirical accelerations $\mathbf{a} = [a_R \quad a_T \quad a_N]$ in radial, tangential and normal direction.

As to Method 2, parameters to estimate are quite different with Method 1 and 3. Not only the LEO satellite position, velocity, receiver clock offset, clock rate, atmosphere drag, radiation pressure and empirical accelerations are needed, but also a vector $\mathbf{b} = (B_1, B_2, \dots, B_n)$ of carrier-phase biases is needed for tracked GPS satellites. So the state vector is

$$X_2 = [\mathbf{r}_{1 \times 3} \quad \mathbf{v}_{1 \times 3} \quad b_r \quad \dot{b}_r \quad C_D \quad C_R \quad \mathbf{a}_{1 \times 3} \quad \mathbf{b}_{1 \times n}] \quad (26.7)$$

n is the number of the observed satellites.

Unlike traditional carrier phase processing, RTK and post-processing, where ambiguities are assumed to be constants [10]. Ambiguities in Method 2 are modeled as random walk process. And if proper parameter of process noise is set, it can absorb part of the slowly time varying ephemeris errors, and improve the real-time orbit determination accuracy with single-frequency GPS data. In this paper, the process noise covariance matrix for bias state is

$$Q_b = \sigma_b^2 \times I_{n \times n} \quad (26.8)$$

The value $\sigma_b^2 = (5 \times 10^{-3} dt)^2$ is chosen as optimal. dt is time interval of Kalman filter.

Evidently, an initial value $b = \rho_{C/A} - \rho_{L1} - 2V_{ion}$ for the bias of GRAPHIC measurement can be obtained. The ionospheric delay V_{ion} is given by the improved Klobuchar model. The transition matrix, coefficient matrix of observation and some part of the covariance matrix are nearly the same with the ones in Method 1 and 3. More details are shown in reference [8, 10, 11].

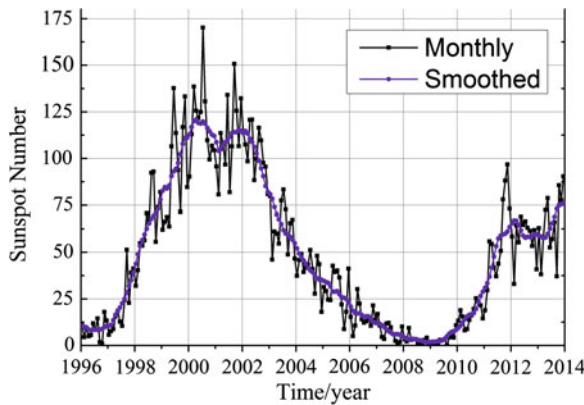


Fig. 26.1 Sunspot number for the past 18 years

26.3 Processing and Analysis

The dominant error source in the processing of the space-born GPS data from single-frequency receiver is the ionospheric delay. The vertical total electron content is tied closely to the solar activity. Usually the solar activity is described by sunspot numbers.

The variance of sunspot numbers in the past 18 years is shown in Fig. 26.1. It is obvious that year 2011 was the year of high solar activity and 2008 was the low one. The data used for this study is GRACE-A GPS tracking data for the 7 days 33–39 of year 2008 and 2011.

When processing the GRACE-A GPS data of year 2008 with the three methods, the 3DRMS position and velocity error are provided in Fig. 26.2. Mean 3DRMS of three orbit determination schemes for days 33-39/2008 are shown in Table 26.3. It is obvious that in the times of low solar activity an accuracy of 0.72 m in position and

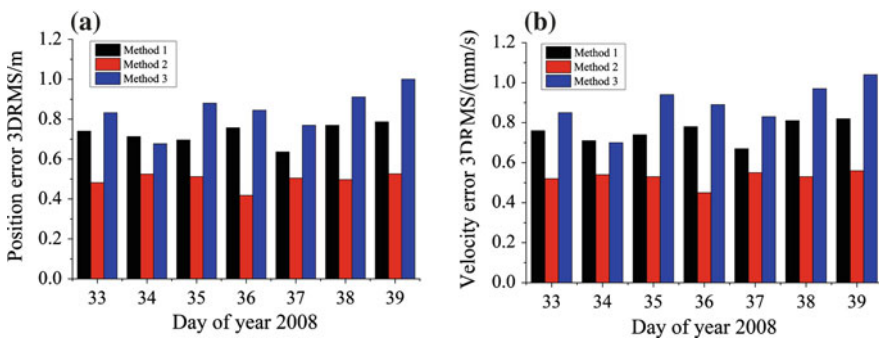


Fig. 26.2 The position error and velocity error (3DRMS) of three real-time orbit determination methods for year 2008. **a** Position error. **b** Velocity error

Table 26.3 Mean 3DRMS of three orbit determination schemes for days 33–39/2008

	R/m	T/m	N/m	RMS/m
Method 1	0.27	0.55	0.38	0.72
Method 2	0.13	0.40	0.25	0.49
Method 3	0.24	0.57	0.57	0.84

0.8 mm/s in velocity are achieved with Method 1. And 0.49 m and 0.5 mm/s with Method 2, as to Method 3 we obtain 0.84 m and 0.9 mm/s. It shows that when processing with the improved Klobucahr model to correct the ionospheric delay, higher accuracy than the dual-frequency pseudo-range ionospheric-free combination is gained. Because in the times of low solar activity a suitable scale factor is obtained with the improved Klobuchar model, the ionospheric delay gets corrected efficiently. A noteworthy performance is also achieved with Method 2, the GRAPHIC combination measurement, whose accuracy is much higher than the other two methods. Because the GRAPHIC combination can totally eliminate the influence of first order ionospheric delay, and GRPHIC combination measurement yields a noise roughly half of single-frequency pseudorange. In this method proper process noise parameter is needed, the bias parameters may absorb part of slowly time varying broadcast ephemeris orbit and clock errors.

In method 3 we get poor accuracy with dual-frequency pseudo-range ionospheric-free combination, that’s because the combination yields a noise much larger than the one of single-frequency pseudo-range. And the discrepancies of the ionospheric delay correction also have influence on the results.

Figure 26.3 shows a sample plot of GRACE-A position errors in R/T/N with the data of year 2008 by the three methods. It is obvious that the use of GRAPHIC combination provides the highest accuracy.

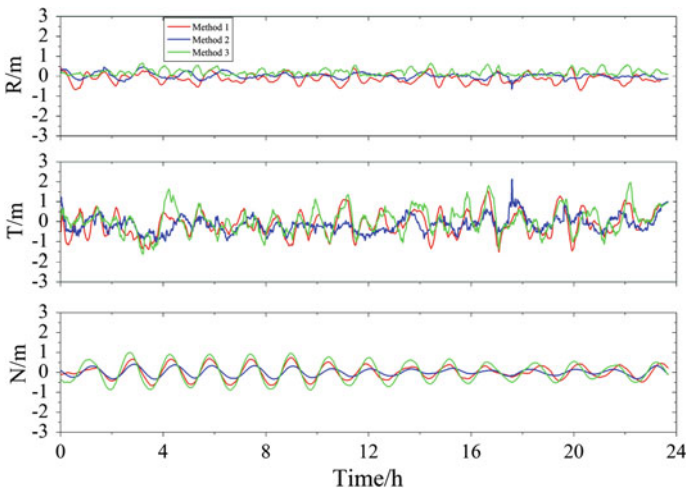


Fig. 26.3 The R/T/N position error of three real-time orbit determination methods for year 2008

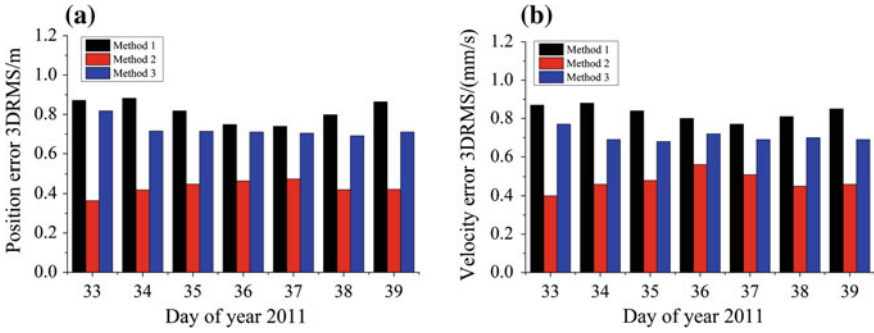


Fig. 26.4 The position error and velocity error (3DRMS) of three real-time orbit determination methods for year 2011. **a** Position error. **b** Velocity error

Figure 26.4 shows the 3DRMS position and velocity errors of year 2011, in the times of high solar activity, with the three methods. Figure 26.5 gives a sample plot of GRACE-A position errors in R/T/N with the data of year 2011 (Table 26.4).

It is prominent that in the times of high solar activity the accuracies of position and velocity (0.82 m, 0.9 mm/s) are worse than 0.72 m and 0.8 mm/s in 2008 with Method 1. In the times of high solar activity the fractional TEC above the receiver estimated by Chapman profile function do not match with the truly fact well. So the scale factor we get from the improved Klobuchar model is not suitable. But as for Method 2, the mean 3DRMS keep at a level of 0.5 m in position and 0.5 mm/s in velocity with different ionospheric circumstances.

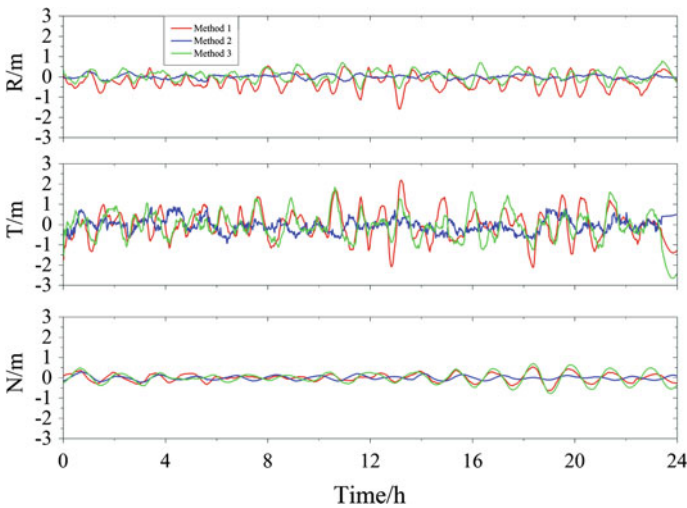


Fig. 26.5 The R/T/N position error of three real-time orbit determination methods for year 2011

Table 26.4 Mean 3DRMS of three orbit determination schemes for days 33–39/2011

	R/m	T/m	N/m	RMS/m
Method 1	0.40	0.67	0.21	0.82
Method 2	0.12	0.37	0.17	0.43
Method 3	0.26	0.63	0.23	0.72

26.4 Summary

A study of GPS-based real-time orbit determination with GRACE-A single-frequency GPS tracking data has been conducted by three methods under different solar activities. We all get satisfactory accuracy with different ways to eliminate ionospheric delay.

- (1) Accuracies of 0.72 m and 0.8 mm/s are obtained by method 1 with improved Klobuchar model in the times of low solar activity, better than the accuracies of dual-frequency pseudo-range orbit determination. But in the times of high solar activity, ionospheric delay influence the method 1 greatly, the accuracies are worse than those of dual-frequency pseudo-range orbit determination, but still accuracies of 0.82 m and 0.9 mm/s are obtained.
- (2) Thanks for the GRAPHIC combination, the ionospheric delay is eliminated efficiently. No matter in the times of high or low solar activity, we can always get accuracies which are better than 0.55 m and 0.55 mm/s in position and velocity 3DRMS. Comparing with dual-frequency pseudo-range orbit determination, the accuracies get improved about 40 %.

Both methods based on single-frequency GPS tracking data can obtain accuracies better than 1 m and 1 mm/s, which meet the accuracy requirement of real-time orbit determination for microsattellites. That could be valuable as reference for the plan of carrying single-frequency GPS receivers on the satellites. High accuracy results are obtained with the GRAPHIC combination method, but the combination increases the complexity of data processing for the use of carrier-phase observation. When using this method, cycle-slip should be taken into account. But there are not many efficient real-time cycle-slip detection algorithms for space-born single-frequency GPS tracking data. So the next work is to improve an efficient algorithm to detect cycle-slips in single-frequency GPS data.

References

1. Junyong C, Yaming D, Pengfei C (2007) Development and progress in GNSS. *J Geodesy Geodyn* 5:1–4
2. Bock H, Jaggi A, Dach R (2009) GPS single-frequency orbit determination for low earth orbiting satellites. *Adv Space Res* 43(05):783–791
3. Montenbruck O (2003) Kinematic GPS positioning of LEO satellites using ionosphere-free single frequency measurements. *Aerosp Sci Technol* 7(05):396–405

4. Montenbruck O, Ramos-Bosch P (2008) Precision real-time navigation of LEO satellites using global positioning system measurements. *GPS Solut* 12:187–198
5. Montenbruck O, Gill E (2002) Ionospheric correction for GPS tracking of LEO satellites. *J Navig* 55:293–304
6. Hwang Y, Born GH (2005) Orbit determination strategy single-frequency global positioning system data. *J Spacecr Rocket* 42:896–891
7. Dong-ju P, Bin W (2012) The application of GIM in precise orbit determination for LEO satellites with single-frequency GPS measurements. *Acta Astron Sinica* 53:36–50
8. Wang F (2006) Theory and software development on autonomous orbit determination with space-borne GPS measurement. Wuhan University, Wuhan
9. Simsky A (2006) Standalone real-time navigation algorithm for single-frequency ionosphere-free positioning based on dynamic ambiguities (DARTS-SF). In: *Proceeding of ION GPS 2006*, pp 1–10
10. Fu-hong W (2010) A Kalman filtering algorithm for precision real-time orbit determination with space-borne GPS. *Geom Inf Sci Wuhan Univ* 35(6):653–656
11. Guo X, Zhang Q, Zhao Q (2013) Precise orbit determination for LEO satellites using single-frequency GPS observations. *Wuhan University*, 2013(2):41–46

Chapter 27

Orbit Determination Using Combined GPS + Beidou Observations for Low Earth Cubesats: Software Validation in Ground Testbed

Yang Yang, Xiaokui Yue, Geshi Tang, Hongzheng Cui
and Baiyan Song

Abstract Nowadays, satellites in low earth orbit (LEO) can benefit from Global Navigation Satellite System (GNSS), such as Global Positioning System (GPS) of United States, to estimate positions and velocities. As China's BeiDou Navigation Satellite System (BDS) has been formally operational since the end of 2012, standalone Beidou and combined GPS + Beidou positioning techniques tend to be applied in the future space missions. However, no LEO satellites have been operated with Beidou receivers at present. Hence Beidou-based precise orbit determination (POD) technique is required to be tested and verified on ground at first stage. This study is to test the GPS + Beidou orbit determination software in ground testbed. GNSS data collected from iGMAS (International GNSS Monitoring and Assessment Service) and MGEX (Multi-GNSS Experiment) stations are processed in both static and kinematic PPP (Precise Point Positioning) modes. Decimetre level of positioning accuracy is achieved. The inter-system biases between GPS and Beidou are estimated and analysed. Results indicate that GPS + Beidou solutions are more precise than the standalone GPS solutions.

Keywords GPS + Beidou · Kinematic PPP · LEO cubesats · Orbit determination

Y. Yang (✉) · X. Yue
School of Astronautics, Northwestern Polytechnical University (NPU),
Xi'an 710072, People's Republic of China
e-mail: yiyinfeixiong@gmail.com

Y. Yang · X. Yue
National Key Laboratory of Aerospace Flight Dynamics (AFDL), NPU,
Xi'an 710072, People's Republic of China

G. Tang · H. Cui · B. Song
National Key Laboratory of Aerospace Flight Dynamics (AFDL),
Beijing Aerospace Control Centre, Beijing 100000, People's Republic of China

27.1 Introduction

Nowadays, many more satellites are equipped with Global Positioning System (GPS) receivers onboard to provide precise orbit determination (POD) solutions. GPS-based POD has become the primary tracking strategy since GPS receivers provide near-continuous observations with excellent geometric information, which allow orbit determination accuracy at the centimetre level for the current satellites missions [1, 2].

As China's Beidou Navigation System (BDS) becomes operational for regional navigation and positioning services since December 2012, which consists of 5 Geostationary Earth Orbit satellites (GEO), 5 Inclined Geosynchronous Orbit (IGSO) satellites and 4 Medium Earth Orbit (MEO) satellites (see Fig. 27.1 for ground tracks), this emerging GNSS system begins making great contributions to the current navigation/positioning applications in Asia-Pacific region specifically. Afterwards, the Beidou global system is expected to be completed by the end 2020 with 5 GEOs, 3 IGSOs and 27 MEOs in the whole constellation [3]. It is inevitable signals from BDS will be utilised for real satellites POD in Low Earth Orbit (LEO) in a standalone system mode or combined with GPS system.

Precise point positioning (PPP) technique [4] has been widely used for LEO satellite POD in post-processing, which is utilising the undifferenced pseudorange and carrier phase measurements, appropriately dealing with the errors and biases in measurements, and applying precise GNSS orbits and clocks. In comparison with differential GNSS (DGNSS) positioning methods, PPP methods can generate positioning results without observations of extra reference stations. Most publications about Beidou-based PPP algorithm are processing observations from static ground stations [5–7]. It is reported that the accuracy of GPS-based PPP can reach at several centimetres, while Beidou-based PPP performance is some worse

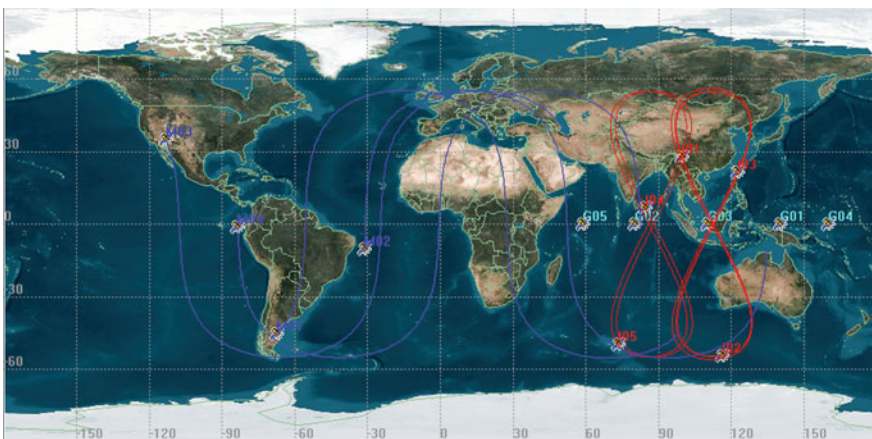


Fig. 27.1 Beidou navigation constellation ground tracks

currently. All the results have been worked out in the terrestrial applications. It is indicated BeiDou can not only provide positioning services as a standalone navigation system but also can be combined with GPS to improve the overall integrity and robustness. Hence, it is of great potential to apply Beidou in the space missions.

This paper aims to implement GPS + Beidou PPP algorithm for LEO satellites POD applications but being tested with observation data collected from ground stations. The positioning performance of the standalone BDS is evaluated and compared with GPS solutions. A kinematic-PPP mode is used to process the combined GPS + Beidou observations. Results indicate that the software are ready for the future real flight data processing.

27.2 GNSS Tracking Stations

The GNSS precise point positioning (PPP) technique requires high quality product (orbits and clocks) application, since their errors directly affect the quality of positioning. iGMAS (International GNSS Monitoring and Assessment Service) aims to provide GNSS users open services with an unified standards, especially for Beidou users [8]. Within this workgroup, many iGMAS stations have been built up around the world to track GNSS satellites (see Fig. 27.2), which are mainly located in mainland of China. The tracking stations are equipped with UR4B0 receivers and different types of antennas, which can receive both GPS and Beidou signals, and are developed by the UNICORE Company in Beijing, China (<http://www.unicorecomm.com/english/>). As one of the working groups of the International GNSS Service (IGS), the Multi-GNSS Experiment (MGEX) has been set-up to extend precise ephemeris data and bias information to all new constellations including Beidou in the product portfolio. Tracking of the BeiDou satellites is supported by a subset of

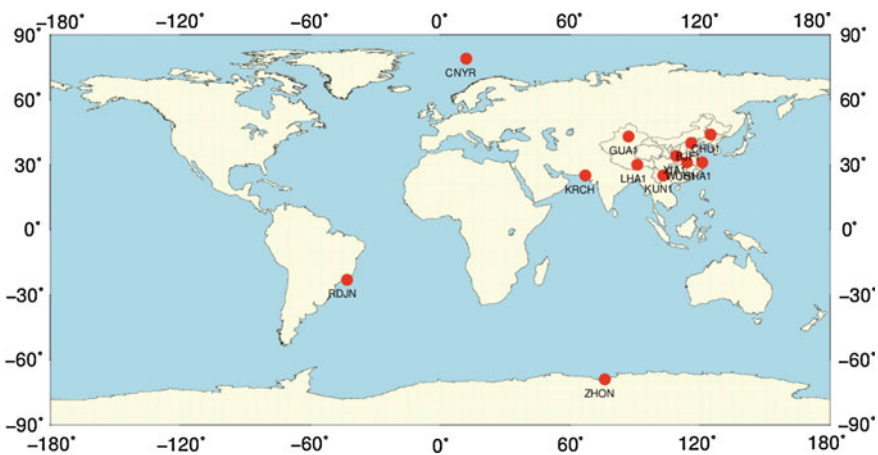


Fig. 27.2 iGMAS tracking stations

stations in the MGEX network as well, especially in Asia-Pacific region. This research will test the PPP-based kinematic POD algorithm using some data collected from some stations in both iGMAS and MGEX working groups.

27.3 Cubesat Mission

A pico-nano LEO cubesat project is proposed jointly by Beijing Aerospace Control Centre and China Spacesat Co. Ltd., which includes a nano-satellite and three pico-satellites for scientific experiments in low earth orbit. More importantly, GPS + Beidou dual-mode dual-frequency receivers are equipped on both nano and pico satellites. For the first time will the spaceborne GPS + Beidou receiver be used for POD solutions in space missions. As a several-centimetre POD accuracy is required to implement the scientific experiments, elaborate ground tests for the GPS + Beidou based POD software validation need to be conducted.

Orbit determination accuracy will highly depend on the quality of spaceborne GNSS receivers. There are many terrestrial-type multi-GNSS receivers in the market, which support multi-system, multi-frequency of GNSS signals. Among them, products of Trimble, Javad and Septentrio have been operating in MGEX stations. The Unicore company from China also provides good industrial GNSS receivers to support the emerging Beidou-based positioning and navigation applications on ground. Up to now, some spaceborne high-quality GPS receivers have achieved good performance in the space missions, such as Blackjack receiver in GRACE mission developed at JPL [9]. However, there are no GPS + Beidou receivers on-the-fly. To support the small satellite project and achieve the scientific tasks on-orbit, a miniaturised GPS + Beidou receiver is under development. Data collected using this receiver will be tested and analysed in the future.

27.4 Methodology

27.4.1 Observation Equations

A kinematic-PPP algorithm is used for satellite POD solutions. The ionosphere-free observations equations for combined GPS and Beidou constellations can be written as in a simple form:

$$\begin{aligned}
 PR_r^{s,GPS} &= \rho_r^{s,GPS} + c(cl_r - cl^s) + v_{r,PR}^{s,GPS} \\
 CP_r^{s,GPS} &= \rho_r^{s,GPS} + c(cl_r - cl^s) + a_r^s + v_{r,CP}^{s,GPS} \\
 PR_r^{s,BDS} &= \rho_r^{s,BDS} + c(cl_r - cl^s + ISB_r^{BDS}) + v_{r,PR}^{s,BDS} \\
 CP_r^{s,BDS} &= \rho_r^{s,BDS} + c(cl_r - cl^s + ISB_r^{BDS}) + a_r^s + v_{r,CP}^{s,BDS}
 \end{aligned} \tag{27.1}$$

where ρ is the geometric distance between GNSS satellites and receiver including relevant corrections, such as antenna phase centre offsets/vari-ations (PCO/PCV) and carrier phase wind-up effects. Due to a lack of available antenna information for the BeiDou satellites, a set of PCO values of (+0.6, 0.0, +1.1 m) recommended by MGEX (<http://www.igs.org/mgex/status-BDS#Characteristics>) are used for Beidou antenna PCO corrections. The PCV value of Beidou satellite is neglected. Some built-in attributes of the receiver antennas are unknown, hence receiver PCO/PCV are neither currently accounted for. Beidou satellite attitude control mode has also not been released yet. In this research, a yaw-steering mode has been assumed for Beidou satellites, being the same as that of the GPS Block IIR satellites. In this sense, the carrier phase wind-up effects of Beidou observations are corrected using the same formulation with GPS satellites.

c is the speed of light in vacuum. cl^s and cl_r denote the epoch-wise GNSS satellite and receiver clock offsets, respectively. For precise modelling of GNSS observations, code phase and carrier phase biases (including both GNSS satellite and LEO satellite), namely DCB and DPB should be corrected. The DCB is often assimilated by the clock offset $c(cl_r - cl^s)$ while the carrier phase bias DPB is combined with carrier phase ambiguity (see Eq. 27.2).

$$\begin{aligned} c(cl_r - cl^s) &= c(\bar{cl}_r - \bar{cl}^s) + DCB_r^s \\ a_r^s &= \lambda N_r^s + DPB_r^s - DCB_r^s \end{aligned} \quad (27.2)$$

The compensation for DCB in Beidou code phase measurements is different from that in GPS code phase measurement. More information can be referred to [10]. GPS time is used as the reference in the GNSS system, hence a parameter symbolised as ‘ISB’ is added to account for the inter-system biases in Eq. 27.1, which can be expressed as [11]:

$$\begin{aligned} ISB_r^{BDS} &= c(cl_r - cl^s)^{BDS} - c(cl_r - cl^s)^{GPS} \\ &= ISTD + \Delta DCB \end{aligned} \quad (27.3)$$

where ISTD denotes the inter-system time difference and ΔDCB is the DCB difference between two systems. It is reported that ISTD is stable in one-day interval [11], hence it is defined as a one-day constant. However, the receiver-dependent varying ΔDCB needs to be considered, which leads to one observation equation ‘sacrificed’ in the combined system.

a_r^s denotes the carrier phase ambiguity, which does not hold the integer property any more due to the ionosphere-free combination and the presence GNSS satellite and receiver hardware delays.

v denotes the unmodelled errors or noises, such as multipath effects and thermal noises.

To sum up, when Beidou observations are combined with GPS observation for POD solutions, an additional ISB parameter needs to be estimated together with LEO satellite's positions, reference receiver clock offset and carrier phase ambiguity. For the receiver on the ground, the tropospheric path delay effects are corrected by estimating the zenith wet delay (ZWD). The dry tropospheric delay error is first corrected using Saastamoinen tropospheric model and the ZWD is then estimated as an unknown parameter. The Niell mapping functions are used as hydrostatic and wet mapping functions. Additionally, site correction including solid earth tides, polar tides, ocean loading are also considered in the PPP algorithm.

27.4.2 PPP Processing Settings

A Kalman filter is used for the PPP based orbit determination. Note that the receivers are static in the ground test, but all the data are processed in both static and kinematic modes. Table 27.1 gives an overview of the PPP processing settings in this research. It should be mentioned that a factor of 3 is added to Beidou GEOs to compensate for the bigger orbit determination errors in comparison with other GNSS satellites.

As mentioned, the precise GNSS orbits and clocks determination is one of the most important factors affecting the PPP performance. For the emerging Beidou navigation system, orbits and clocks determination results have been published in many literatures [6, 12, 13]. iGMAS analysis centres are also working on Beidou precise orbit and clock determination, but these products are not public until now. Currently there are some institutions (universities) providing Beidou orbit and clock products in the MGEX framework, such as GFZ (<http://www.igs.org/mgex/products>). This research will use these products for GPS + Beidou PPP algorithm validation, which can be accessed at the CDDIS MGEX product archive (<ftp://cddis.gsfc.nasa.gov/pub/gps/products/mgex/>).

Table 27.1 PPP processing settings

Observation date	20–22 July 2014
Observation interval	30 s
Minimum elevation	10°
Positioning mode	Static/Kinematic
Processing mode	Forward/Combined
Observation types (Beidou)	B1 and B2 code and carrier phase
Observation types (GPS)	L1 and L2 code and carrier phase
Observation noise	$\sigma_{CP} = 0.005$ m for carrier phase
σ_{PR}/σ_{CP} ratio	100
Weighting function	sin(Elevation)

27.5 Ground Data Analysis

Data from two iGMAS stations and two MGEX stations with an additional CUBB station at Curtin University of Technology are processed to test the GPS + Beidou PPP algorithm. The facility information can be found in Table 27.2. The PPP solutions of CUT0 and CUBB are compared with the known reference coordinates (cf. <http://saegnss2.curtin.edu.au/lcdc/>). On the other hand, a forward-backward Kalman smoother is used to solve the reference coordinates of the rest three stations. A smoothed solution is given by:

$$X = \frac{P_{fwd}^{-1}X_{fwd} + P_{bck}^{-1}X_{bck}}{P_{fwd}^{-1} + P_{bck}^{-1}} \quad (27.4)$$

where P denotes the covariance matrix, X denotes the estimated states and subscript ‘fwd’ and ‘bck’ denote the forward and backward runs, respectively. Then the average of the estimated states are regarded as the reference coordinates of the stations.

Table 27.3 gives the PPP accuracies in terms of RMS errors using both GPS and Beidou observations. All the statistics are calculated using the positioning solutions excluding the first two hours of convergence. Several centimetres of 3D RMS of the positioning performance could be achieved in the static-PPP processing, but varying from station to station. Positioning performance of SHA1 station on Day 201 is somewhat worse in terms of static-PPP solution. Particularly, 3 days’ observations from KUN1 station are tested for inter-system bias analysis.

The detailed results of KUN1 station on Day 203 are presented as follows. The Beidou satellites visibility and DOP values are depicted in Fig. 27.3. At least 8 Beidou satellites are visible at KUN1 station. The average GDOP for the whole day’s observation is 2.7. The skyplot of Beidou satellites is shown in Fig. 27.4. The positioning errors in ENU (East, North, Up) coordinates using static PPP algorithm are shown in Fig. 27.5. RMS errors of 0.01, 0.008 and 0.012 m are achieved with the 3D RMS error reaching 0.018 m.

To evaluate the PPP algorithm for satellite orbit determination applications, several tests in a kinematic mode are executed using the static data collected from KUN1 station on the same day (see Fig. 27.6). The positioning error of 0.067 m in terms of 3D RMS is obtained processing GPS + Beidou observations

Table 27.2 Station information

Station	Receiver type	Antenna type
CUT0	TRIMBLE NETR9	TRM 59800.00 SCIS
CUBB	JAVAD TRE_G3TH_8	TRM 59800.00 SCIS
GMSD	TRIMBLE NETR9	TRM 59800.00 SCIS
SHA1	UNICORE UB4BOI	NOV750.R4
KUN1	UNICORE UB4BOI	NOV750.R4

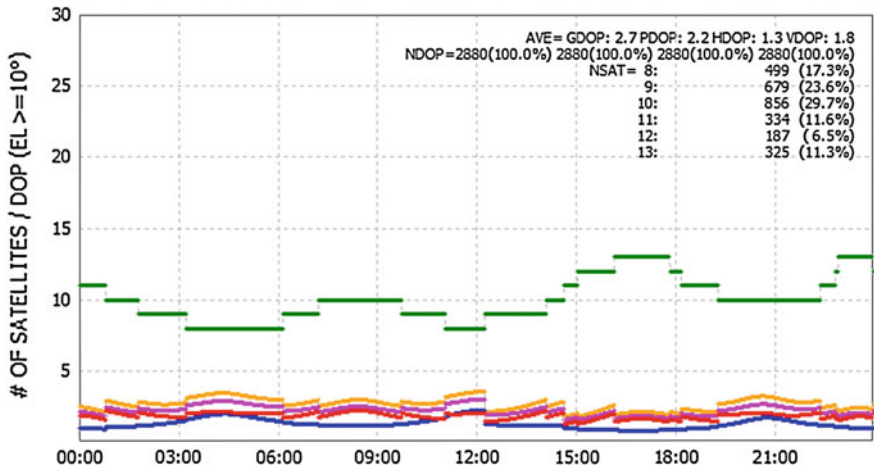


Fig. 27.3 Beidou satellites visibility and DOP values at KUN1 station (green line denotes the total numbers of visible GNSS satellites, orange line denotes GDOP values, magenta line denotes PDOP values, red line denotes VDOP values, red line denotes HDOP values)

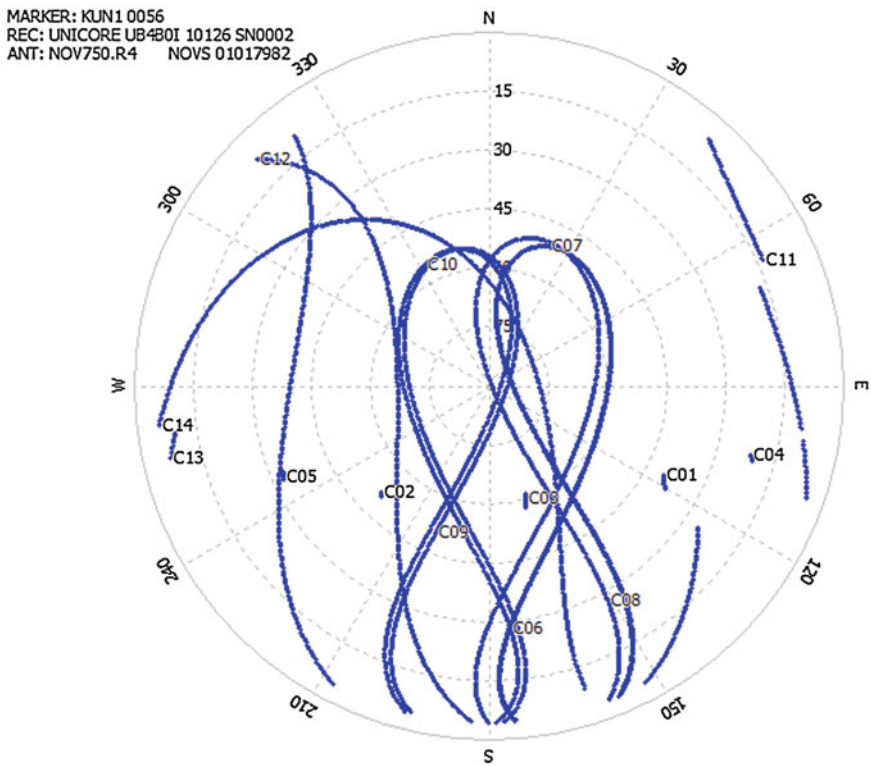


Fig. 27.4 Beidou satellites skyplot at KUN1 station

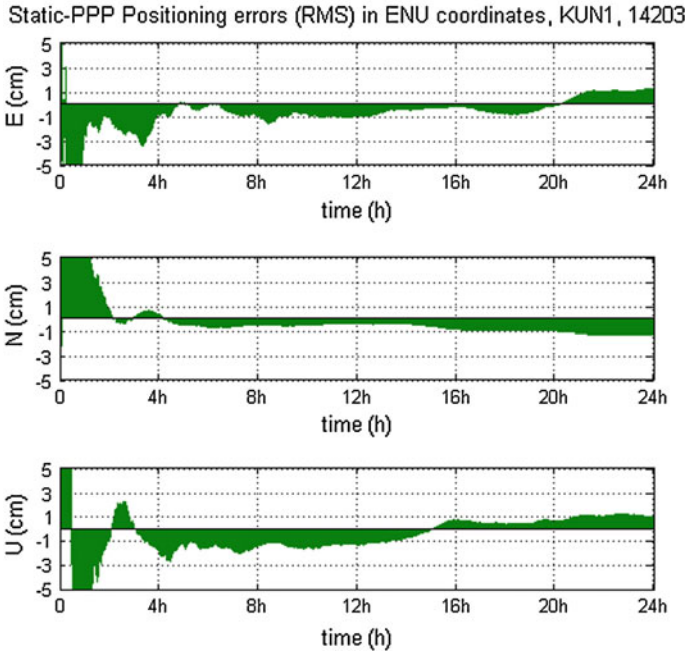


Fig. 27.5 GPS + Beidou static PPP solutions for KUN1 station on Day 203

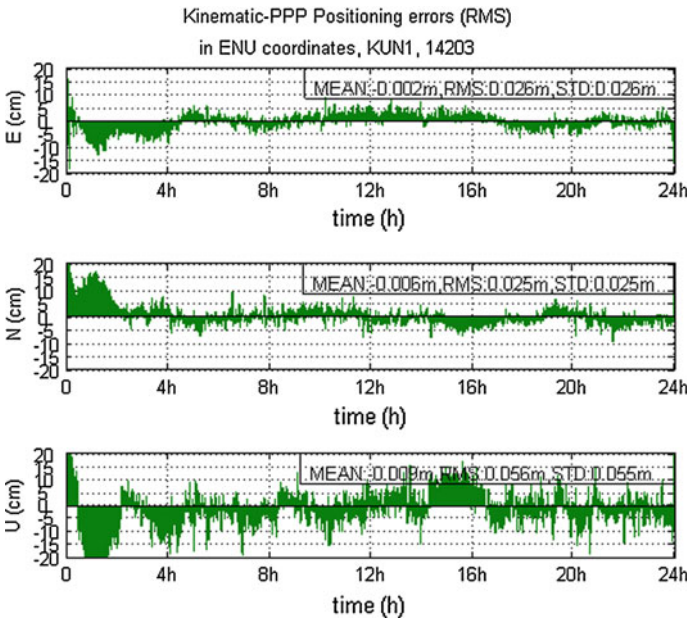


Fig. 27.6 GPS + Beidou kinematic PPP solutions for KUN1 station on Day 203

simultaneously. The Up component accuracy is obviously deteriorated in comparison with other two components. Since Beidou GEOs can be visible only in Asia-Oceania region, there is a long arc LEO satellites cannot receive GEO signals. Therefore it is significant to evaluate the PPP performance without Beidou GEO satellites. The positioning errors are depicted in Fig. 27.7. Without observations from 5 GEOs, obvious accuracy decreases occur in East and Up components. The positioning accuracy becomes more deteriorated in East component when only GPS observations are processed in the PPP algorithm. This is different from the results in static-PPP scheme, where GPS + Beidou solutions are not obviously improved relative to standalone GPS solutions (see Fig. 27.8) in terms of positioning accuracy. In addition, the convergence time becomes longer in the East component.

The ionosphere-free observations residuals are depicted in Figs. 27.9 and 27.10 for the kinematic GPS + Beidou PPP processing of KUN1 on Day 203. The RMS of GPS code phase and carrier phase observations are 1.210 and 0.027 m, respectively. However, these values are worse for Beidou observation residuals. More specifically, the ionosphere-free code observation residuals are fluctuating heavily for some satellites, resulting in a non-zero distribution.

Finally, estimated inter-system biases are depicted in Fig. 27.11. According to Eq. 27.3, these values consist of an approximate constant component and a hardware bias component. It is clearly shown in Fig. 27.11 that estimated ISB are

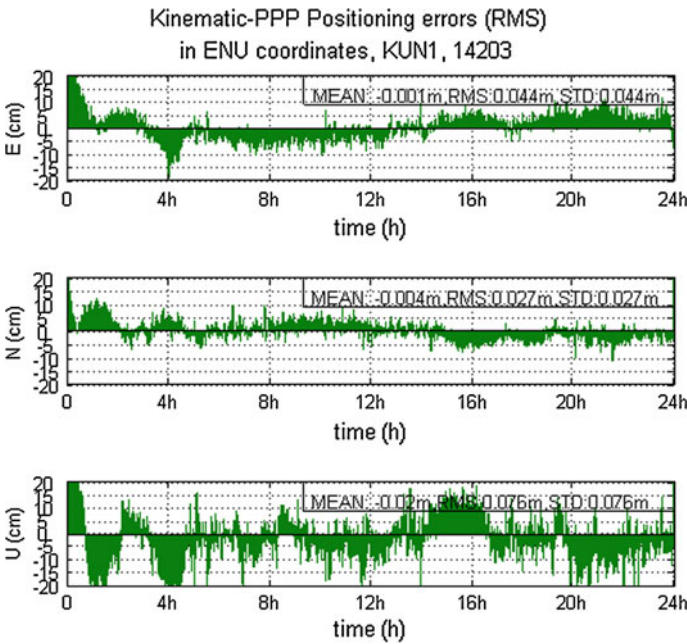


Fig. 27.7 GPS + Beidou kinematic PPP solutions for KUN1 station on Day 203 (without Beidou GEOs)

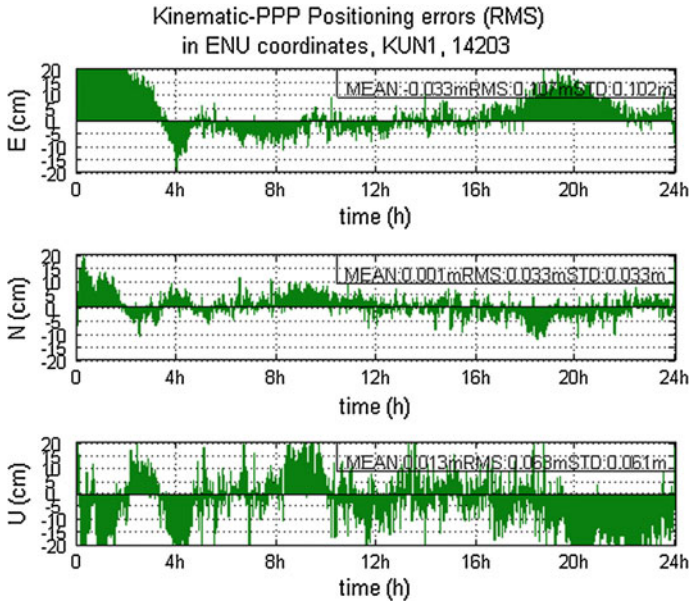


Fig. 27.8 GPS kinematic PPP solutions for KUN1 station on Day 203

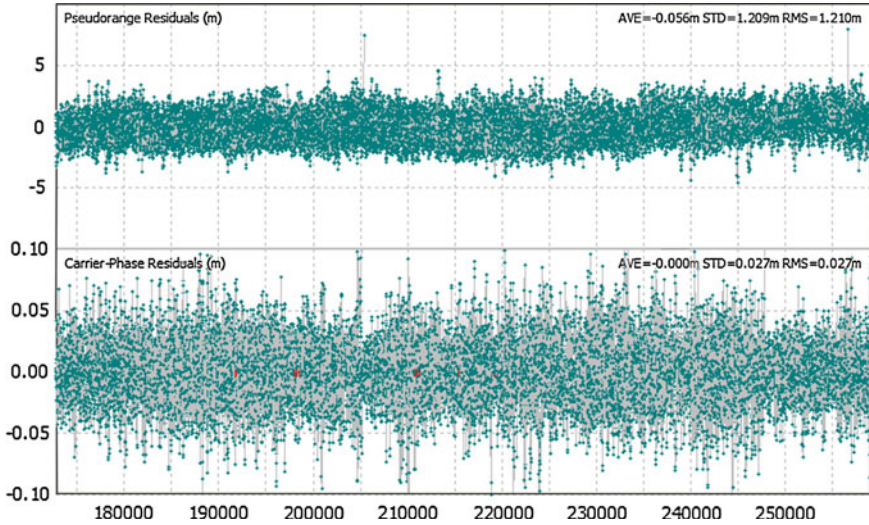


Fig. 27.9 GPS ionosphere-free observation residuals in GPS + Beidou Kinematic PPP solutions for KUN1 station on Day 203

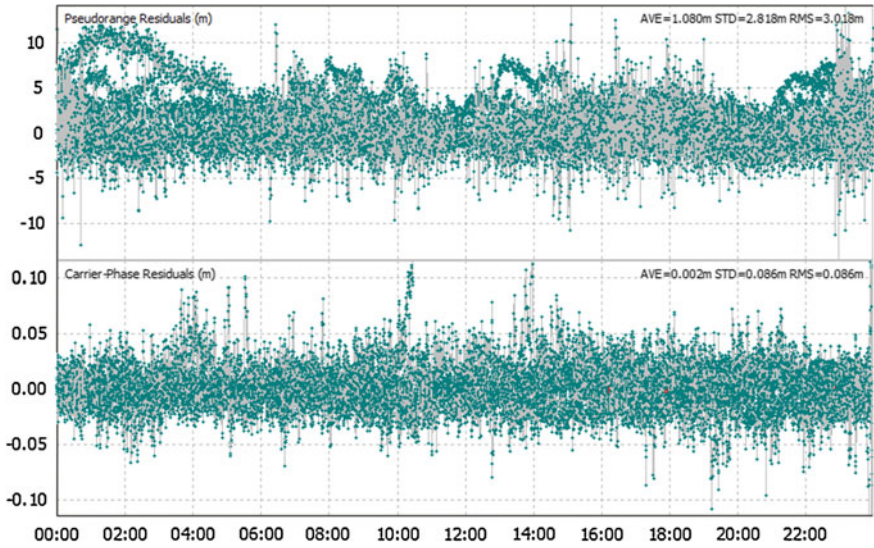


Fig. 27.10 Beidou ionosphere-free observation residuals in GPS + Beidou Kinematic PPP solutions for KUN1 station on Day 203

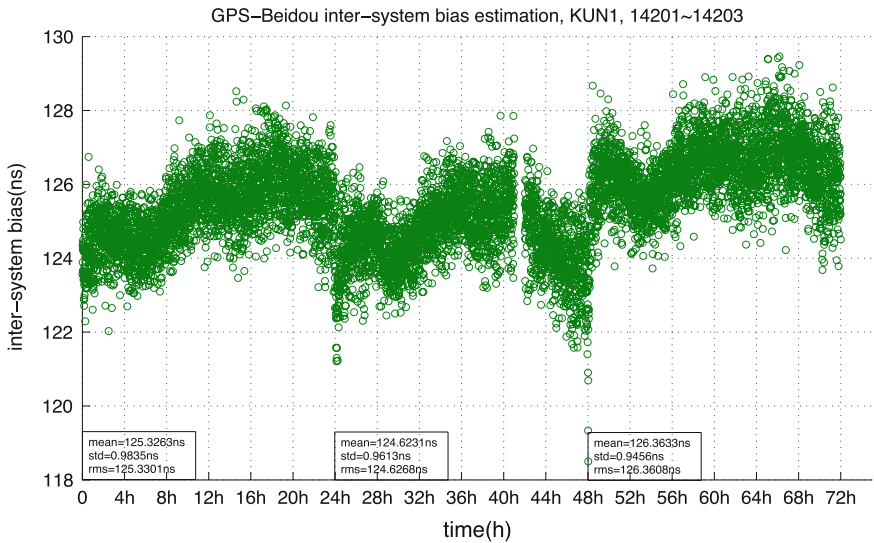


Fig. 27.11 Inter-system biases estimation using static PPP for KUN1 station during Day 201–203

keeping stable during a whole day’s solutions. Approximately 1 ns deviations are observed, which may lead to 0.3 m in positioning error. If the hardware bias ΔDCB is calibrated carefully and the system time difference $ISTD$ is estimated before hand,

Table 27.3 GPS + Beidou PPP solutions statistics (RMS, m)

Stations	PPP modes	RMS errors			
		E	N	U	3D
CUT0	Static	0.008	0.003	0.019	0.021
	Kinematic	0.031	0.027	0.103	0.111
CUBB	Static	0.007	0.004	0.013	0.015
	Kinematic	0.035	0.031	0.086	0.098
GMSD	Static	0.011	0.012	0.025	0.030
	Kinematic	0.039	0.033	0.062	0.080
SHA1	Static	0.031	0.018	0.051	0.063
	Kinematic	0.034	0.028	0.069	0.081
KUN1	Static	0.025	0.010	0.013	0.030
	Kinematic	0.042	0.031	0.072	0.089
	Static	0.019	0.010	0.020	0.029
	Kinematic	0.063	0.029	0.120	0.139
	Static	0.010	0.008	0.012	0.018
	Kinematic	0.026	0.025	0.056	0.067

or the total *ISB* is estimated as one unknown parameters in advance, it is no more necessary to ‘sacrifice’ one observation to solve this parameter in the combined GNSS positioning system (Figs. 27.6 and 27.8 and Table 27.3).

27.6 Concluding Remarks

This paper presents initial assessment of the ground testbed for LEO cubesat orbit determination. GPS + Beidou data collected from tracking stations are processed in kinematic mode to test the PPP based orbit determination algorithm. It is shown that decimetre accuracy could be obtained in terms of the positioning performance. In the future work, more tests will be conducted with real data collected using developed GPS + Beidou receiver and receiver-in-loop simulated data in the LEO scenarios.

Acknowledgments iGMAS BACC centre, MGEX and Curtin University of Technology (CUT) are acknowledged for providing GNSS data. Some figures are plotted by RTKPLOT (<http://www.rtklib.com>). This study is financially supported by the National Natural Science Foundation of China (11172235) and the Doctoral Fund of Ministry of Education of China (20106102110003).

References

1. Kang Z, Tapley B, Bettadpur S, Ries J, Nagel P, Pastor R (2006) Precise orbit determination for the GRACE mission using only GPS data. *J Geodesy* 80(6):322–331
2. Visser P, van den IJssel J, Helleputte TV, Bock H, Jäggi A, Beutler G, Švehla D, Hugentobler U, Heinze M (2009) Orbit determination for the GOCE satellite. *Adv Space Res* 43(5):760–768

3. China Satellite Navigation Office (2013) Beidou navigation satellite system signal in space interface control document open service signal (version 2.0)
4. Zumberge J, Heflin M, Jefferson D, Watkins M, Webb F (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. *J Geophys Res Solid Earth* (1978–2012) 102(B3):5005–5017
5. Shi C, Zhao Q, Hu Z, Liu J (2013) Precise relative positioning using real tracking data from COMPASS GEO and IGSO satellites. *GPS Solut* 17(1):103–119
6. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Teunissen P, Nakamura S (2013) Initial assessment of the COMPASS/BeiDou-2 regional navigation satellite system. *GPS Solut* 17(2):211–222
7. Li M, Qu L, Zhao Q, Guo J, Su X, Li X (2014) Precise point positioning with the beidou navigation satellite system. *Sensors* 14(1):927–943
8. Jiao W, Ding Q, Li J, Lu X, Feng L, Ma J, Chen G (2011) Monitoring and assessment of GNSS open services. *J Navig* 64(11):S19–S29
9. Montenbruck O, Kroes R (2003) In-flight performance analysis of the CHAMP BlackJack GPS receiver. *GPS Solut* 7(2):74–86
10. Montenbruck O, Steigenberger P (2013) The BeiDou navigation message. *J Global Position Syst* 12(1):1–12
11. Chen J, Xiao P, Zhang Y, Wu B (2013) GPS/GLONASS system bias estimation and application in GPS/GLONASS combined positioning. In: 2013 Proceedings China Satellite Navigation Conference (CSNC), Springer, pp 323–333
12. Shi C, Zhao Q, Li M, Tang W, Hu Z, Lou Y, Zhang H, Niu X, Liu J (2012) Precise orbit determination of Beidou satellites with precise positioning. *Sci China Earth Sci* 55(7):1079–1086
13. Steigenberger P, Hugentobler U, Hauschild A, Montenbruck O (2013) Orbit and clock analysis of compass GEO and IGSO satellites. *J Geodesy* 87(6):515–525

Chapter 28

Preliminary Analysis of Positioning Performance with BDS Virtual Reference Station Technology

Dongfeng Yu, Pengbo Li, Guangxing Wang, Mingzhi Zhou and Zhigang Hu

Abstract The principle of GNSS VRS (Virtual Reference Station) was discussed in the paper. The critical technologies of VRS, like virtual station error modeling with reference networks and generation of virtual observables, were studied. The feasibility of BDS VRS technology, as well as the current position performance with BDS VRS, was verified. It is shown that, with BDS VRS, the positioning accuracy inside the local reference networks can reach 1–2 cm, provided baselines of 10–20 km.

Keywords Beidou/BDS VRS · CORS · Relative positioning · RTK

28.1 Introduction

On Dec 27, 2012, the regional deployment of BeiDou satellite navigation system (BDS) had been finished, with five GEO satellites, five IGSO and four MEO ones in orbit. BDS began to formally provide positioning, navigation and timing (PNT) service in Asia-Pacific area, covering most of the area in 55°S–55°N and 0°E–150°E [1]. Pivoting on high precision navigation and positioning with BDS

D. Yu
R&D Center of CASC, Beijing 100034, China

P. Li · G. Wang · M. Zhou · Z. Hu (✉)
GNSS Research Center of Wuhan University, Wuhan 430079, China
e-mail: zhigang.hu@whu.edu.cn

G. Wang
e-mail: gxwang0113@126.com

M. Zhou
e-mail: nanwang_luojia@foxmail.com

[2, 3], study of BDS applications in fields like meteorology and deformation monitoring sprung up. Various technologies based on relative positioning mode are still the important means in obtaining coordinate information with high precision. The continuous operating reference system (CORS) is able to build a positioning frame, long-term, continuous, and reflecting the variations of spatial and temporal information, which helps provide fundamental information directly, fast and effectively and ensures the consistency and integrity of the data. The advent of CORS promotes the wide applications of RTK technologies. The technique of VRS, with its wide coverage, high precision and flexibility, becomes one of the prevalent network RTK technologies.

As the first BDS ground-based augmentation networks in China, the Hubei BDS CORS Network was completed in 2013. Based on CORS network, the BDS VRS technology was able to further improve the positioning accuracy within the CORS-covering area, without any additional cost in hardware. In the paper, the algorithm of BDS VRS was investigated, and its performance in positioning was assessed with real observables.

28.2 Fundamental of VRS

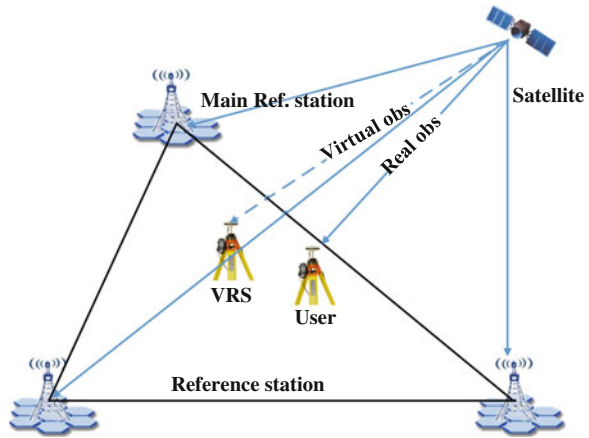
Virtual Reference Station (VRS) technology, which is a kind of network RTK technology, The main idea is to make full use of multiple reference stations' accurate coordinate information to fixed phase ambiguity between reference station at first, then inverse double differential ionosphere error and troposphere error, build the function of the related error within the region and coordinate difference among stations by differential correction, estimate ionosphere error and troposphere error of virtual reference station rapidly according to the approximate coordinates of the user station (virtual reference station) and coordinate difference of the main reference station, and then generate the virtual observations with geometric distance between virtual reference station and the satellite, finally get short baseline of user station and virtual reference station to help the user to achieve high accuracy, real-time positioning, as shown in Fig. 28.1.

The key technology of VRS is calculating the ionosphere error and troposphere error using observation data of known reference stations, building the accurate function model of relative position among virtual reference station and the main reference station and ionosphere error and troposphere error, getting the virtual observation value finally.

28.2.1 Main Error Models of VRS

According to the principle of GNSS relative positioning, the double difference observation equation of reference station can be expressed as:

Fig. 28.1 VRS technical principles



$$\lambda_k \left(\nabla \Delta \varphi_{m,n,k}^{ij} + \nabla \Delta N_{m,n,k}^{ij} \right) = \left(\nabla \Delta \rho_{m,n,k}^{ij} - \nabla \Delta I_{m,n,k}^{ij} + \nabla \Delta T_{m,n,k}^{ij} \right) + \nabla \Delta \varepsilon_{m,n,k}^{ij} \quad (28.1)$$

where $\nabla \Delta \varphi_{m,n,k}^{ij}$, $\nabla \Delta N_{m,n,k}^{ij}$, $\nabla \Delta \rho_{m,n,k}^{ij}$, $\nabla \Delta I_{m,n,k}^{ij}$, $\nabla \Delta T_{m,n,k}^{ij}$, $\nabla \Delta \varepsilon_{m,n,k}^{ij}$ are double difference carrier phase observation value, double difference ambiguity, double difference geometric distance, double difference ionosphere residual error, double difference troposphere residual error and double difference observation noise of satellites i, j and stations m, n respectively, and $k = 1, 2, 3$, λ_k , is the carrier wavelength.

If the reference station coordinates is known and double difference ambiguity has been fixed by other methods, Eq. (28.1) can be rewritten as:

$$V_k = \nabla \Delta T_{m,n,k}^{ij} - \nabla \Delta I_{m,n,k}^{ij} = \lambda_k \left(\nabla \Delta \varphi_{m,n,k}^{ij} + \nabla \Delta N_{m,n,k}^{ij} \right) - \nabla \Delta \rho_{m,n,k}^{ij} \quad (28.2)$$

According to GNSS dual frequency carrier phase observation value and Eq. (28.1), we can get L1 double difference ionosphere residual as follows:

$$\nabla \Delta I_{m,n,1}^{ij} = \frac{f_2^2}{f_1^2 - f_2^2} \left(\nabla \Delta \varphi_{m,n,1}^{ij} + \nabla \Delta N_{m,n,1}^{ij} - \left(\nabla \Delta \varphi_{m,n,2}^{ij} + \nabla \Delta N_{m,n,2}^{ij} \right) \right) \quad (28.3)$$

Obviously, double difference ionosphere residual error, double difference troposphere residual error can be isolated through substituting Eq. (28.2) with Eq. (28.3). According to Eqs. (28.1–28.3), double difference ionosphere error and troposphere error of arbitrary two baselines among reference stations can be get. Next is how to get the function relation of the geometry between the virtual reference station and the main station with computed double difference ionosphere and troposphere. Station coordinate of virtual reference station can be arbitrary given, but in order to facilitate, users tend to use the approximate coordinates as virtual reference

station coordinates. Study shows that the established model accuracy is often associated with the fitting method of using [4, 5]. Commonly used linear combination method, linear interpolation method, the low order surface fitting considering the plane and the elevation, the least square collocation method. Because of the limitation of length, above methods will not be discussed in this paper.

28.2.2 Observation Generation of VRS

Through making single difference of the satellite i, j among the reference station r , virtual reference station v and user, we can get:

$$\Delta\varphi_{r_k}^{i,j} = \frac{1}{\lambda_k} \left(\Delta\rho_{r_k}^{i,j} - \Delta I_{r_k}^{i,j} + \Delta T_{r_k}^{i,j} \right) - \Delta N_{r_k}^{i,j} \quad (28.4)$$

$$\Delta\varphi_{v_k}^{i,j} = \frac{1}{\lambda_k} \left(\Delta\rho_{v_k}^{i,j} - \Delta I_{v_k}^{i,j} + \Delta T_{v_k}^{i,j} \right) - \Delta N_{v_k}^{i,j} \quad (28.5)$$

$$\Delta\varphi_{u_k}^{i,j} = \frac{1}{\lambda_k} \left(\Delta\rho_{u_k}^{i,j} - \Delta I_{u_k}^{i,j} + \Delta T_{u_k}^{i,j} \right) - \Delta N_{u_k}^{i,j} \quad (28.6)$$

Considering the double difference ionosphere residual error $\nabla\Delta I_{r,v,k}^{i,j}$, double difference troposphere residual error $\nabla\Delta T_{r,v,k}^{i,j}$ of the virtual reference station and the main reference station have been calculated in Sect. 28.2.1, Eq. (28.5) can be further expressed as:

$$\Delta\varphi_{v_k}^{i,j} + \Delta N_{v_k}^{i,j} = \frac{1}{\lambda_k} \left(\Delta\rho_{v_k}^{i,j} - \Delta\rho_{r_k}^{i,j} + \Delta\varphi_{r_k}^{i,j} + \Delta N_{r_k}^{i,j} - \nabla\Delta I_{r,v,k}^{i,j} + \nabla\Delta T_{r,v,k}^{i,j} \right) \quad (28.7)$$

According to Eq. (28.7), $\Delta\rho_{v_k}^{i,j}$, $\Delta\rho_{r_k}^{i,j}$, $\Delta\varphi_{v_k}^{i,j}$ all have been known, so the single difference observations of virtual reference station can be calculated. It should be noted that the virtual observations contain the single difference carrier observations $\Delta\varphi_{r_k}^{i,j}$ among the main reference station and ambiguity values $\Delta N_{r_k}^{i,j}$. After obtaining the virtual observation value $\Delta\varphi_{v_k}^{i,j}$, make Eqs. (28.7) and (28.6) consists of double difference, taking into account the close distance of virtual reference station and user station, the composition of double difference troposphere and ionosphere residual error can be neglected, so:

$$\Delta\varphi_{v_k}^{i,j} - \Delta\varphi_{u_k}^{i,j} = \frac{1}{\lambda_k} \left(\nabla\Delta\rho_{v,r,k}^{i,j} - \Delta\rho_{u_k}^{i,j} + \Delta\varphi_{r_k}^{i,j} - \Delta\varphi_{u_k}^{i,j} + \nabla\Delta N_{v,u,k}^{i,j} - \nabla\Delta I_{r,v,k}^{i,j} + \nabla\Delta T_{r,v,k}^{i,j} \right) \quad (28.8)$$

User coordinate information is contained in $\Delta\rho_{u_k}^{i,j}$, thus user coordinate information can be calculated by Eq. (28.8).

28.3 Positioning Experiments

Shown in Fig. 28.2 is the BDS VRS experimental network, composed of five stations equipped with BDS receivers. Four of the stations are reference stations, with BASE1 the master reference. The lengths of the baselines are between 11.2 and 23.4 km. The coordinates of virtual reference station were obtained through pseudorange positioning at the rover station ROVE. The receiver and antenna types were Trimble R9 and Trimble ZEPHYR GEODETIC, respectively. Observables were collected from 12:00 pm to 15:00 pm on May 27, 2014. Firstly, the baselines between reference stations were solved in relative positioning mode. After the ambiguities were fixed with Modified LAMBDA algorithm [6], the double-differenced ionosphere and troposphere residuals were separated as discussed above. The functions between errors and the relative position (virtual with respect to master reference station) were established with the method of linear combination, and then the virtual observables were generated. The precise coordinates of rover station are obtained with ambiguity-fixed solution of static baseline in post-processing mode.

Short baseline was formed with observables at rover and virtual observables at virtual reference station, and solved in static positioning mode. The ambiguities were all fixed, and the difference between the solution and precise coordinate is shown in Table 28.1. The 3D errors of static positioning with BDS VRS are 1–2 cm.

To investigate the convergence of kinematic positioning with BDS VRS, the virtual observables and rover station observables were processed in kinematic mode, and the results are shown in Fig. 28.3. It could be concluded that the convergence time for the first-time positioning is less than 10 min.

Fig. 28.2 BDS VRS experimental network

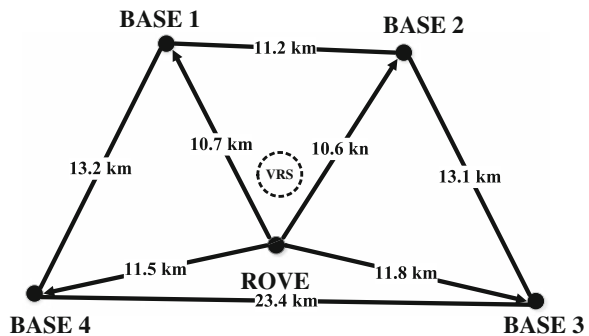


Table 28.1 The results of static positioning with BDS VRS

	ΔX (m)	ΔY (m)	ΔZ (m)	RMS (m)
Difference	-0.014	0.005	-0.003	0.015

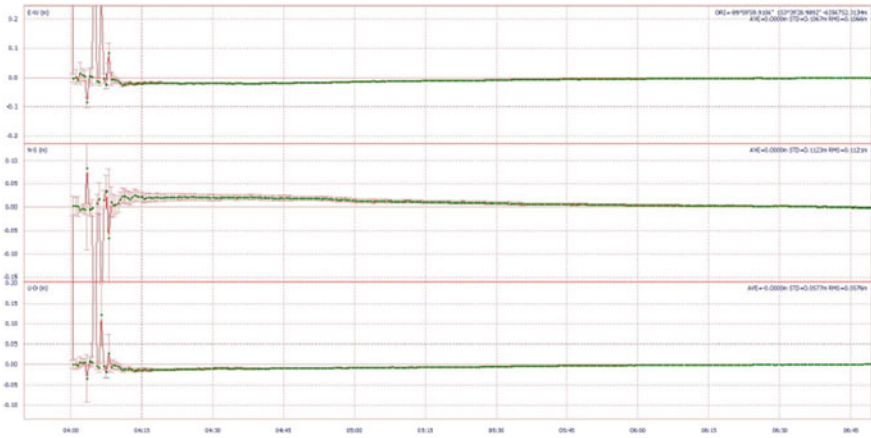


Fig. 28.3 Time series of BDS kinematic positioning results

28.4 Conclusions

In the paper, the algorithm of BDS VRS was studied, and the positioning performance with BDS VRS was verified using BDS real observables. In the CORS network with sides of 10–20 km, the positioning accuracy with BDS VRS could reach 1–2 cm, and the converging time is shorter than 15 min. The feasibility of combining BDS and VRS was also validated. With the rapid development of BDS, the constructions of infrastructures including BDS CORS network are also stepping up. The advantages, such as high precision, flexibility and low cost, will help make VRS an important trend of BDS applications.

Acknowledgments This work was sponsored by China Postdoctoral Science Foundation funded project (Grant No. 2014M562063). The authors would like to thank the editor in chief and the anonymous reviewers for their valuable comments and improvements to this manuscript.

References

1. C.C.S.N. Office (2013) BeiDou Navigation Satellite System Open Service Performance Standard (Version 1.0). In: C.S.N. Office (ed), Beijing
2. Shi C, Zhao Q, Hu Z, Liu J (2012) Precise relative positioning using real tracking data from COMPASS GEO and IGSO satellites. *GPS Solut* 17:103–119
3. Shi C, Zhao Q, Li M, Tang W, Hu Z, Lou Y, Zhang H, Niu X, Liu J (2012) Precise orbit determination of Beidou satellites with precise positioning. *Sci China Earth Sci* 55:1079–1086
4. Wei E, Chai H, Liu J (2006) On algorithm of VRS. *Geom Inf Sci Wunan Univ* 31(11):1007–1010
5. Chenggang L (2007) Network GPS/VRS system's high accuracy difference correcting information generation and release. Southwest Jiaotong University
6. Chang X-W, Yang X, Zhou T (2005) MLAMBDA: a modified LAMBDA method for integer least-squares estimation. *J Geodesy* 79:552–565

Part II
Atomic Clock Technique
and Time-Frequency System

Chapter 29

Development of Space Mini Passive Hydrogen Maser

Yonghui Xie, Pengfei Chen, Shanmin Liu, Yuxian Pei,
Shuai Tao and Chuanfu Lin

Abstract Development of space mini passive hydrogen maser has been carried out in Shanghai Astronomical Observatory. Research on key technologies such as miniaturized cavity-bulb assembly, absorption pump, magnetic state selection assembly and magnetic shield has made progress, Parameters of microwave cavity, storage bulb and state selection have been optimized to ensure system performance. Weight of physical package is less than 10 kg and atomic signal is up to 3 dB. Integration with electronic package, optimization and engineering modification are the main work at present. Less than 15 kg of total weight and frequency stability better than $1 \times 10^{-12} \tau^{-1/2}$ ($1\text{s} \leq \tau \leq 10,000\text{ s}$) is expected.

Keywords Shanghai astronomical observatory · Mini · Passive hydrogen maser · Development

29.1 Introduction

On-board atomic clocks are the key payloads for global navigation satellite system (GNSS), whose positioning accuracy is greatly affected by the clock performance. Passive hydrogen maser (PHM) has better performance in frequency stability, frequency drift and temperature coefficient, which is very helpful for autonomous navigation of the GNSS. In state of art best solution for on-board atomic clock is cooperation between PHM and rubidium clock, which benefit from its small size and low cost, so as to improve system performance and reliability. The main constraint on widely application of PHM is its weight, which is up to 23 kg for domestic product and 18 kg for aboard one. Some overseas companies have carried out research on mini PHM with lighter weight of 12 kg [1].

Y. Xie (✉) · P. Chen · S. Liu · Y. Pei · S. Tao · C. Lin
Shanghai Astronomical Observatory, 80 Nandan Road, Shanghai, China
e-mail: xyh@shao.ac.cn

Shanghai Astronomical Observatory has long history in research of hydrogen masers, and has been developing the space PHM since 2005, now stage of engineering prototypes have been achieved. On this basis, study of key technologies of mini PHM has been carried out. 14 kg of weight, 40 % of volume reduction and 20 % of power reduction are expected, while performance and life time remain the same. Attention on optimization of magnetic field and thermal design have also been paid.

29.2 Physics Package

29.2.1 Cavity-Bulb Assembly

Cavity-bulb assembly is the key component of PHM. Its parameters such as Q factor, filling factor, atomic signal gain, etc. directly affect performance of PHM [2].

$$\sigma_Y(t) = K \frac{(1 + S_0 - \alpha)^2}{Q_0 \alpha \sqrt{S_0} (1 + S_0)} \tau^{-\frac{1}{2}} \quad (29.1)$$

where S_0 is the saturation factor at resonance, Q_0 is the atomic line quality factor, α is oscillation factor, which can be expressed by [3]:

$$G = \frac{1 + S_0}{1 + S_0 - \alpha} \quad (29.2)$$

where G is atomic gain signal.

In mini PHM novel electrode microwave cavity is still adopted for its better cavity quality factor and filling factor, which has already been applied successfully in present PHM (Fig. 29.1).

Most important task in development of mini PHM is miniature of the cavity-bulb assembly, while performance should not be worsen. Parameters of cavity structure and storage bulb are optimized. Cavity coupling factor is adjusted by reduce coupling loops of microwave and CAT. Finally 40 % of volume reduction of the cavity-bulb assembly is achieved, while Q factor of the cavity is maintain as 6000, filling factor as 0.5 and atomic storage time as present PHM.

29.2.2 Vacuum System

Vacuum assembly of mini PHM is featured by long life time, large volume of hydrogen absorption and high reliability. Absorption pump is necessary to meet these requirements. Getter material of Absorption pump in present PHM is made of titanium, which has higher absorption efficiency, and corresponding higher activation temperature. More careful activation protection is required. Now new getter

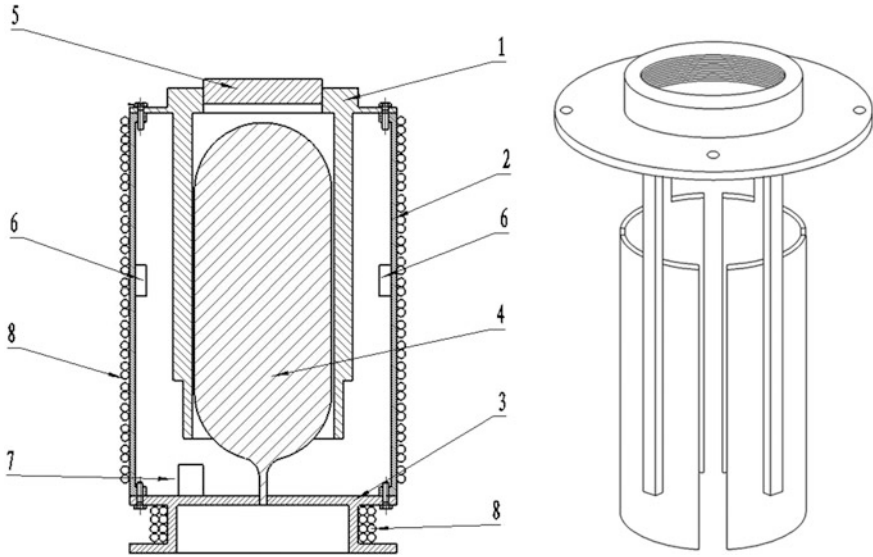


Fig. 29.1 Structure of novel kind of electrode cavity(*left*) and electrode tube(*right*). 1 Electrode tube; 2 cylinder cavity; 3 base plate; 4 storage bulb; 5 frequency adjust knob; 6 coupling loop; 7 varactor diode loop; 8 thermal control coil

material is applied in mini PHM, activation temperature is limited within 500 °C. Pump structure and application amount of getter material are also optimized, so as to ensure more conventional activation. Absorption ability of $40 \text{ bar} \times \text{L}$ of hydrogen is realized under this design.

A small ion pump is equipped in the vacuum system to eliminate residual gases. Operating load of the small ion pump is greatly reduced to less than 0.5 μA under new pump structure design, so application of a relatively small-sized pump is possible.

Vacuum environment of the cavity-bulb assemble is provided by external vacuum system to prevent turbulence of environment temperature and atmosphere. Considering that most working condition of mini PHM would be in space, which can easily provide vacuum environment, external vacuum system is only valuable in period of ground test, so temporary vacuum valve instead of vacuum pump could be applied to simplified system design and reduce the weight.

29.2.3 Magnetic Field Design and Other Features

3 layers of magnetic shield made of IJ85 are applied surrounding the cavity-bulb assembly in the physics package. Diameter and length of the shields shrink corresponding to size of the cavity. By optimization of thickness and distance of the shields, shielding factor up to 50,000 is ensured.

Supporting structure is designed and applied between magnetic shields, which are then firmly fixed. Displacement in vibration, shock or long term operation is avoided.

Magnetic iron of the ion pump would cause disturbance to homogeneity of system magnetic field. Its position and orientation are optimized in mini PHM. Private magnetic shield is also applied for the ion pump.

Cavity-bulb assembly is surrounded solenoid made of copper wire, which generate C field about 1 mGs for atomic transition. (0–1) transition can be observed if C field is shut down.

Physics package of mini PHM has other features for its small size and light weight:

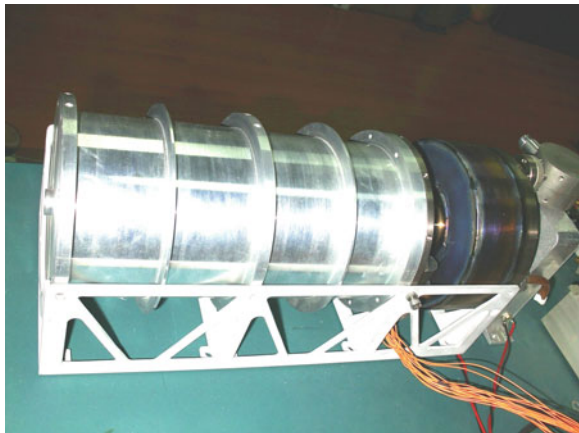
- miniaturized dissociator
- miniaturized CAT assembly
- light weight hydrogen source
- optimized state selection system

29.2.4 Performance of Physics Package

Development of physics package of mini PHM is close to complete, with atomic gain up to 3 dB and line-width of 3.5 Hz (Figs. 29.2 and 29.3).

Saturation factor S_0 at -80 dBm of microwave input is about 0.37. 3 dB of atomic gain is corresponding to G of 1.41. From Eq. (29.2), α can be induced as 0.4. In Eq. (29.2), K can be set as $1.33E-4$ from conventional experience, Q_0 is $4.06E+8$ for line-width of 3.5 Hz, then frequency stability of $9.2E-13/\tau^{-1/2}$ can be calculated, which is theoretical performance of mini PHM.

Fig. 29.2 physics package of mini PHM



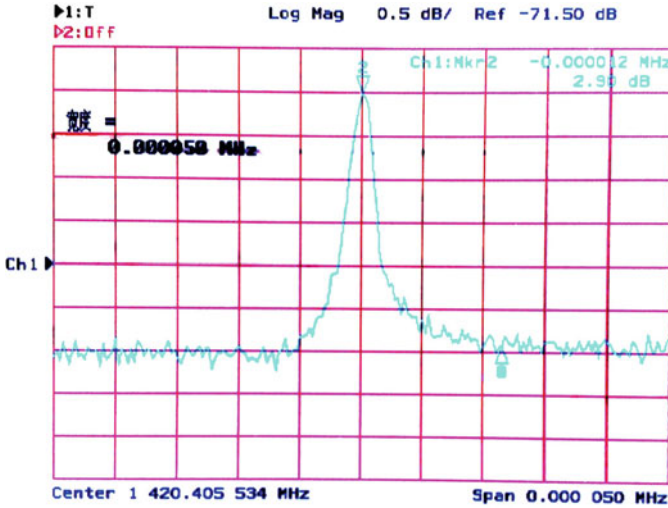


Fig. 29.3 Atomic signal at -80 dBm input

29.3 Electronics Package

Scheme of time-separated control method is still adopted in electronics package of mini PHM. There are two frequency control loops, VCXO control loop and cavity control loop, in electronics package. The former for lock of VCXO frequency and the latter for lock of cavity frequency, they are time separated to avoid mutual interference. This scheme is quite mature and has show advantage in long term frequency stability and drift. Several improvements are made for mini PHM:

- Wider CAT control range
- optimized DA bits and control algorithm
- optimized temperature control parameter
- adjusted arrangement and small size

Microwave cavity frequency would show slow drift in operation. Wider CAT control range brings smaller possibility for cavity frequency to get out of control, which improve reliability of CAT and the clock. Additional noise is suppressed and temperature stability is improved by the second and the third adjustment separately. Finally size of electronics package is compressed to fit miniature physics package, this work would be carried out after system debug and test (Fig. 29.4).

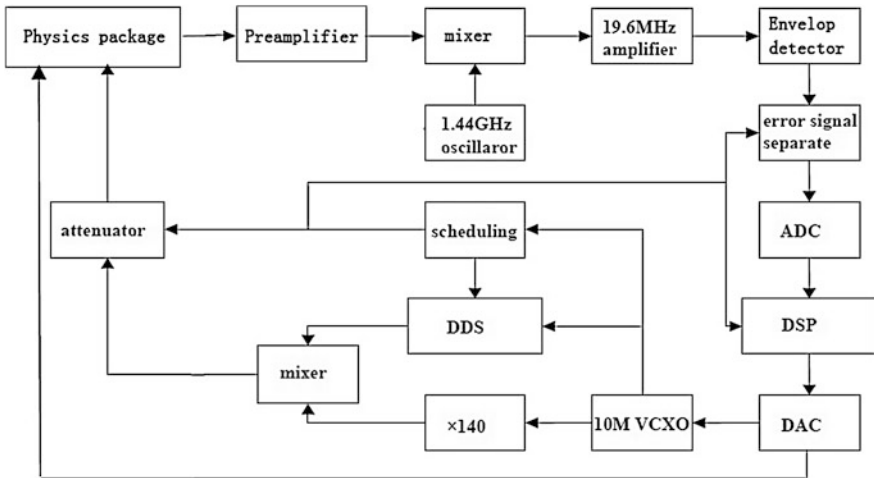


Fig. 29.4 Block-diagram of electronics package

29.4 System Status and Performance Test

Development of physics package of mini PHM is close to accomplished, with satisfied atomic gain and weight about 9 kg. Scheme of electric package is confirmed and parameters is under optimizing. A preliminary test has shown frequency stability of $1.8E-12/\tau^{-1/2}$, potential of further improvement is expected.

Weight of electronics package is expected to be limited within 5 kg after structure optimized and miniature. Total weight about 14 kg is expected for the clock. With reduction of weight and improvement of thermal structure, power consumption could be reduced by about 10 W (Fig. 29.5).

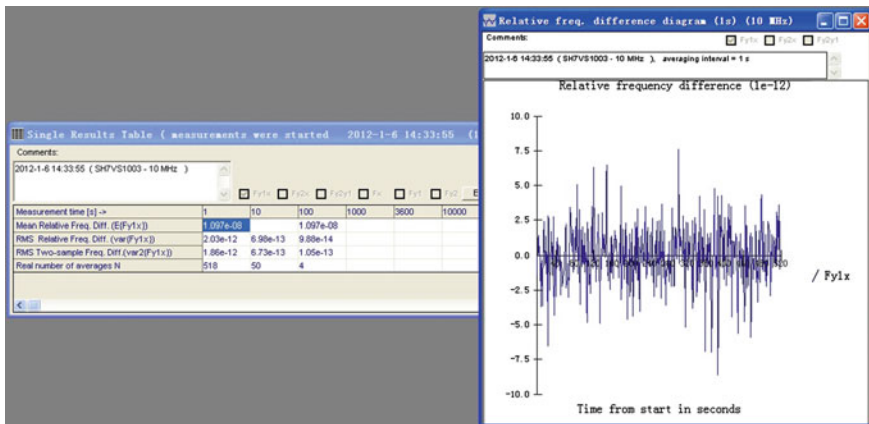


Fig. 29.5 preliminary test result of frequency stability

29.5 Conclusion

Comparing to conventional PHM, Mini PHM has 40 % of weight reduction and corresponding size and power consumption compression, while its performance, such as frequency stability and drift, is maintained, which contribute to wider application of PHM.

Study of key technologies of mini PHM has been carried out and physics package has been achieved. Prototype of mini PHM is expected in 2 years.

References

1. Belloni M, Gioia M, Beretta S et al (2011) Space mini passive hydrogen maser—a compact passive hydrogen maser for space applications, frequency control and the european frequency and time forum (FCS). In: 2011 joint conference of the IEEE international, pp 1–5
2. Vanier J, Audoin C (1989) The quantum physics of atomic frequency standards. Adam Hilger, Bristol, Chapter 6
3. Wang Q, Mosset P, Droz F (2006) Verification and optimization of the physics parameters of the onboard galileo passive hydrogen maser. In: 38th PTI meeting, Washington

Chapter 30

Development of High Frequency-Temperature Stability of OCXO for Aerospace Applications

Lei Yang, Yuhao Qin, Jian Huang, Xiaoqiang Zhang, Wei Jiang, Bing Cheng, Zhifu Feng and Fanghong Guo

Abstract With the performance upgrade of aerospace applications, the frequency source have raised higher requirement on frequency-temperature stability of OCXO (oven-controlled crystal oscillator). Based on the temperature control theory of OCXO and the character of frequency fluctuation in different temperature, we use analog temperature compensation technology in the design of OCXO. Finally, we realize a miniaturization 10 MHz OCXO for aerospace applications. This OCXO test result shows, the frequency-temperature stability is better than $\pm 5.0E-9$ with temperature range -30 to 70 °C.

Keywords OCXO · Frequency-temperature stability · Analog temperature compensation technology

30.1 Introduction

Oven-controlled crystal oscillator (OCXO) as the excellent frequency equipment, It has a very wide range of applications in electronics, communications, navigation, aerospace and other fields. Frequency-temperature stability, an important indicator of space borne OCXO, reflects the temperature control capability of OCXO products. It is mainly indicate the changes of frequency within a certain temperature range (-30 – 70 °C). Frequency temperature coefficient is greater, temperature sensitivity worse, the greater of the impact on the frequency stability. Frequency temperature coefficient is smaller, temperature sensitivity better, the smaller of the impact on the frequency stability.

This paper presents an OCXO based on analog temperature compensation network, which can further improve the frequency-temperature stability index on

L. Yang (✉) · Y. Qin · J. Huang · X. Zhang · W. Jiang · B. Cheng · Z. Feng · F. Guo
China Academy of Space Technology Xi'an, Xi'an, China
e-mail: gongdaxiaozi@163.com

the basis of the OCXO, but do not change the power and volume. At the same time, it can further improve the applicability of the product space environment, to provide protection for the onboard frequency source applications.

30.2 Status of Frequency-Temperature Stability

In the onboard OCXO development process, in order to meet the accuracy of the output frequency at different temperatures, the turn point temperature of crystal is needed to be control in high precision temperature.

At present, it is mainly carried out through the following methods to improve the frequency-temperature stability indicators,

Method 1. Double oven structure control [1].

Method 2. Digital compensation network [2, 3].

Method 3. Temperature control with the temperature following [4].

However, the above methods have the following deficiencies in space applications.

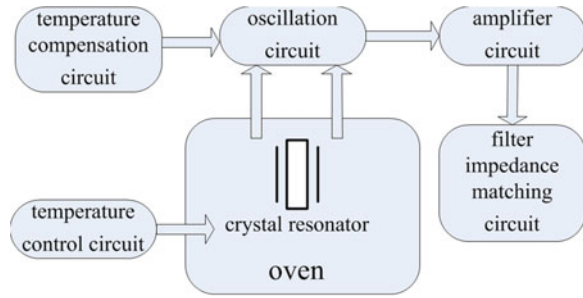
- (1) Although the method 1 can solve a space borne OCXO frequency-temperature stability index fundamentally, but the design requires a great deal of power in a wide range of start-up and steady status. Meanwhile, it needs a development materials and higher processes, which is not conducive to system application.
- (2) Method 2 use digital potentiometers, semiconductor temperature sensors, the proportion of the op amp and varactors network to complete thermostatic temperature compensated crystal oscillator (TCXO). The method can achieve better temperature compensation feature, but large volume and poor reliability is not conducive to space applications.
- (3) Method 3 uses the temperature processor to implement and maintain the temperature in a range between the internal temperature of the crystal oscillator and the external ambient temperature. It has a problem as follows, to maintain the temperature range smaller, a better frequency-temperature stability index with greater power consumption. keep the temperature range larger and smaller power consumption, indicators may not meet the requirements.

30.3 Analog Temperature Compensation OCXO Design

30.3.1 System Components

Analog temperature compensation OCXO circuit consist of the main oscillation, amplifier, filter impedance matching circuit, temperature control systems and analog temperature compensation circuit, which is presented on Fig. 30.1. Main

Fig. 30.1 Temperature compensation OCXO components



oscillation circuit is responsible for converting the DC power into AC power, crystal resonator as frequency stabilization device existing in the main oscillator circuit. The amplifier circuit amplifies the AC signal and plays a role as impedance matching and isolation. Filter impedance matching circuit is mainly used to filter out unwanted signals. Temperature control systems typically include temperature control circuit and an oven, which used to make crystal stability works at the turn point temperature to achieve high stability. The warming of analog temperature compensation circuit, according to the change of the ambient temperature (temperature control system and the outside temperature interaction), is varied within a certain range minutely adjust the output frequency, to achieve the higher frequency-temperature stability.

30.3.2 Crystal Frequency-Temperature Characteristics

Because of the characteristics in the start-up, short-term stability and long-term aging, anti-radiation and other aspects, SC-cut crystal have incomparable advantages than other cut crystals, which has been widely used in the high stability crystal oscillator. But SC-cut crystal has its drawbacks, difficult directional and more complex spectrum. Addition to the single-cut corner frequency spectrum, SC-cut crystal is used only C mode with to suppress the existing A, B, C oscillation modes and other unwanted modes.

The typical frequency temperature characteristics of 10 MHz SC-cut crystal resonator is shown in Fig. 30.2.

30.3.3 Temperature Circuit Analysis

As temperature compensation block diagram of OCXO shown in Fig. 30.1, temperature control circuit mainly serve to maintain crystal resonator temperature constant. Its working principle is to use the heat generated by the power tube heating the crystal resonator working at the turn point temperature, in which its

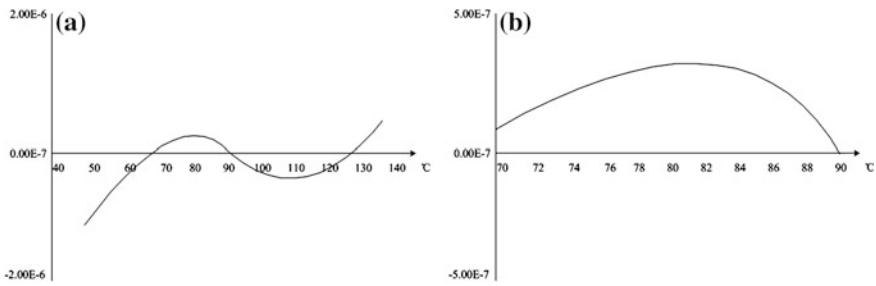


Fig. 30.2 10 MHz SC-cut crystal resonator typical frequency temperature characteristics. **a** The frequency temperature characteristics, **b** An enlarged view of the turn point temperature

frequency stability is best. The control temperature is higher than the ambient temperature, thereby significantly reducing the effects of ambient temperature variation to achieve the best performance. The characteristics of temperature control circuit will directly affect the frequency-temperature stability of the crystal oscillator [5].

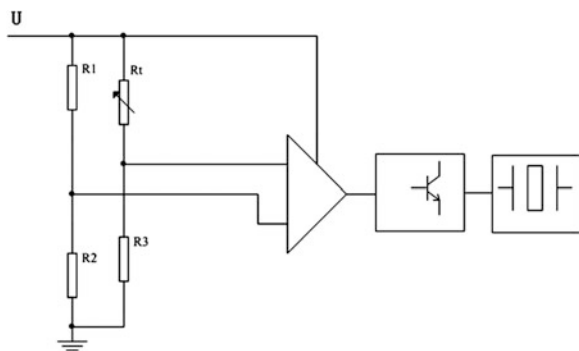
The temperature control circuit is a direct amplification continuous temperature control circuit, with the advantage of high precision, small temperature fluctuations, high reliability, continuity of work, and small components. The circuit can achieve smaller temperature control structure and be used for the high frequency-temperature stability of OCXO. The principle structure is shown in Fig. 30.3.

The bridge circuit consists of resistors R1, R2, R3 and thermistor Rt, of which R1 = R2. The input voltage U for the bridge, the bridge output voltage ΔU can be drawn from Fig. 30.3.

$$\Delta U = \frac{U}{2} \times \frac{R_t - R_3}{R_t + R_3} \tag{30.1}$$

As the temperature increasing, Rt decreases. When Rt = R3, the bridge balance, the output voltage is zero. For the selected thermistor Rt, the different values of R3 correspond to different control temperatures. R3 is determined on the actual debugging inflection crystals and the crystal has the best short-term stability, aging

Fig. 30.3 Direct amplification continuous temperature control circuit



and frequency temperature stability. But the actual temperature control system is bound to convection, radiation, part of the heat conduction contact, etc., which have some dissipation of heat, so the bridge always output a certain voltage to control the heating current equal to the heat dissipated by the oven.

30.3.4 Analog Temperature Compensation Network

The oscillation frequency of the quartz crystal is easily affected by temperature change, which can lead to failure. It is typically to use a variety of methods in engineering for temperature compensation, such as, capacitance compensation, polycrystalline complementary, digital temperature compensation and bath compensation. The accuracy of the thermal network compensation is not the highest, but it has been dominant in the temperature compensation field with small size, low power consumption, low cost and good stability.

It is difference about the temperature compensation network between temperature compensation OCXO and TCXO. The main difference between them is the different temperature ranges of compensation. The former compensated temperature range is 1–2 °C or smaller. The latter is wide, generally –40–80 °C or wider. As shown in Fig. 30.4, a common compensation circuit is consisted by a thermistor and the varactor diode with inversely proportional to the bias network of a quartz crystal. Oscillator load capacitance is approximately equal to the capacitance of the varactor diode. When the frequency of the quartz resonator changed with ambient temperature variation, a thermistor network output voltage varies with temperature and changes the junction capacitance of the varactor diode. It can remain the sum of junction capacitance of the varactor diode and the internal capacitance of the crystal resonator unchanged to achieve the purpose of compensation. The thermal network consists of the number of the thermistor and the resistor formed by series and parallel. Because the temperature change of OCXO is small, a simple form of a thermistor network can meet the requirements.

When the temperature compensation network structure is selected, the most important is the location of thermistor. Unlike OCXO, TCXO has little heat and the warming network thermistor temperature measurement is generally ambient

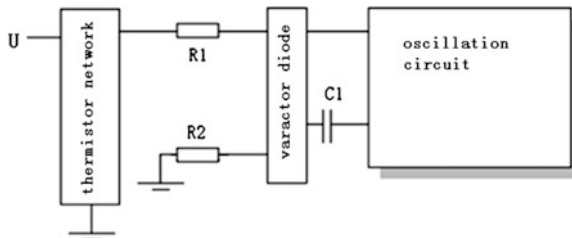
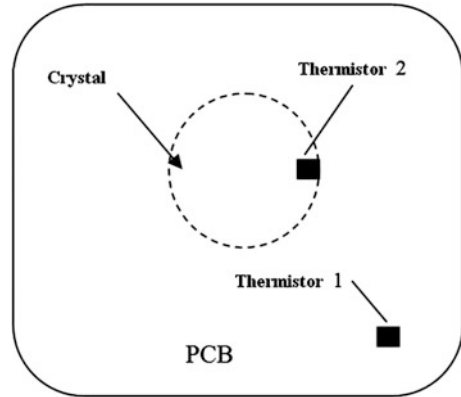


Fig. 30.4 Analog temperature compensation network structure

Fig. 30.5 Position of the thermistors



temperature. Large ranges of temperature sensing thermistor sensitivity requirements are not high. Because of OCXO needing greater heat at different ambient temperatures, temperature sensing thermistor sensitivity is raised higher requirements. In consideration of the stability of the temperature compensation network, the position of the thermistor is located as shown in Fig. 30.5. Thermistor 1 sensing the external environment can be sensitive to temperature changes. Thermistor 2 near the crystal heating zone, the temperature range is relatively small, has certain degree stability to temperature compensation network.

30.4 Test Results

In order to test the applicability of the analog temperature compensation network, two 10 MHz OCXO are developed with 3# and 5#. The 3# and 5# OCXO have respectively done with temperature compensation test. Frequency temperature compensated voltage curve and the test measured curves are shown in Figs. 30.6, 30.7 and 30.8.

As shown in Figs. 30.7 and 30.8 by the experimental data in two 10 MHz OCXO, in the temperature range of -30 to 70 °C, the frequency-temperature stability of the two products were respectively increased from before compensated $\pm 1.48\text{E}-08$ and $\pm 2.39\text{E}-08$ to compensated $\pm 4.34\text{E}-09$ and $\pm 4.75\text{E}-09$. After the test of the two compensated OCXO, the phase noise and other parameters were not affected.

Compared with the existing compensation methods to analog temperature compensation network, the method has the following advantages,

- (1) This method is applicable to any kind of improvement in upgrading the existing OCXO frequency-temperature stability indicators, while the simple structure network, high compensation precision, low power consumption and virtually no change in the original structure advantages.
- (2) The method has easy miniaturization, high reliability, and good implement and so on.

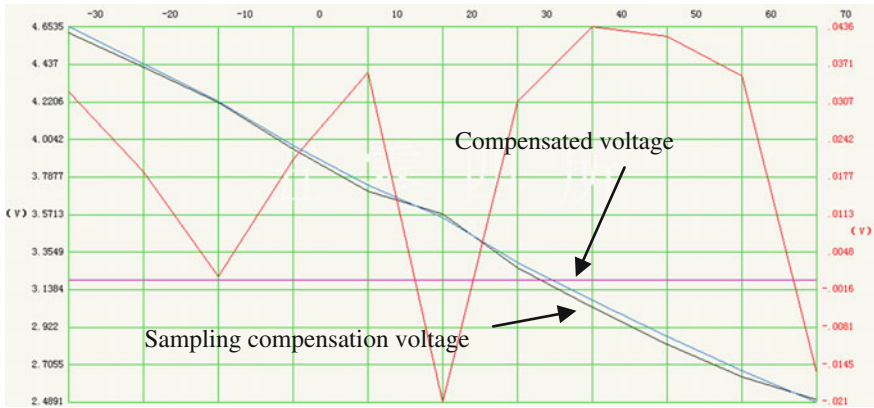


Fig. 30.6 The sampling compensation voltage curve

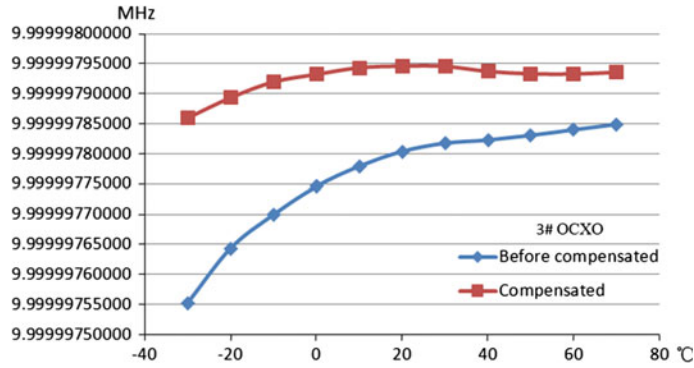


Fig. 30.7 The frequency-temperature stability of 3# OCXO

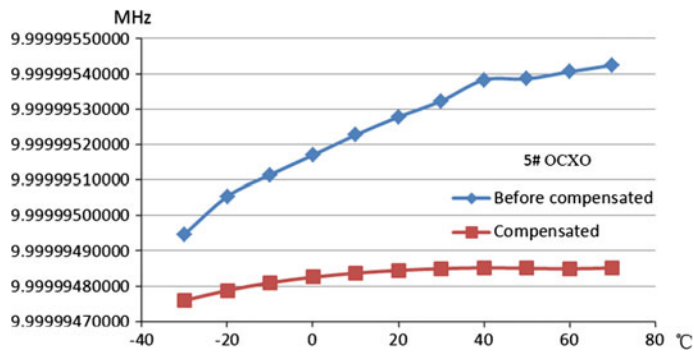


Fig. 30.8 The frequency-temperature stability of 5# OCXO

30.5 Conclusions

Based on the analysis of the OCXO technology, combined with analog temperature compensation technology, we have designed and implemented the temperature compensation OCXO. By testing two products, the frequency-temperature stability compared to the index before compensated has achieved good results. Of course, for enhancing the temperature stability OCXO frequency indicators as well as a lot of work to do, we need to continue to study on it.

Acknowledgments In the course of this product, the authors thank the help of colleagues and the author citation.

References

1. Chauvin J (2007) A new generation of very high stability BVA oscillators. In: IEEE international frequency control symposium and 21th European frequency and time forum, pp 1261–1268
2. Liu H, Wu D (2011) A high—precision mixed—signal voltage-controlled crystal oscillator with temperature. *Microelectronics* 41(6):830–833
3. Kong G, Zhao S, Wang J (2011) High precision temperature compensation crystal oscillator. In: 2011 China time and frequency symposium, pp 227–230
4. Zhen JING (1975) High stability crystal oscillator. National Defense Industry Press, Beijing, p 90
5. Qin Y, Jiang W, He Y et al (2011) Development of 10 MHz high stability mini-OCXO. In: 2011 China time and frequency symposium, pp 188–189

Chapter 31

Realization and Influencing Factors

Analysis of ACES (Atomic Clock Ensemble in Space) Management

Jun Xie, Yunfeng Sun, Yongsheng Qu, Dong He and Ming Zhao

Abstract In navigation satellites constellation comprised of multi-satellite, introducing an atomic clock ensemble in space is beneficial for improvement of its measurement accuracy and ability of autonomous operation. This paper summarizes the methods adopted currently at home and abroad towards atomic clock ensemble management in ground or in space and then proposes an initial design scheme of ACES management on a single satellite or among different satellite clocks within constellation which is based on the designing conception of atomic clock ensemble in navigation satellite constellation. Considering generation and characterize measurement of clock ensemble signal, anomaly detection and processing of atomic clock, this paper gives analysis and recommendation as for the key technologies and factors affecting ACES management, which will provide technical basis in designing a continuous, reliable, stable, accurate ACES running system for BEIDOU global navigation satellite system project.

Keywords ACES · ACES management · Influencing factors

31.1 Introduction

Atomic clock ensemble management is adopted widely by time-keeping laboratory around the world. Generally in local laboratory or several laboratories an ensemble of precision clocks used for time-keeping is called one clock ensemble. During clock ensemble operating, acquiring phase differences between the clock signals

J. Xie
China Academy of Space Technology (CAST), Beijing 100094, China
e-mail: XiebookK@sina.com

Y. Sun (✉) · Y. Qu · D. He · M. Zhao
Xi'an Branch CAST, Xi'an 710100, China
e-mail: sxsyf28@126.com

within clock ensemble with high accuracy phase measurement technology, generating a paper time with a proper algorithm and those phase difference results, finally outputting a final signal after using the clock difference to micro-change an actual clock.

Clock ensemble management consists of an ensemble of atomic clocks, clock differences measurement system, clock difference data operation, phase micro-step and physical signal output through the master reference clock. Clock ensemble management is in essential a feedback control system of which the important thing is reasonable determination for control parameter and acquisition of clock measurement results. Figure 31.1 gives schematic diagram of realization of clock ensemble management.

The application of clock ensemble management on ground is very familiar. The generation of time scale, such as TAI, UTC (k) etc. is based on clock ensemble management. With advanced clocks produced the performance of the clock ensemble is continuously developing. The aim of clock ensemble management is to improve clock noise and maintain higher accuracy, stability and reliability.

With continuously developing of the global navigation satellite system, the number of atomic clock used in space is increasing, if one can acquire the performance parameters of master each atomic clock on-board each navigation satellite, it is beneficial for improvement of system measurement accuracy and ability of autonomous operation.

At the same time, each navigation satellite itself is equipped with three to four sets of atomic clocks, which also involves the ACES management. If one does not take atomic clock ensemble management on-board navigation satellite it would degrade the MTR 10.23 MHz signal stability and availability and at last system performance while the master clock was in anomaly such as large frequency drift, frequency jumping or phase jumping. For example, on January 1, 2004 an anomaly of atomic clock on the GPS SVN 23 occurred. Because of no timely detection, the frequency of the atomic clock had been drifting for nearly three hours and in some places the GPS positioning error even reached 40 km. If one has taken ACES management on-board

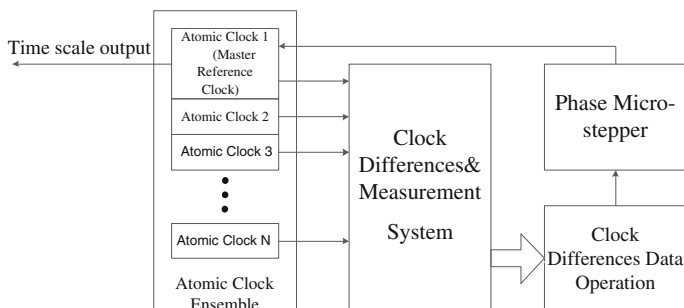


Fig. 31.1 Schematic diagram of realization of clock ensemble management

navigation satellite the ACES management can predict and correct these clock anomalies to elevate satellite system reliability and ability of autonomous operation. Moreover the ACES management can improve satellite clock long term frequency stability and system measurement accuracy and performance.

31.2 Overview of Atomic Clock Ensemble Management at Home and Abroad

31.2.1 Atomic Clock Ensemble Management on Ground

As previously mentioned, the applications of atomic clock ensemble management on ground are quite common, especially in terms of timekeeping. “PHM clock + cesium clock” is assembled popularly in timekeeping laboratory around the world which combines the advantages of each clock, the short-term frequency stability for PHM clock, the long-term frequency stability and accuracy to output the signal with optimal stability and accuracy. Such as the US Naval Observatory (UNSO) clock ensemble consists of nineteen PHM clocks and twenty five commercial cesium clocks (also added four subsequent rubidium fountain clocks) [1], Chinese Academy of Sciences National Time Service Center (NTSC) clock ensemble consists of four PHM clocks and nineteen commercial cesium clocks [2].

In terms of time scale algorithms there are three main algorithms: ALGOS algorithm, AT1 algorithm, Kalman filter algorithm. Bureau International des Poids et Mesures (BIPM) Time Division, Chinese National Institute of Metrology (NIM) and the Chinese Academy of Sciences National Time Service Center (NTSC) use weighted average algorithm ALGOS for timekeeping service; American National Institute of Standards and Technology (NIST) uses a weighted average of AT1 algorithm for research work such as atomic time scale computing. American GPS navigation satellite system uses the Kalman filter algorithm to generate GPS time [3]. Among these three algorithms, the first two are the weighted average algorithms; the third algorithm utilizes an estimation theory to make the optimal estimate of the clock differences in clock ensemble.

ALGOS is a delayed algorithm which is mainly used in the TAI imputed and the requirement of TAI is to have very reliable and excellent long-term frequency stability. Delayed algorithm is obviously not suitable for navigation satellite systems (longer lag time causes the satellite clock error accumulation, and affects the performance of the navigation system).

AT1 algorithm and Kalman algorithm are real-time algorithms. AT1 algorithm is to assign a weight to each clock and unable to suppress all noise, so it is more suitable for research on the atomic clock noise characteristics. Kalman filter algorithm can be used to build a variety of atomic clock noise modules and suppress these noises, which is adapted to generate and maintain the system time such as the navigation system. Thus, the selection of time scale algorithm has no uniform or the best solution which is determined by the requirements for usage.

31.2.2 Atomic Cock Ensemble in Space (ACES) Management

Currently there is fewer publicly report about ACES management. Soualle from Germany Astrium GmbH Corporation proposed the concept of using atomic clock ensemble management on the Galileo navigation satellite [4]. The clock ensemble onboard Galileo satellite hosts two passive PHM clocks and two rubidium atomic clocks. A 10.23 MHz reference signal is generated by the Clock Monitoring Control Unit (CMCU) on Galileo satellite. There is no clock ensemble management in early CMCU. Two clocks selected from four onboard clocks via a matrix switcher are powered and their output signals are used as the reference signals of two synchronizers in CMCU which generate respectively a 10.23 MHz time reference signals. One of these two 10.23 MHz signals is active master time reference and the other is in hot redundancy [5].

It can be seen that when master atomic clock has anomalies such as abnormal frequency drift, frequency or phase abnormal jump, the early design of CMCU will not effectively prevent 10.23 MHz signal from getting impacted. (Although having a smooth switch, before switch the fault has been reflected in the 10.23 MHz signal). For this reason, Soualle, Zenzinger have respectively given a proposal of making atomic clock ensemble in space management onboard CMCU [4, 5], of which the main idea is: all of the four atomic clocks onboard Galileo satellite together generate 10.23 MHz signal, with a Kalman filter algorithm achieving the estimation of clock differences among these clocks so that the out-put signal of clock ensemble obtain optimal time and frequency characteristic; excluding clock anomaly in advance through adopting the clock anomaly (frequency or phase abnormal jump) detection algorithm to ensure that the output signal of clock ensemble is not affected. Words in all headings (even run-in headings) begin with a capital letter. Articles, conjunctions and prepositions are the only words which should begin with a lower case letter.

31.3 Proposed Scheme and Key Technologies of ACES Management

31.3.1 Proposed Scheme of ACES Management

Currently the number of atomic clock used in space is increasing, most of which are applied to the navigation satellite system. ACES management given herein is for the purposes of navigation satellites, but it also applies to other situations of ACES management.

Within navigation satellite systems ACES management should include two aspects: ACES management on a single satellite, ACES management between each satellite clock within constellation. Both are managed for multiple atomic clocks, through processing clock difference data to obtain the final clock ensemble time

(paper time), and output a physical signal. The difference is a measure of the clock difference and the different ways to get the data. For the former clock difference is acquired in a single satellite through high-precision measurement system. For the latter clock difference is acquired via precision measurement and data transmission way of the inter-satellite among different satellite clocks.

ACES management on a single satellite is taken to ensure the continuity and reliability of satellite clock output signal through managing the atomic clocks equipped onboard navigation satellite. Based on the currently realized scheme onboard BEIDOU navigation satellites through which the 10.23 MHz reference signal is generated from rubidium atomic clocks the proposed scheme of ACES management is given, as shown in Fig. 31.2. The basic principle for scheme to achieve is: four 230 kHz signals are generated respectively from four atomic clocks with the corresponding DDS; and then mix with 10.23 MHz of a voltage-controlled crystal Oscillator to generate 10 MHz signal; four phase meters are employed to achieve phase measurement between any two 10 MHz signals, being imported the clock ensemble signal generation module in FPGA/ASIC the results of phase measurement are used as initial data to produce clock ensemble output. Using Kalman filter algorithm and anomaly detection algorithm for clock difference data processing, the optimal clock differences output from the signal generation module and then enter into phase detection (PD) module along with a multiplexer output signal. Passing through D/A converter and filter the output of the PD module enters into the VCXO control terminal. Then the final 10.23 MHz reference signal onboard satellites output from the VCXO.

In principle the hardware used for scheme should change as small as possible, and the increasing parts should make use of the mature circuit module. As for the requirements of clock ensemble management the corresponding function module should be introduced in the FPGA. For the four phase meter modules in Fig. 31.2, a

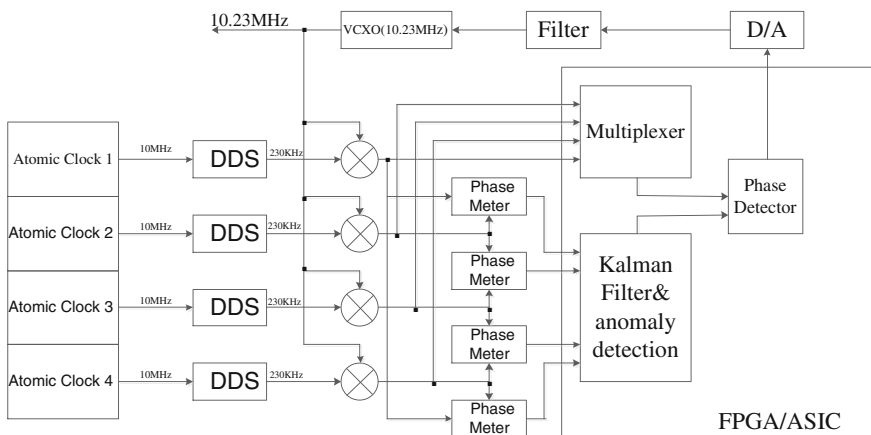


Fig. 31.2 Block diagram of generation of ACES management (single satellite)

4-channel phase comparison meter can be substituted for them. If the FPGA resources permitting, the four DDS modules can also be implemented in FPGA to save hardware resources and space.

Clock ensemble management between each satellite clock within constellation is mainly to achieve production and maintenance the system time in the autonomous operation mode. The proposed scheme to achieve the following: Based on precision measurement capabilities and data transmission functions of the inter-satellite measurements and communication transmission links, GNSS achieves pseudo-range measurements and time synchronization, gets the clock difference between the satellite clocks. All the clock differences are transmitted to the master satellite. The clock difference of the clocks on master satellite and other satellites are measured by precise time measurement by which the satellite system time (paper time) is obtained through selecting appropriate real-time algorithm such as Kalman filter algorithm. Calculation the deviation of the satellite system time and ground system time is used to adjust the satellite system time to ensure synchronization between the two. When the autonomous operation mode of the constellation is switched to the ground control operation mode, the system can guarantee the constellation time smooth switch. At the same time fitting clock difference for each satellite clock data through the satellite system time is taken to achieve their respective parameter estimation of clock error model. The parameter estimates are sent via inter-satellite link to each satellite and then as a part of navigation message issued to the user. A block diagram of generation of ACES management within a satellite constellation is shown in Fig. 31.3.

In the algorithm selection for atomic clock in space management, the real-time requirement is mainly considered, and Kalman filter algorithm nicely meets this requirement. At the same time the GPS system in its full operation since the 1990s has been adopting the Kalman filter algorithm to achieve and maintain the generation of GPS time, from which it can be concluded that the algorithm should be stable and reliable. So it is the Kalman filter algorithm that is selected to achieve the generation of clock ensemble signal. As for the anomaly detection algorithm, it is suggested that according to the properties of the atomic clock onboard satellites a

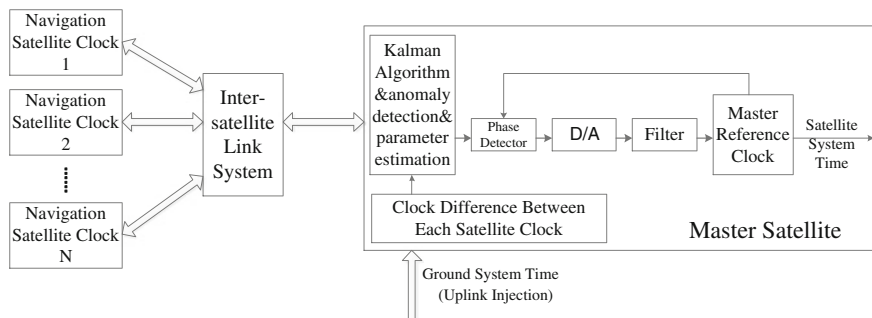


Fig. 31.3 Block diagram of generation of ACES management (multi-satellite)

lot of simulation analysis and experimental validation should be made before selection, which is one key-point of the achievement for the clock ensemble management scheme. There will be detailed description in the following sections.

31.3.2 Key Technologies of ACES Management

For ACES management scheme to achieve, the following key technologies needed to refine and research:

The high-precision phase measurement and phase micro-step technology. Due to the phase measurement results will be used as raw data for the clock ensemble signal generation the accuracy of the phase measurement of the atomic clocks must meet performance requirements. PHM clocks will be equipped with the sub-sequent BEIDOU navigation satellites and its frequency stability specifies $1e-12\tau^{-1/2}$, for which a single phase measurement resolution will require within 1 ps (1 s measurement time interval required) at least. At present, more general phase measurement is achieved by the dual mixer time difference measurement method, so one can consider modifying dual mixer time difference method to achieve high precision phase measurement requirements. Of course, one can also consider other better measurement techniques in order to meet the technical requirement. On the basis of high-precision phase measurement achieved the micro-step phase technology will be required to implement fine-tune control of VCXO in order to ensure the adjustment should not lead to visible discontinuous in frequency and phase for output signal. The technology depends on the high-resolution D/A convertor. However, due to the satellite engineering implementation constraints, resolution of the device cannot be too high and it should be considered an effective method for the improved resolution.

Selection and implementation of algorithms in ACES management is another key technology. A Kalman filter algorithm and clock anomaly detection algorithm are more critical. For the former, firstly the selection of state equation for clock error state model should be able to accurately reflect the onboard clock state. In order to achieve the algorithm simply prophase algorithm can be used two-state model (such as clock offset, clock speed) for the design, follow-up needs to consider three state model (such as adding the clock drift rate state) to more accurately reflect the clock state. Secondly, since the systems constructed by Kalman filter algorithm exist unobserved factors [3], which led to filter divergence, i.e., the estimation results error will be increase endless with time. So during applying the algorithm, one should consider covariance matrix deduced by the algorithm to be divided into unobservable parts and observable parts and to remove the unobservable parts at a fixed time interval. Finally, in the case of clock remove or join, the algorithm should be able to ensure clock ensemble output signal to be continuous in frequency and phase. For clock anomaly detection algorithm, an appropriate algorithm is selected according to the characteristics of the clock anomaly. For a continuous clock frequency or/and phase abnormal jump, due to

randomness it is possible to consider the use of the algorithm with statistical properties; for abnormal frequency drift rate caused by a single frequency or/and phase abnormal jump, one should perform a variety of algorithms to compare in detection performance and then select the appropriate algorithm. Finally, one should consider a compromise between the selection and implementation of algorithms and actual hardware resources.

In addition to these two key technologies, there is need to focus on the following technologies: satellite atomic clock configurations, considering the performance requirements for atomic clock ensemble and realization cost, one should make optimal configuration for the type and quantity of atomic clocks; Transmission delay error among different satellite clocks, constraints must be made for within inter-satellite link signal transmission path delay consistency and stability, also periodic calibration of delay must be provided, which will have an impact on the inter-satellite satellite clock error measurement accuracy.

31.4 Influencing Factors and Control Measures of ACES Management

For ACES normal running the following factors need to be considered and measures to be controlled:

Currently onboard the BEIDOU-II Navigation Satellite atomic clock ensemble consist of “4 rubidium atomic clocks” or “2 PHM clocks + 2 rubidium atomic clocks”. Although the clock ensemble output signal maintain its optimal performance before being actualized, in the actualization process one should consider selecting the optimal performance clock of the clock ensemble as the primary reference clock as possible. For “4 rubidium atomic clocks” clock ensemble, due to similar performance among each clock, one can select either as primary reference clock. For the “2 PHM clocks + 2 rubidium atomic clocks” clock ensemble, it should be given priority in the selection between the two PHM clocks a primary reference clock.

The required power of clock ensemble increases relative to that of current two clocks. It can be estimated from the power requirements of atomic clocks onboard the BEIDOU navigation satellite that the power consumed has increased about 140–170 W, so the satellite platform must provide adequate and stable power supply for clock ensemble.

For reliable operation of ACES management algorithms firstly one must choose a reliable algorithm. As described above, Kalman filter algorithm proves to be reliable because of a long stable running of American GPS system adopting the algorithm. Clock anomaly detection algorithm should be applied with detection theory, and according to the atomic clock characteristics and system requirements for detecting in time, one can determine the anomaly detection thresholds and detection methods to identify the requirements of the anomaly detection probability

and false alarm probability. A variety of available algorithms should be fully validated and then one can select the optimal algorithm from them. Secondly, one should make a compromise with hardware resources and necessary resources for algorithms. On the basis of the full realization of the early cases performance, it can be considered that a key device used for algorithms achievement ASIC would substitute for FPGA to ensure the project implementation.

Telemetry information of clock ensemble running should accurately reflect fully operational state. When the satellite clock fails, an accurate operation must be given according to the telemetry information, such as isolation from the clock ensemble; if confirming clock back to normal in time to rejoin clock ensemble. It should be avoided interference from Ground control with the normal operation of the clock ensemble. According to the telemetry information one can give accurate assessment of in-orbit performance of each clock from the clock ensemble and provide a reference for the subsequent engineering development of atomic clock.

31.5 Conclusions

Based on the establishment of ACES for navigation satellite, one can predict clock anomaly and take measures at the maximum extent possible to avoid the impact of clock anomaly on the 10.23 MHz reference signal and to guarantee the availability of navigation satellites. It is beneficial for improvement of navigation satellite system measurement accuracy and ability of autonomous operation. At the same time the ACES for navigation satellite can also improve long-term frequency stability of the satellite clocks and the measurement accuracy of the navigation satellite system. For the establishment of ACES on navigation satellites this paper presents a preliminary scheme and gives analysis and recommendation as for the key technologies and factors affecting ACES management, which aimed to provide reference for the project implementation of ACES on the following BEIDOU navigation satellite system.

References

1. <http://www.usno.navy.mil/unso/time/master-clock>. Accessed 19 Dec 2014
2. Shaowu D (2007) Study on several important technical issues in time-keeping. A dissertation for the degree of Doctor
3. Wright JR (2007) GPS composite clock analysis. In: IEEE frequency control symposium, pp 523–528
4. Soualle F et al (2010) New concept for the on-board master clock generation unit for future galileo satellites. In: ION international technical meeting (ITM)
5. Zenzinger A et al (2012) Failure detection and correction for clock ensemble in space. In: IEEE conference publication, pp 1–8

Chapter 32

Research of Navigation Constellation Independent Punctuality Based on Different Configuration of Satellite Clocks

Yang Yang, Fan Jian jun, Hong Yuan, Jin Shu xin and Yang Yu fei

Abstract The time-synchronization has no ground time criterion in the model of autonomous time-synchronization, time precision keep on the navigation satellite itself. When can make GEO satellite as time norm, and use satellite-satellite chain to give other satellite. This paper design the navigation constellation of the new satellite-navigation-system, and configure the satellite clocks, simulate all kinds of constellation time-synchronization precision, educe the satellite clock configure scheme in the new navigation system. Due to the introduction of high-precision optical clocks, the overall time of the entire navigation constellations drift is smaller, the accuracy is much higher than the link between self-ordinary; and because the presence of high-precision clock, the time synchronization algorithm's requirements are lower, general EKF can meet the requirements.

Keywords Navigation system · Time synchronization · Satellite clock

32.1 Introduction

Atomic clock is the time reference ranging navigation system; it is the core part as the payload in the satellite navigation system. Its performance directly determines the user's navigation timing accuracy. In the autonomous time synchronization mode, time synchronization is lacking of ground-based reference, time accuracy depends on satellite navigation autonomously maintained.

Y. Yang (✉) · F.J. jun · H. Yuan · J.S. xin · Y.Y. fei
Beijing Satellite Navigation Center, Beijing, China
e-mail: 112632507@qq.com

32.2 Application of Atomic Clocks Investigation Status Quo

In the application of navigation system atomic clocks, the magnetic state of the cesium clock and optically pumped cesium clocks have been used. Magnetic state cesium clock has accurate frequency reference, but it is rather bulky. The HP507A clock represented optically pumped cesium clock, it has small size, light weight, long-term stability, high accuracy and long life. Rubidium atomic clock can meet the requirements during the life of the satellite; it has been verified in the GPS-2R satellite. And rubidium clock has low price, small size, light weight, good short-term stability; it has been widely used in satellite navigation systems. Hydrogen clock has been divided into active and passive two types. GPS uses cesium and rubidium clock frequency as the time standard, GLONASS mainly use cesium clocks, European “Galileo” (GALILEO) satellite navigation system is mainly used rubidium clocks and hydrogen masers [1].

In order to meet the establishment of satellite navigation system “modernization” and third-generation satellite systems, we improve the traditional space borne Rb and Cs atomic clock, while the use of new physical principles and techniques of atomic clocks onboard is developed very rapid, it has made good progress, mainly in two aspects of laser cooling and trapping atoms technology [2]. States are committed to develop higher precision, smaller atomic clocks [3–6].

With the research of new atomic clocks, such as the high-performance optical atomic clocks, and applied the research results in satellite navigation system [7], the performance of satellite navigation systems will be increased dramatically. This paper presents a hybrid navigation constellation, by setting different atomic clocks in different satellites, analysis of constellation autonomous navigation methods in different satellite configurations.

32.3 Satellite Autonomous Navigation Algorithms

For autonomous navigation constellation time synchronization, using distributed navigation algorithm is obviously more reasonable. Because the distributed computing process use parallel processing to reduce the amount of computation of single stars, it can improve computational efficiency while distribute the computing process only with their related observational data. Avoided a lot of inter-satellite transmission, it provides a more flexible and stable structure, making the calculation of the satellite constellation, and it is conducive to the expansion of the constellation.

In a distributed architecture, each navigation satellites only need to accurately estimate their own state. However, due to the observed values ranging between inter-satellite have associate status, they need each other satellite clocks’ face time

and covariance matrix, since each satellite is estimated that only its own state. EKF equations for distributed process can be represented as

$$Z_{i,k} = h_i(X_k) + V_{i,k} \quad (32.1)$$

$$Z_{i,k} = H_{i,k}X_{i,k} + V_{i,k} + v_{i,k} \quad (32.2)$$

Measurement update process

$$K_{i,k} = P_{i,k/k-1}H_{i,k}^T(H_{i,k}^TP_{i,k/k-1}H_{i,k} + R_{i,k})^{-1} \quad (32.3)$$

$$\hat{X}_{i,k} = \hat{X}_{i,k-1} + K_{i,k}[Z_{i,k} - H\hat{X}_{i,k-1}] \quad (32.4)$$

$$P_{i,k} = (I - K_{i,k}H_{i,k})P_{i,k/k-1} \quad (32.5)$$

32.4 Satellite Clock Modeling

The atomic clocks which used in satellite navigation system have been changed into the system and random variation. Atomic clock difference $x(t)$ can be expressed as the difference between the instantaneous clock time and the same time the standard, clock error is modelling with quadratic polynomial [8].

$$x(t) = a_0 + a_1(t - t_0) + \frac{1}{2}a_2(t - t_0)^2 + \varepsilon_x(t) \quad (32.6)$$

a_0 is the initial phase of the clock (time) bias, a_1 is the initial atomic clock frequency deviation, a_2 is atomic linear frequency drift rate, $\varepsilon_x(t)$ is the time deviation which affected by clock noise caused by random variation component, t_0 is reference time.

Space borne atomic clock system model can be represented as

$$x_k = \phi_{k-1}x_{k-1} + W_{k-1} \quad (32.7)$$

where

$$x_k = \begin{bmatrix} a_{0,k} \\ a_{1,k} \\ a_{2,k} \end{bmatrix}, \phi_{k-1,k} = \begin{bmatrix} 1 & \tau & \frac{\tau^2}{2} \\ 0 & 1 & \tau \\ 0 & 0 & 1 \end{bmatrix}, W_{k-1} = \begin{bmatrix} w_{\phi,k-1} \\ w_{f,k-1} \\ w_{a,k-1} \end{bmatrix}$$

$a_{0,k}$, $a_{1,k}$, $a_{2,k}$ represent t_k epochs satellite clock bias, frequency deviation and frequency drift rate; τ is filtering cycle; $w_{\phi,k-1}$, $w_{f,k-1}$, $w_{a,k-1}$ represent t_{k-1} epochs satellite clock noise, frequency noise and frequency drift noise.

32.5 Simulation and Results Analysis

32.5.1 Simulation of Atomic and Optical Clocks Error

The most common expression of frequency stability is Allan variance, In this paper, the satellite clock error is calculated with Allan variance model [9] (Table 32.1).

After each of the above types of noise generated, atomic clock error is calculated by the following formula

$$x_i = x_{i-1} + \tau(y_i^{WP} + y_i^{WF} + y_i^{FF} + y_i^{RW}) \quad i = 1, 2, \dots, N. \quad (32.8)$$

32.5.2 Simulation of Ultra-stable Oscillator Data

There is no good mathematical model for oscillator. In this paper, allan variance provided by ACES [10] (Table 32.2).

Because of the PM white noise, FM white noise, FM flicker noise, FM random walk noise and phase white noise are independent of each other, its difference can be expressed as the sum of the variance, it is represented as

$$\sigma_y^2(\tau) = \sigma_{-2}^2(\tau) + \sigma_{-1}^2(\tau) + \sigma_0^2(\tau) + \sigma_1^2(\tau) + \sigma_2^2(\tau)$$

You can utilize these five separate noise respective characteristics, separated these noises and Inversion error sequence, the formula shows in Eq. (32.8).

Table 32.1 Typical values of the atomic frequency standard Allan variance

Noise types	PM white noise	FM white noise	FM flicker noise	FM random walk noise
Rubidium clock	–	$(1.5 \cdot 10^{-12} \tau^{-1/2})^2$	$(5 \cdot 10^{-14} \tau^0)^2$	$[(1 \cdot 10^{-13})^2 / \text{day}] \tau$
Cesium clock	–	$(8.5 \cdot 10^{-12} \tau^{-1/2})^2$	$(2 \cdot 10^{-14} \tau^0)^2$	–
Hydrogen clock	$(1.5 \cdot 10^{-13} \tau^{-1})^2$	$(4 \cdot 10^{-14} \tau^{-1/2})^2$	$(2 \cdot 10^{-15} \tau^0)^2$	$[(3 \cdot 10^{-13})^2 / \text{day}] \tau$
Optical clock	–	$(1 \cdot 10^{-14} \tau^{-1/2})^2$	$(1 \cdot 10^{-16} \tau^0)^2$	–

Table 32.2 ACES ultra-stable crystal oscillator (USO) allan variance with interval

Measurement interval(s)	1	2	4	10	20
Sigma	1.49e-13	1.26e-13	1.00e-13	8.82e-14	8.88e-14
Measurement interval(s)	40	100	200	400	1000
Sigma	9.67e-14	1.23e-13	1.30e-13	1.20e-13	2.25e-13

32.5.3 Simulation Conditions

In this paper, a hybrid navigation constellation has been simulated by MEO\GEO\IGSO, satellites are composed from 1–35, 1–27 for MEO satellites, 28–30 for IGSO satellites, 31–35 for GEO satellites. ISLs (MEO-MEO/IGSO) use Ka link, Considering the total measurement error is 0.1 m (1σ). Satellite clock error σ_{X_0} and covariance P_{X_0} design for hydrogen maser, rubidium and cesium clock, optical clock, ultra-stable oscillator respectively as

$$\begin{aligned} \sigma_{X_0, H} &= [0.2 \times 10^{-9} \quad 0 \quad 0] & P_{X_0, H} &= \text{diag}[0.04 \times 10^{-18} \quad 0 \quad 0] \\ \sigma_{X_0, Rb, Cs} &= [2 \times 10^{-9} \quad 0 \quad 0] & P_{X_0, Rb, Cs} &= \text{diag}[0.04 \times 10^{-18} \quad 0 \quad 0] \\ \sigma_{X_0, O} &= [0.1 \times 10^{-9} \quad 0 \quad 0] & P_{X_0, O} &= \text{diag}[0.01 \times 10^{-18} \quad 0 \quad 0] \\ \sigma_{X_0, uso} &= [1 \times 10^{-9} \quad 0 \quad 0] & P_{X_0, uso} &= \text{diag}[1 \times 10^{-18} \quad 0 \quad 0] \end{aligned}$$

Simulation time is 60 days, simulation program shows in Table 32.3.

32.5.4 Simulation Results Analysis

Option I:

Figures 32.1 and 32.2 are the optical clock + ultra-stable oscillator time synchronization statistics chart. Non-GEO satellites use ultra-stable oscillator, because ultra-stable’s short-term stability of the crystal is better, but less stable long, when there is no two-way time synchronization, the entire constellation average time of 60 days maximum drift is about 3500 ns (Fig. 32.2). But when they join Ka-band satellite link, the use of high-precision optical clocks on the GEO satellite, constellation average time (constellation whole time drift) does not exceed the maximum drift 0.3 ns (Fig. 32.1).

Option II:

Figures 32.3 and 32.4 are optical clock + rubidium and cesium clocks time synchronization statistics charts. Non-GEO satellites use atomic clocks (rubidium, cesium clocks), due to atomic clocks (rubidium, cesium clocks) long stable are better than ultra-stable oscillator, when there is no two-way time synchronization,

Table 32.3 Simulation algorithm design of independent time synchronization

Program	GEO satellites	Non-GEO satellites	Measurement	Algorithm
Option I	Optical	Ultra-stable oscillator	Ka	Distributed EKF
Option II	Optical	Rubidium/cesium clock	Ka	Distributed EKF
Option III	Hydrogen	Ultra-stable oscillator	Ka	Distributed EKF
Option IV	Hydrogen	Rubidium/cesium clock	Ka	Distributed EKF

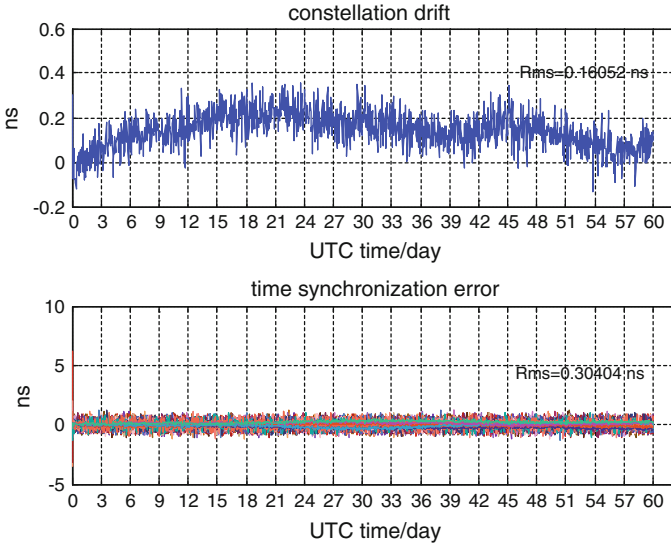


Fig. 32.1 Option I constellation drift and satellite time synchronization error

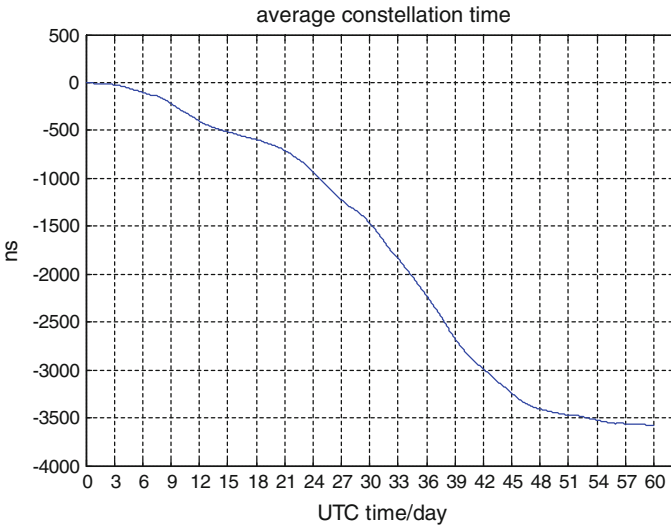


Fig. 32.2 Option I constellation time drift error

the entire constellation of 60 days maximum drift about for 450 ns (Fig. 32.4). When added the Ka-band inter-satellite link, the use of high-precision optical clocks on the GEO satellite constellation average time (constellation whole time drift) does not exceed the maximum drift 0.3 ns (Fig. 32.3).

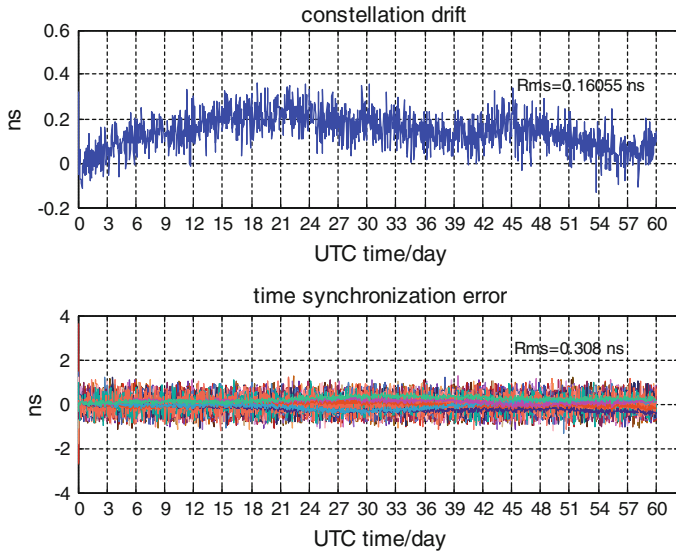


Fig. 32.3 Option II constellation drift and satellite time synchronization error

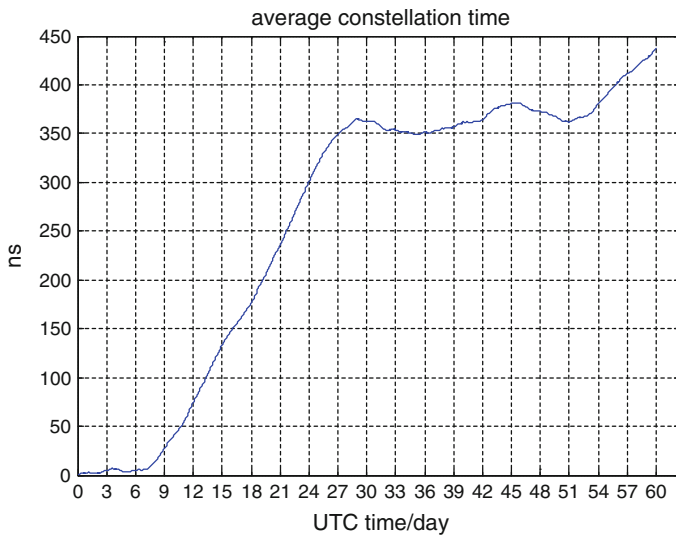


Fig. 32.4 Option II constellation time drift error

Option III:

Figures 32.5 and 32.6 are the hydrogen + ultra-stable oscillator clock synchronization accuracy statistics chart. Due to the non-GEO satellite use ultra-stable oscillator, its long steady accuracy is poor, when there has no two-way time

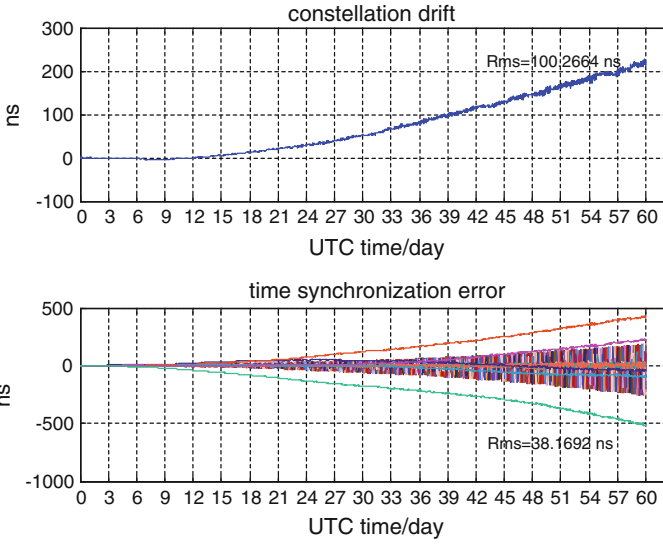


Fig. 32.5 Option III constellation drift and satellite time synchronization error

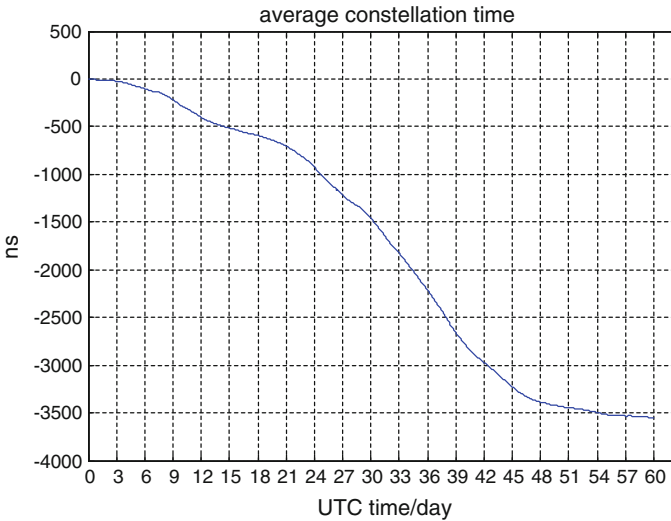


Fig. 32.6 Option III constellation time drift error

synchronization, its constellation overall drift is 3500 ns (Fig. 32.6). Placed passive hydrogen on the GEO satellite as a time synchronization reference, the constellation average time is the maximum of 200 ns (Fig. 32.5), we can see the constellation of time synchronization error is diverged.

Option IV:

Figures 32.7 and 32.8 are the hydrogen clock + rubidium, cesium clocks synchronization accuracy statistics charts. The results are similar to the option III, the whole constellation of time drift is also larger (Fig. 32.8), the constellation of time

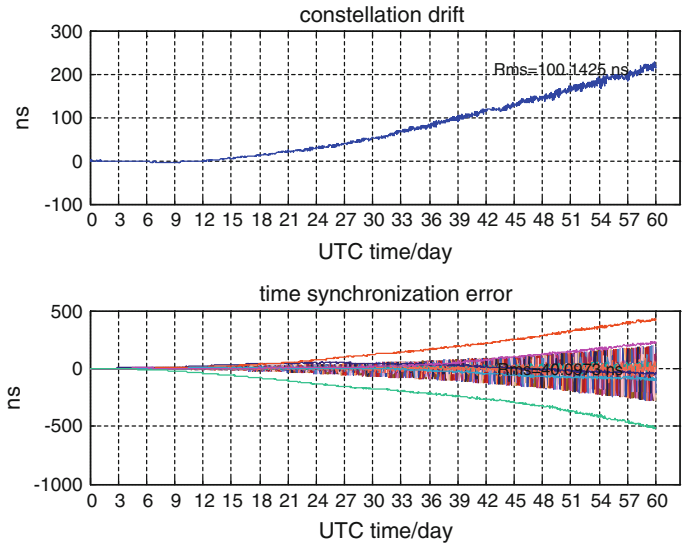


Fig. 32.7 Option IV constellation drift and satellite time synchronization error

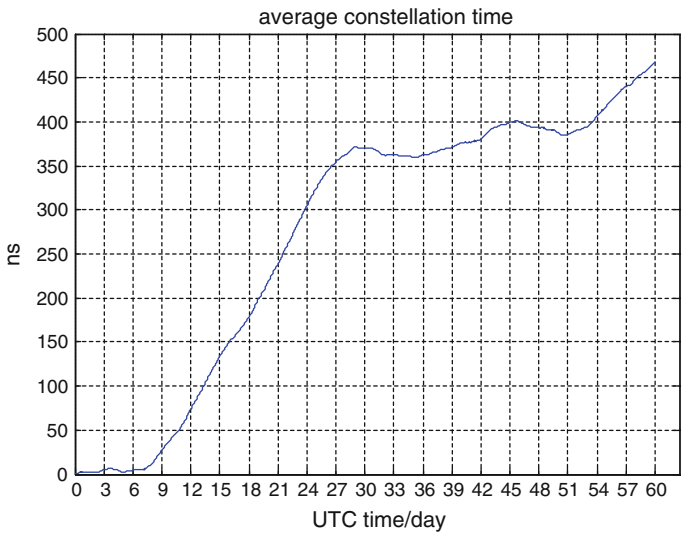


Fig. 32.8 Option IV constellation time drift error

synchronization error is diverged, combine the results of option I and II, it could prove that hydrogen program is difficult to meet the requirements of autonomy time synchronization.

32.6 Conclusion

Satellite atomic clocks have frequency drift problems, when time lose; the accuracy will be getting worse. In this paper, in the independent time synchronization mode, the synchronization accuracy of hybrid navigation constellation depends on the accuracy of GEO satellite clock, the higher the accuracy of GEO clock, the accuracy of the navigation constellation is also higher. Hydrogen maser clock's accuracy is far than optical clock, when the time passed, the entire constellation of time is drifted. It can set high-precision satellite clock such as optical atomic clock when can achieve the purpose of self-time synchronization.

Due to the introduction of high-precision optical clocks, the overall time of the entire navigation constellations drift is small, the accuracy is much higher than the link between self-ordinary star mode for time synchronization accuracy (typically tens of ns); and because the presence of high-precision clock, the time synchronization algorithm's requirements are lower, general EKF can meet the requirements.

References

1. Cheng ZZ, Ying LY (2010) Next-generation onboard atomic clocks. *Glob Positioning Syst* 35 (5):1–5
2. André H (2011) Precise GNSS clock-estimation for real-time navigation and precise point positioning. In: DLR deutsches zentrum fur luft- und Raumfahrt e.V.-Forschungsberichte, pp 28–44
3. Levi F, Godone A, Micalizio S et al (2004) CPT maser clockevaluation for Galileo. Guildford, UK, Frequency and Time Forum, pp 233–238
4. Godone A, Levi F, Micalizio S et al (2004) Coherent population trapping maser: noise spectrum and frequency stability. *Phys Rev A* 70(1):2508–2519
5. Tremine S, Guerandel S, Holleville D et al (2006) Limitations to the short term frequency stability in a compact cold atom clock. In: Proceedings of the 2005 IEEE international on frequency control symposium and exposition, Paris, France, pp 591–598
6. Esnault FX, Perrin S, Tremine S et al (2007) Stability of the compact cold atom clock HORACE. In: 2007 joint with the 21st European Frequency and time forum on frequency control symposium IEEE international Geneva, pp 1342–1345
7. Svehla D (2008) A novel design for the navigation system and proposal to unify the timing and the positioning system using GIOVE follow-on. In: ACES and future GNSS-based earth observation and navigation, Munich, Germany, 26–27 May 2008, pp 1–27
8. Ping S, Guang-ji Q (2005) Time synchronization techniques of the autonomous navigation of navigation constellation. *J Astronaut* 26(6):768–772
9. Diez J, D'Angelo P, Fernández A (2006) Clock errors simulation and characterisation. In: 19th international technical meeting of the satellite division on ION GNSS, Fort Worth, Texa, pp 815–821
10. Weaver G (2010) The performance of ultra-stable oscillators for the gravity recovery and interior laboratory. In: 42nd annual precise time and time interval (PTTI) meeting, pp 369–379

Chapter 33

Fiber Based Radio Frequency Dissemination Scheme to Multiple Users

Wei Chen, Dan Xu, Nan Cheng, Qin Liu, Fei Yang, Youzhen Z. Gui
and Haiwen W. Cai

Abstract High-precision time and frequency signals play an important role in many areas such as modern communication, navigation, and baseline interferometry. Optical fiber based time and frequency networking not only has a higher precision which can meet the demand of new generation time and frequency standard but also constructs to the ground based time and frequency network as a complementary choice of satellite based methods. As a consequence, we propose a fiber based radio frequency dissemination scheme to multiple users with a tree-like topology. The noise compensation system acquiring the noise information by measuring multiple reflection signal is put in remote terminals, which allow users to access the frequency signal along the fiber without affecting others. With this scheme, a high precise frequency transfer experiment is demonstrated along a 60 km fiber link to multiple users. The frequency stability of 7.1×10^{-14} @1 s and 6.8×10^{-17} @ 10^4 s are obtained. Furthermore, the influence induced by adding or cancelling a new terminal is tested when other terminals are operating. The result shows that the stability of radio frequency networking scheme which is robust and flexible can fulfil most atomic clock that is currently used and would bring benefit to establish a national or even continental ultra-stable frequency dissemination network.

Keywords Laser technique · Frequency transfer · Optical fiber · Time and frequency networking · Atomic clock

W. Chen (✉) · D. Xu · N. Cheng · F. Yang · H.W. Cai (✉)
Key Laboratory of All Solid-State Laser and Applied Techniques,
Shanghai Institute of Optics and Fine Mechanics, CAS, Shanghai, China
e-mail: xmchenwei@163.com

H.W. Cai
e-mail: hwcai@siom.ac.cn

Q. Liu · Y.Z. Gui (✉)
Key Laboratory for Quantum Optics, Shanghai Institute of Optics
and Fine Mechanics, CAS, Shanghai, China
e-mail: yzgui@siom.ac.cn

Y.Z. Gui · H.W. Cai
Shanghai Institute of Optics and Fine Mechanics, CAS, Shanghai 201800, China

33.1 Introduction

With the progress of the atomic clocks, the stability of commercial hydrogen clocks has reached to $10^{-13}@1$ s. Moreover, the stability of the optical clocks has come to $10^{-16}@1$ s [1]. As a matter of fact it leads to the rise of the fiber based frequency transfer which can be as an alternative choice of the GPS methods. With characteristic of low transmission loss, anti-electromagnetic interference and high stability, fiber based frequency transfer can fulfil the needs of the transmission of high-precision clock. In recent years, the developed countries have rushed to carry out researches on the fiber based frequency transfer techniques. Highlights covers: simultaneous transfer of radio frequency and time [2]; radio frequency transfer alongside internet data traffic [3]; optical frequency transfer over a single span 1840 km fiber with ultra- stable performance [4].

However compared to traditional satellite-based transmission ways, point to point fiber based frequency transfer scheme has its inevitable limitations. Its coverage cannot meet the demand of most practical application such as modern large linear accelerators, very long base lines interferometry and satellite positioning [5]. Multi-points frequency transfer is needed in all these applications. Therefore, fiber based frequency networking technology has a broad range of applications and urgent demands. Its difficulties lies in how to take advantages of the topological structure, simplify the complexity of the system and enhance the system's convenience and robustness. As a linear topology, the midpoint extraction in a trunk fiber for multiple-access has been firstly proposed and demonstrated to solve the problem [6, 7].

In order to better adapt to branching network structure. We propose a precise radio frequency dissemination scheme to multiple users with a tree-like topology. In this scheme, phase noise can be detected and compensated in remote terminals. Each remote terminal can work independently by using wavelength division multiplexing (WDM) technology. It leaves the advantage for users to access the high precise radio frequency online. Network adaptability and system robustness are greatly enhanced.

33.2 Networking Scheme

33.2.1 Theory and Set-up

The schematic diagram of the fiber based frequency dissemination network is shown in Fig. 33.1. Radio frequency generated from a precise atomic clock is sent to multiple remote terminals from the local terminal via a tree-like fiber network. The high precise radio frequency signal can be exactly reproduced by identifying and suppressing the phase noise in the far end of the link. In the experiment, the 10 MHz frequency signal is firstly boost to 1 GHz through a frequency multiplier

passes through an optical delay line (ODL) including a temperature controlled fiber ring (dynamic range about 3.4 ns, sensitivity of 40 ps/°C) and a fast fiber stretcher (dynamic range 17 ps, response speed of a few hundreds kHz) as a compensated structure. The loop bandwidth B is limited by the propagation delay τ . It satisfies the relations of $B < 1/(4\tau)$ [9]. For 25 and 60 km links, they are 2 kHz and 833 Hz respectively.

After passing through a circulator, the frequency signal goes through the corresponding channel of a WDM. Detected by the low noise photo detector, the frequency signal can be expressed as:

$$V_{a1} = \cos(\omega t + \varphi_0 + \varphi_p). \quad (33.2)$$

It includes one trip phase noise term φ_p . Then it splits into three parts. The first part is the final frequency output. The second part sends to the phase discriminator as a reference signal and the last part as a measuring signal modulates another DFB laser ($\lambda_2 = 1550.9$ nm). This measuring signal sends back into the long haul fiber. After the back-and-forth structure, the measuring frequency signal comes back to the remote site once again. The optic path is assumed to be symmetrical. The detected measuring signal after de-multiplexing with the corresponding channel is expressed as:

$$V_{a3} = \cos(\omega t + \varphi_0 + 3\varphi_p). \quad (33.3)$$

It passes three times of the optical path, thus the phase noise term is $3\varphi_p$. This detected measuring signal is also sent to the phase discriminator. Therefore we get the error signal of $2\varphi_p$ and it feedback controls the ODL for compensation after the proportion-integration-differentiation (PID) arithmetic. The ODL generates a phase drift of $-\varphi_p$. Thus the final output frequency will be changed to:

$$V'_{a1} = \cos(\omega t + \varphi_0). \quad (33.4)$$

It is stable and will be down converted to 10 MHz for users by a frequency divider (FD).

WDM technology is used for measuring frequency signals of different remote terminals. In our experiment, another wavelength ($\lambda_3 = 1549.3$ nm) is applied to the remote terminal 2. It ensures that each remote terminal can only measure the corresponding phase noise of the fiber link. Each remote terminal works independently without disturbing the others. In this way, high precise frequency dissemination to multiple users is realized.

33.2.2 Experimental Results

The experimental set-up is constructed as we proposed before. The regenerated frequency signals of two remote terminals are compared to the local clock. The results of relative frequency stability (typically represented by overlapping Allan

deviation (ADEV)) are shown in Fig. 33.2. The dark curves in the figure are free running link. Taking the advantages of optical fiber, the short time stability can reach the order of 10^{-14} in experiment circumstance. In contrast, the long term stability is drifting owing to the influence of ambient temperature. With the increase of distance, the influence of optical lose, backscattering light, ambient temperature and other nonlinear effects are obviously observed. We can observe from the curve that free running state of remote terminal 1 (60 km away from local terminal) is worse than remote terminal 2. The performances of compensated link are also shown in the figure (curves). The short time stability improves not much because of the limitation of intensity noise generated from some active optoelectronic components (such as relative intensity noise of laser,spontaneous emission noise of EDFA, shot noise and thermal noise of detector). The ADEV are $7.1 \times 10^{-14}@1$ s and $4.6 \times 10^{-14}@1$ s respectively. However the long term performance is dramatically enhanced. Even though the optical path is not strictly symmetrical (e.g. the asymmetry of the back-and-forth structure, the different wavelength between back and forth), it can reach $6.8 \times 10^{-17}@10^4$ s which has three orders of improvement.

33.3 Robustness Discussion

For cyber structure, the traditional star-like networking topology may achieve high precise frequency distribution as well by increasing the complexity of noise suppression apparatus at local terminal. However for practical application, a reasonable frequency distribution networking structure which can accommodate new terminals conveniently without affecting others is expected [10]. The tree-like radio frequency dissemination scheme we proposed can theoretically fulfil the requirement. The robustness is discussed.

For analysing the influence of phase fluctuations when a new terminal is withdrew or added,three circumstances are taken into consideration. First, when remote

Fig. 33.2 Relative frequency stability

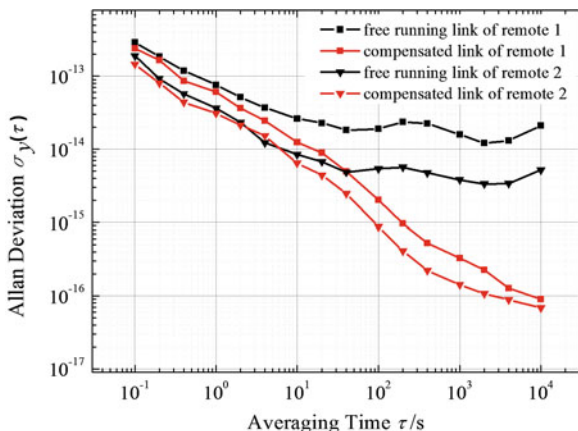
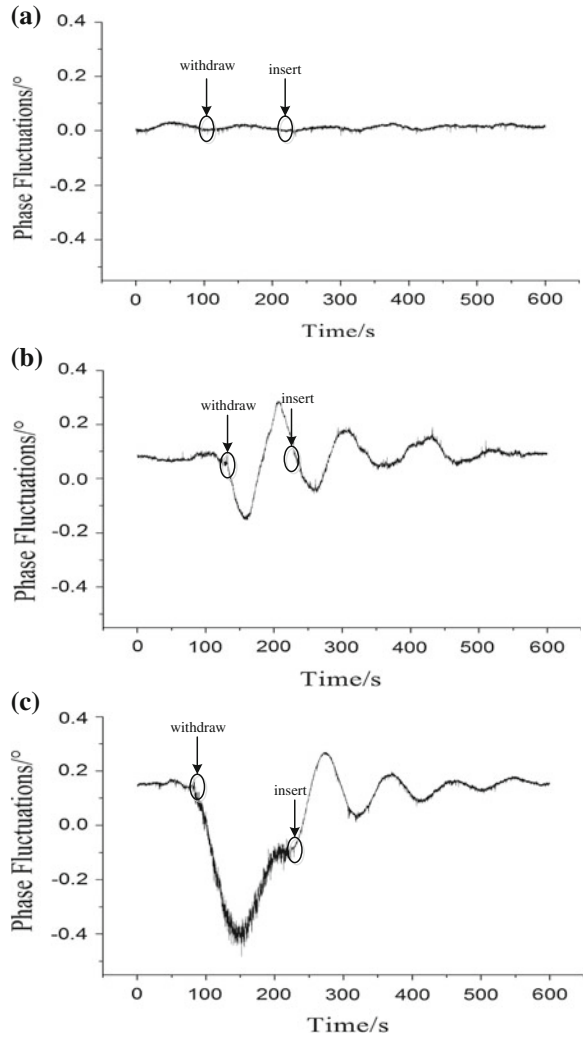


Fig. 33.3 Phase fluctuations while online access



terminal is not far away from the local station, optical power budget is sufficient. In this circumstance EDFA is not needed in the back-and-forth structure. We take an experiment based on the tree-like dissemination scheme, but replace the length of the two spooled fiber to 1 and 10 km. So the remote terminals are 1 and 11 km away from the local station. The EDFA is taken away. When the dissemination system is working on phase compensated state, remote terminal 1 is suddenly withdrew and then inserted two minutes later. The phase fluctuation of radio frequency output of remote terminal 2 is detected. It is shown in Fig. 33.3a. It indicates that the stable radio frequency signal will not be affected in this circumstance.

Second, when a remote terminal is further, optical power budget is insufficient for the further terminal. For an experiment, the tree-like networking dissemination

set-up is still used, but the length of the two spooled fiber is changed to 10 and 25 km. An EDFA is used in the back-and-forth structure to amplify the measuring signals which are generated from the two remote terminals. The gain of the EDFA is carefully controlled. We make sure that the total output power after the EDFA is below stimulated Brillouin threshold. Just like the first circumstance, remote terminal 1 is suddenly withdrew and then inserted two minutes later. The phase fluctuation of the remote terminal 2 is monitored. It is shown in Fig. 33.3b. Some turbulence are observed. But because the compensated structure is working, the phase fluctuation is stabilized after several cycles of vibration. The main reasons for this phenomenon is the transient effects of the EDFA [11]. Due to constant pump power, the light energy of the withdrawn terminal will changed to the online terminal. That is to say that the gain of the remote terminal 2 will be enhanced, which will induce the phase fluctuation. In practical applications, the pump light power can be feedback controlled. The phase fluctuation will be further suppressed.

Third, when remote terminals are far away from the local station, optical power is not enough for all terminals. In the experiment, the length of the two fiber spools is chosen to 25 and 35 km. We just use one EDFA for experimental test. The output power of each measuring signal after the EDFA will be controlled below stimulated Brillouin threshold. But the total optical power is exceed the threshold. Then the remote terminal 1 is withdrawn and inserted just like we tested before. The phase fluctuation is shown in Fig. 33.3c. The gain of EDFA for remote terminal 2 is enlarged while the remote terminal 1 is withdrawn. The optical power of the measuring signal exceeds the stimulated Brillouin threshold in this circumstance. So the curve is much thicker than before. It may influence the performance of the dissemination system. However when the terminal inserted back, the phase fluctuation is stabilised quickly.

Therefore, for practical applications, the total optical power of all wavelength channels should keep below the stimulated Brillouin threshold all along the fiber link. Meanwhile the fiber length of the asymmetry part should be as short as it can and ambient temperature should be controlled for reduce the influence of unsuppressed delay noise. In general, whichever the circumstance is, the tree-like dissemination scheme allows users to insert or withdraw online terminals and the phase fluctuation can rapidly restore to stable. That is to say, even if some remote terminal is broken or the branching fiber is cut off, other terminals will not be seriously affected and the high precise frequency signal can be accessed at any time.

33.4 Conclusion

A tree-like radio frequency networking scheme is proposed. The performance of the dissemination scheme is test by an experiment. Relative stability of 7.1×10^{-14} @1 s and 6.8×10^{-17} @ 10^4 s are obtained. The influence to phase fluctuation when a terminal is added or canceled is further discussed. The result shows that the fiber based frequency dissemination scheme to multiple users can fulfil the most practical applications which is currently used.

References

1. Jiang YY, Ludlow AD, Lemke ND, Fox RW, Sherman JA, Ma LS, Oates CW (2011) Making optical atomic clocks more stable with 10(-16)-level laser stabilization. *Nat Photonics* 5(3):158–161
2. Krehlik P, Śliwczyński Ł, Buczek Ł, Lipiński M (2012) Fiber-optic joint time and frequency transfer with active stabilization of the propagation delay. *IEEE Trans Instrum Measur* 61(10):2844–2851
3. Lopez O, Haboucha A, Chanteau B, Chardonnet C, Amy-Klein A, Santarelli G (2012) Ultra-stable long distance optical frequency distribution using the Internet fiber network. *Opt Express* 20:23518–23526
4. Droste S, Ozimek F, Udem Th, Predehl K, Hansch TW, Schnatz H, Grosche G, Holzwarth R (2013) Optical-frequency transfer over a single-span 1840 km fiber link. *PRL* 111(11):110801
5. Calhoun M, Huang S, Tjoelker RL (2007) Stable photonic links for frequency and time transfer in the deep-space network and antenna arrays. *Proc IEEE* 95(10):1931–1946
6. Grosche G (2010) Method for making available a reference frequency, German patent application DE, 10.2008.062.139
7. Gao C, Wang B, Chen WL, Bai Y, Miao J, Zhu X, Li TC, Wang LJ (2013) Fiber-based multiple-access ultrastable frequency dissemination. *Opt Lett* 37(22):4690–4692
8. Lopez O, Amy-Klein A, Daussy C, Chardonnet C, Narbonneau F, Lours M, Santarelli G (2008) 86 km optical link with a resolution of 2×10^{-18} for RF frequency transfer. *Eur Phys J D* 48:35–41
9. Williams PA, Swann WC, Newbury NR (2008) High-stability transfer of an optical frequency over long fiber-optic links. *J Opt Soc Am B* 25(8):1284–1293
10. Sascha WS, David G, Baldwin KGH, Brian JO, Warrington RB, Guido A, Luiten AN (2013) High-precision optical-frequency dissemination on branching optical-fiber networks, *Opt Lett* 38(15):2893–2896
11. Jiang N (2007) Measure of recover time and research of gain-flattening technology of Erbium-doped fiber. Beijing Jiaotong University, Beijing, p 8–14

Chapter 34

Influence of Lamp Spectral Profile on Short-Term Stability and Light Shift of a Rubidium Atomic Clock

Qiang Hao, Shengguo He, Feng Xu, Feng Zhao and Ganghua Mei

Abstract Passive gas-cell rubidium atomic clock has been widely used in Global Navigation Satellite System (GNSS). Further improvement on the frequency stability of rubidium atomic clocks has extreme significance for promoting the positioning and timing precision of GNSS. Distortion of the lamp spectral profile, caused by self-absorption, is a common concern for the design of a rubidium clock. However, few literature has systematically investigated the impact of lamp spectral profile on a Rb clock. In this work, the influence of lamp spectral profile on both short-term frequency stability and light shift of a rubidium clock was studied theoretically and experimentally. The results showed that serious distortion of the lamp spectral profile could lead to one times deterioration to short-term frequency stability, and change the zero light shift point by a few degrees centigrade. Thus, we demonstrated that optimization of the lamp profile may be an effective way to improve the performance of a Rb clock, which should be paid more attention when designing a high performance spaceborne Rb clock.

Keywords Rb atomic clock · Lamp spectral profile · Frequency stability · Light shift

Q. Hao (✉) · S. He · F. Xu · F. Zhao · G. Mei

Key Laboratory of Atomic Frequency Standards of Chinese Academy of Sciences,
Wuhan Institute of Physics and Mathematics, Wuhan 430071, China
e-mail: wlsichuang@163.com

G. Mei

e-mail: mei@wipm.ac.cn

Q. Hao · S. He · F. Xu

University of Chinese Academy of Sciences, Beijing 100049, China

© Springer-Verlag Berlin Heidelberg 2015

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015*

Proceedings: Volume III, Lecture Notes in Electrical Engineering 342,

DOI 10.1007/978-3-662-46632-2_34

34.1 Introduction

Being of compact volume, light weight, low power consumption and high reliability, the passive gas-cell rubidium atomic clock has been widely used in GNSS. Further improvement on Rb clocks frequency stability is critical for the positioning and timing precision of GNSS.

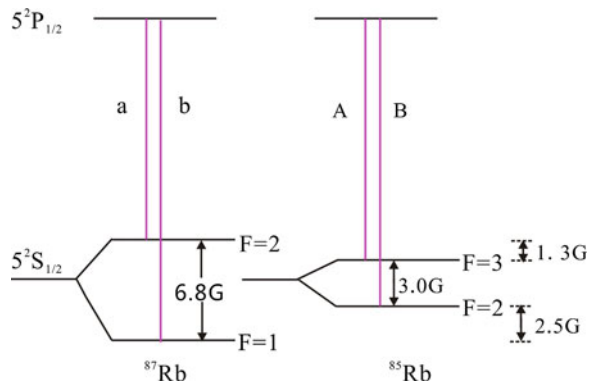
In a lamp pumped Rb clock, the discharge lamp is served as light source to create population inversion between the ground state hyperfine energy levels $F = 1$ and $F = 2$. Spectrum of the lamp mainly consist of D_1 (795 nm) and D_2 (780 nm) lines, each line is composed of the a and the b components. The b component is beneficial while the a component is adverse for creating population inversion. For this reason, hyperfine filter is commonly employed to remove the a component. The energy levels related to D_1 line of the two isotopes, ^{87}Rb and ^{85}Rb , are shown in Fig. 34.1.

As shown in Fig. 34.1, the A component of ^{85}Rb is near to the a component while the B component is significantly far from the b component. Furthermore, the hyperfine levels of ^{85}Rb atom are broadened and shifted by the buffer gas. Thus, when the light emitted from the lamp propagates in the filter cell, the a component will be absorbed dramatically.

The so-called self-absorption effect, i.e. light emitted by specific atoms might be absorbed by other atoms in the bulb, could lead to broadening, asymmetry and even dips of the lamp spectral profile [1]. Variation of The lamp spectral profile could influence both the short-term and long-term frequency stability of a Rb clock. As for the short-term stability, it affects the optical power spectral density directly, which dominates the optical pumping efficiency. For the long-term stability, it could influence light shift, which may be a limit for long-term stability (at average time greater than 10,000 s) [2].

In general, distortion of the lamp spectral profile, caused by self-absorption, is an ineluctable problem to various extent. Its impact may be negligible for a lower performance Rb clock, whereas may be notable for a high performance one. To the

Fig. 34.1 Energy levels of the ground and the 1st excited states of ^{87}Rb and ^{85}Rb atom



best of our knowledge, no previous investigation focuses on the influence of the lamp spectral profile on the performance of a Rb clock. In this work, the influence of lamp spectral profile on both short-term frequency stability and light shift of a Rb clock was studied theoretically and experimentally.

34.2 Theoretical Method

34.2.1 Optical Power Spectral Density and Light Shift

Electric field of the incident light could induce an electric dipole moment in each atom, which is proportional to the electric field. The dipole moment can be expressed as $\vec{p} = \alpha \vec{E}$, where α denotes polarizability, \vec{E} is electric field vector. Interaction between atoms and light can be expressed by Liouville equation [3]:

$$i\hbar(d/dt)\rho = [H_0 + \delta H, \rho] \quad (34.1)$$

$$\delta H = -(|E_0|^2/4)\vec{e}^* \cdot \alpha \cdot \vec{e} \quad (34.2)$$

where δH is a perturbation caused by light, \vec{e} is the polarization vector of light, α is the complex polarizability of Rb atom, in which the real and the imaginary part are related to absorption and dispersion respectively, it can be equivalent to four terms [4]:

$$\alpha = \alpha_{eq} + \alpha_{hfs}\vec{I} \cdot \vec{J} + i \sum_{ff'} a_{gt}(ff')\vec{J}(ff') + \sum_{ff'} a_{br}(ff')\vec{Q}(ff') \quad (34.3)$$

The first term is named equilibrium polarizability, causing an equal displacement for all of the ground state sublevels, so it affects the light propagation while doesn't affect the light shift. The second term is hyperfine structure polarizability, which relates to the variation of absorption and dispersion when the pumping light alters the population of the ground state hyperfine levels. The third term is gyrotropic polarization, whose contribution is zero for the light emitted by lamp is linear polarization. The last term is birefringent polarizability, it could make a certain contribution when the splitting of excited state hyperfine structure couldn't be ignored compared with Doppler broadening.

On basis of Eq. (34.3) and ignoring the third term, the absorptivity, relating to the imaginary part of α , can be expressed as:

$$t = 4\pi kN(\text{Im}\alpha_{eq} + \text{Im}\alpha_{hfs}\langle \vec{I} \cdot \vec{J} \rangle + \sum_{ff'} \text{Im}\alpha_{br}(ff')\vec{e}^* \cdot \langle \vec{Q}(ff') \rangle \cdot \vec{e}) \quad (34.4)$$

where k is Boltzmann constant, N is atomic density.

Light shift is related to the real part of polarizability, thus light shift response function $S(\nu)$ is expressed as [5]:

$$S(\nu) = S_{hfs}(\nu) + S_t(\nu) \quad (34.5)$$

where $S_{hfs}(\nu)$ and $S_t(\nu)$ represent hyperfine structure response function and tensor response function respectively.

Considering light absorption caused by microwave resonance in absorption cell is very small, so power spectral density of the light in both filter and absorption cell could be acquired by the same expression:

$$F(\nu, z) = \Phi(\nu) \exp[-t(\nu)z] \quad (34.6)$$

where $\Phi(\nu)$ is power spectral density of incident light, $t(\nu)$ is absorptivity, z is position in the filter or absorption cell.

In general, the measured light shift is an average of all resonance atoms pumped by the light with various optical power spectral density, so the absorption cell is divided into successive layers with a number of n , and the light shift $\delta\nu_{hfs}$ is approximately calculated as an average of all layers:

$$\delta\nu_{hfs} = \frac{1}{n} \sum_{z=L/n}^L \int_{-\infty}^{+\infty} [S_{hfs}(\nu) + S_t(\nu)] A(\nu, z) d\nu \quad (34.7)$$

where L is the length of absorption cell, $A(\nu, z)$ is power spectral density of the pumping light at position z , which could be calculated from Eq. (34.6).

For a specific lamp spectral profile, the zero light shift point of a Rb clock is the temperature of cavity-cell assembly when $\delta\nu_{hfs}$ is zero.

34.2.2 Short-Term Frequency Stability

According to the theory developed in [6], pumping rate $\Gamma_i(z)$ can be expressed as:

$$\Gamma_i(z) = \frac{1}{h\nu S} \int_{-\infty}^{+\infty} A_i(\nu, z) \sigma_i(\nu) d\nu \quad (34.8)$$

where $i = 1$ corresponds to the a component, $i = 2$ corresponds to the b component. $\sigma_i(\nu)$ is absorption cross section, S is cross-section area of the light beam, h is Planck constant.

The microwave-optical double resonance technique is used in Rb clock, so the absorbed light intensity $\Delta I(\omega)$ in the absorption cell is a function of the microwave frequency:

$$\Delta I(\omega) = Shv \int_0^L \Gamma_1(z)n_1(z, \omega) + \Gamma_2(z)n_2(z, \omega)dz \tag{34.9}$$

where n_1 and n_2 are the population of ground state hyperfine levels $F = 2$ and $F = 1$, which can be acquired by solving the rate equation. Then short-term stability $\sigma(\tau)$ can be estimated by [7]:

$$\sigma(\tau) = 1E(-10)\sqrt{2eI_0}/S_d\tau^{-1/2} \tag{34.10}$$

where e is electronic charge, I_0 is background current, S_d is discrimination slope, which can be obtained from $\Delta I(\omega)$ [8].

34.3 Experimental and Theoretical Results

34.3.1 Spectral Profiles

By adjusting excitation frequency and power, temperature, species of buffer gas of spectral lamps, three lamps with different spectral profiles were acquired. In the spectral profiles measurement, a Fabry–Pérot interferometer with a free spectral range of 15 GHz and a fineness of 30 was employed, and optical filters were adopted to ensure either D_1 or D_2 line could transmit.

The spectral profiles of the lamps are illustrated in Fig. 34.2, the lamps showed a tendency of alleviation in distortion of spectral profiles and are marked as L1, L2, L3 respectively, the buffer gas in L1 is Ar and in L2, L3 are Xe. It is found that the spectral profile of L1 has an obvious dip and the widest width, distortion of the spectral profile of L2 is lower and the spectral profile of L3 has the best symmetry

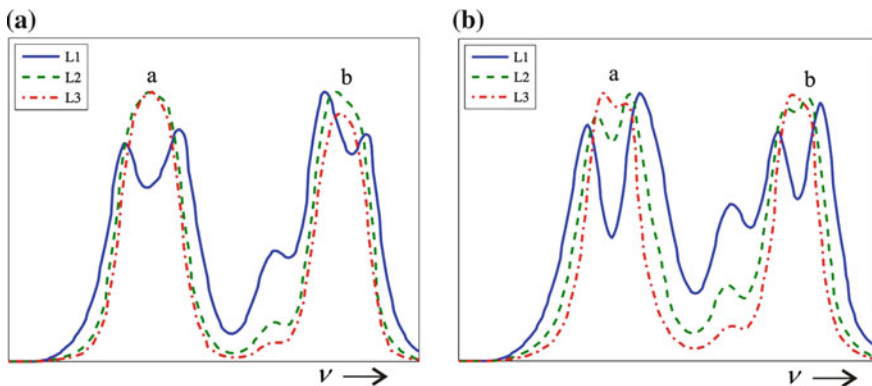


Fig. 34.2 D_1 (a) and D_2 (b) line spectral profiles of the three lamps

and the narrowest width. It is also found that distortion of D_1 line is lower than D_2 line, meanwhile the b component presents lower distortion than the a component, that is due to the differences in their absorption cross section.

The relative light intensity of D_1 and D_2 were also observed by a spectrometer. The percentage of D_1 line in total light intensity of L1, L2 and L3 were 45, 45 and 42 % respectively. The influence of the differences will be discussed below.

In the calculation of light shift and short-term stability, power spectral density of the pumping light at arbitrary position is required. So spectral profiles of the light transmitted through the filter and the absorption cell were measured. The results of D_2 line are shown in Fig. 34.3a, b. The situation of D_1 line is similar with D_2 line. Then simulation spectral profiles of the above two positions of D_2 line was obtained on basis of Eq. (34.6). In the calculation, the required α_{eq} and α_{hfs} were described in [4], $\langle \vec{I} \cdot \vec{J} \rangle$ was taken as -0.2 for filter cell and 0.1 for the absorption cell, estimating by the measured spectral profiles, birefringence effect produced a tiny influence on the light absorption in our assumption. Figure 34.3c shows the

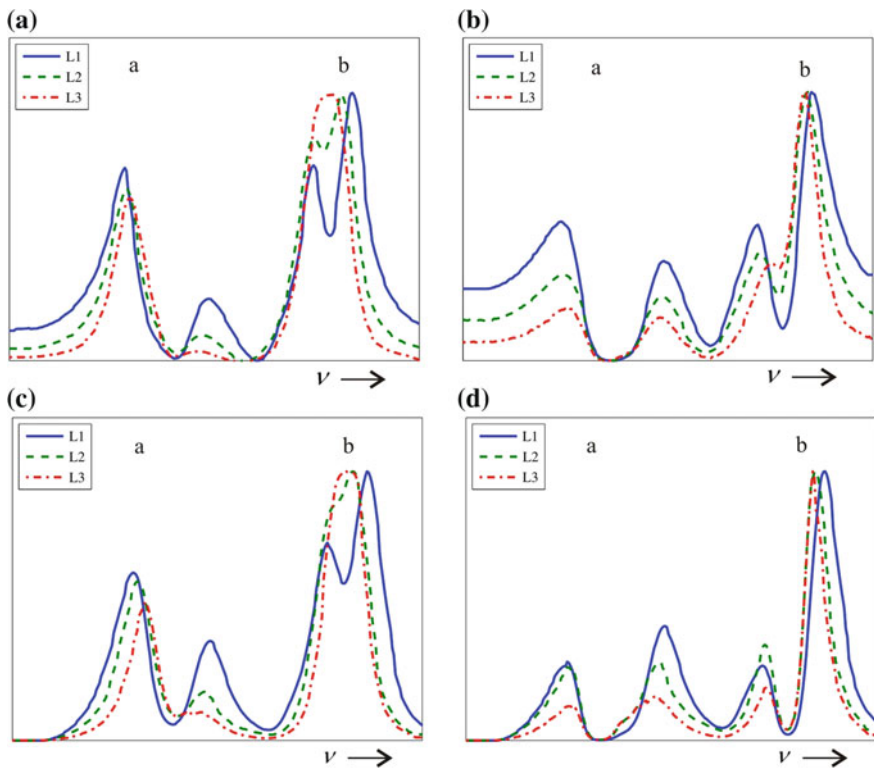


Fig. 34.3 Measured spectral profiles of the D_2 line transmitted through the filter cell (a) and the absorption cell (b) of the three lamps. c and d are the theoretically calculated profiles corresponding to the spectral profiles in (a) and (b), respectively

theoretical spectral profiles of the light though filter cell, and (d) shows the theoretical spectral profiles of the light though absorption cell. As theoretical results of the two typical positions present a good agreement with the experiment, so it demonstrates that the optical power spectral density at arbitrary position for the both cells could be acquired comparably accurately by the theoretical method.

34.3.2 Light Shift

Light shift of a Rb clock was experimentally measured with the three lamps respectively, as is shown in Fig. 34.4. In the experiment, both the filter and absorption cell were put in a slotted-tube cavity [9], light intensity of the lamp was approximately controlled at 110 uA with optical neutral attenuators, an optical filter with a center wavelength of 786 nm and half width of 30 nm was adopted to ensure only D_1 and D_2 lines can transmit.

Light shift, due to AC Stark effect, is related to symmetry and width of the lamp spectral profile. As shown in Fig. 34.4, it is found that the zero light shift point of the cavity-cell assembly is raised several degrees with distortion of the spectral profile aggravating.

For the theoretical study, on basis of the expressions in [5], the light shift response curves of D_1 and D_2 were calculated, as is shown in Fig. 34.5. The curve of D_2 line presents better symmetry while the curve of D_1 line appears obvious distortion, this is due to hyperfine structure splitting of the state $5^2P_{1/2}$ is much larger than $5^2P_{3/2}$. Besides, it can be generalized that D_2 line has greater influence on light shift since the intensity (55 % of L1 and L2, 58 % of L3) and the light shift response (much higher than D_1 line).

Fig. 34.4 Measured light shift related to the three lamps, the legend illustrates relative light intensity

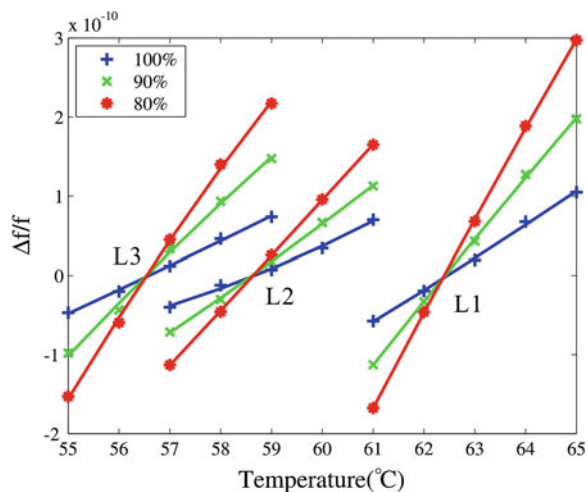
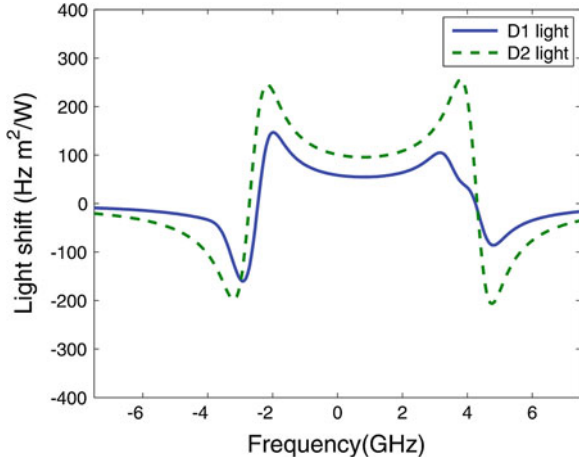


Fig. 34.5 Light shift response curve of ^{87}Rb atom



Given that optical power spectral density of the light at arbitrary position could be acquired based on Eq. (34.6), thus zero light shift points corresponding to the three lamps could be calculated according to Eq. (34.7) and the results are listed in Table 34.1 with the experimental results. In the calculation, the length of absorption cell L is 14 mm and the value of n is took as 14. It can be observed that the theoretical results present the same variation trend with experiment, although there is a deviation between the theoretical and experimental values, which mainly origins from the distinction in theoretical model and practical situation.

It should be mentioned that, in the above experiment and calculation, D_1 line ratio to the total light intensity for the three lamps are not the same, 45 and 45 % for L1 and L2, and 42 % for L3. To clarify influence of the differences, a calculation for L3 was performed, in which the percentage intensity of D_1 line was assumed to be 45 %. The obtained zero light shift point was 58 °C, which is much close to 58.1 °C as given in Table 34.1. So we conclude that a few percent change of relatively light intensity of D_1 and D_2 lines didn't affect the zero light shift point so much, and differences of the zero light shift points were due to the variation of lamp spectral profile.

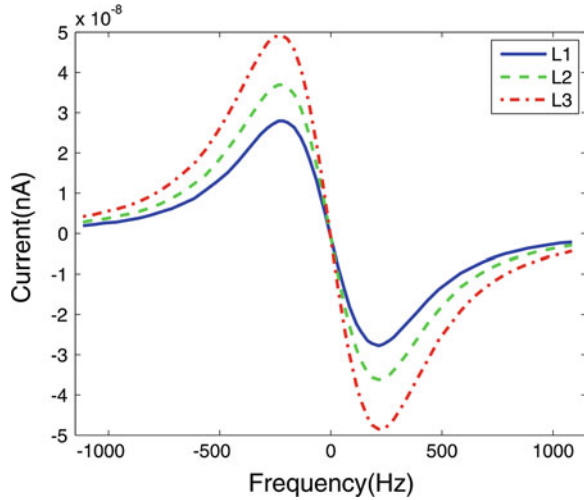
34.3.3 Short-Term Stability

Discrimination slopes of a Rb clock with the three spectral lamps as pumping sources were measured respectively. In experiment, the cavity-cell assembly worked at 68 °C and the optical filter, neutral attenuators used in light shift measurement were still used.

Table 34.1 Experimental and theoretical results of zero light shift points for the three lamps

		L1	L2	L3
Zero light shift point (°C)	Experiment	62.4	58.6	56.5
	Theory	64.2	60.4	58.1

Fig. 34.6 Experimentally obtained S curves corresponding to the three lamps



The measured S curves are shown in Fig. 34.6. It is obvious that slope of center part of the S curve, i.e. the discrimination slope, decreases with distortion of the lamp spectral profile aggravating.

The theoretical calculation was performed in three steps. Firstly, we obtained the pumping rate according to Eq. (34.8) and the results are shown in Fig. 34.7. Γ_2 represents pumping rate of the Zeeman sublevels of $F = 1$, which is beneficial to creating population imbalance at the two ground state hyperfine levels. Γ_1 represents pumping rate of the Zeeman sublevels of $F = 2$, which is adverse to the population imbalance. Therefore, total population difference is related to $(\Gamma_2 - \Gamma_1)$, the greater integral area between Γ_2 and Γ_1 indicates the higher signal to noise ratio

Fig. 34.7 Variation of pumping rate in the absorption cell

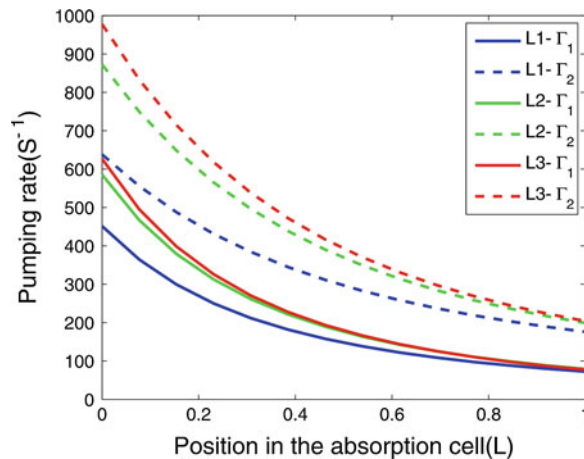


Table 34.2 Experimental and theoretical results of short-term frequency stability

		L1	L2	L3
Incident light current (uA)		111	112	110
Background light current (uA)	Experiment	37	30	25
	Theory	29	24	22
Slope of discriminator (E-10 A/Hz)	Experiment	2.1	2.8	3.7
	Theory	2.4	3.2	3.8
Short-term stability (E-12 $\tau^{-1/2}$)	Experiment	1.7	1.1	0.8
	Theory	1.3	0.9	0.7

of the double resonance and thus better short-term stability. It can be found from Fig. 34.7, the area has a trend of decreasing for L3, L2, and L1.

Secondly, we acquired the population of the two ground state hyperfine levels by solving rate equation.

Finally, we calculated discrimination slope and short-term frequency stability by Eqs. (34.9) and (34.10). Experimental and theoretical incident light intensity, background light intensity, discrimination slope and short-term stability are listed in Table 34.2. Assuming the intensity proportion of D_1 is 45 % for L3, the calculated short-term stability is $0.7E-12\tau^{-1/2}$, thus the influence of relatively intensity variation of D_1 and D_2 line could be neglected in our situation.

As the results listed in Table 34.2, it demonstrates that the distortion of lamp spectral profile deteriorates the short-term frequency stability to a great extent. This deterioration effect is easy to be understood, with distortion of the lamp spectral profile aggravating, the spectral profile will be wider, even appear dips, thus the spectral power density in the center will decrease. On one hand, the optical pumping efficiency will decrease. On the other hand, the shot noise will increase. Both the effects would lead to deterioration of short-term stability.

34.4 Discussion

Both theoretical and experimental investigations were performed on the influence of lamp spectral profile on short-term stability and light shift. It was found that serious distortion of the lamp spectral profile could lead to one times deterioration to short-term frequency stability. Besides, variation of the lamp spectral profile could change the zero light shift point by a few degrees, which indicates that, in addition to the cavity-cell assembly, the lamp should be taken into consideration when setting the zero light shift point.

We noted that the lamp (L3) with the lowest distortion profile showed a significant advantage in short-term stability compared with the other two lamps (L1 and L2), while the corresponding zero light shift point of L3 was lower than the appropriate operation temperature of absorption cell in a magnitude of 10 degrees.

Given that a filter cell with buffer gas Ar of 5.3 kPa was used in the experiment, the zero light shift could be raised by increasing the buffer gas pressure. We have been preliminarily demonstrated that not only the zero light shift point could be raised but also the filtering efficiency could be improved by this approach.

References

1. Cowan RD, Dieke GH (1948) Self-absorption of spectrum lines. *Rev Mod Phys* 20(2):418–455
2. Volk CH, Frueholz RP (1985) The role of long term lamp fluctuations in the random walk of frequency behavior of the rubidium frequency standard: a case study. *J Appl Phys* 57(3):980–983
3. Happer W, Mathur BS (1967) Effective operator formalism in optical pumping. *Phys Rev* 163(1):12–25
4. Mathur BS, Tang HY, Happer W (1970) Light propagation in optically pumped alkali vapors. *Phys Rev A* 2(3):648–660
5. Mathur BS, Tang H, Happer W (1968) Light shifts in the alkali atoms. *Phys Rev* 171(1):11–20
6. Missout G, Vanier J (1975) Some aspects of the theory of passive rubidium frequency standards. *Can J Phys* 53(11):1030–1043
7. Vanier J, Bernier LG (1981) On the signal-to-noise ratio and short-term stability of passive rubidium frequency standards. *IEEE Trans Instrum Meas* IM-30(4):277–282
8. Camparo J, Frueholz R (1986) A nonempirical model of the gas-cell atomic frequency standard. *J Appl Phys* 59(2):301–312
9. Zhong D, Xia B, An S, Wu H, Wang Q, Zhao F, Qi F, Wang F, Mei G (2009) Investigation on physics package with slotted-tube microwave cavity for rubidium atomic frequency standard. In: *Proceedings of the EFTF-IFCS-2009 joint conference*, pp 1019–1022

Chapter 35

Evaluation and Monitoring on the Single Station Time Difference Based on the BDS, GPS and GLONASS Data

Guang Sun, Hua Lu, Lirong Shen, Xiaolin Jia, Meijun Guo,
Yijun Mo and Yingjie Hong

Abstract With the improvement and updating of GPS, GLONASS navigation system, and the gradual establishment of GALILEO and China's BeiDou system (BDS) navigation system, Multi-system GNSS navigation systems have become the main research directions. In order to achieve the compatibility and interoperability among the GNSS systems, accurately determine the time difference between different systems is the key. Based on this, this paper adopted the GNSS observation data, which is the output of the multi-mode dual-band receiver, and the GNSS system's navigation message to monitor the time difference between different GNSS navigation systems. Because the BIPM T bulletin has no BDS data, so the measured time difference data of GLONASS and GPS and the corresponding time difference data of BIPM T bulletin were used to be evaluated. In this paper, the maximum, minimum, mean, mean square error and the RMSE of BDS, GPS and GLONASS time difference data were accumulated. 310 days' time difference data were used to evaluation the difference. The results show that the residual' standard deviation between GLONASS, GPS time difference and the results of T bulletin is 4.28 ns, and the residual' standard deviation between BDS and GPS, GLONASS system can reach 5 and 10 ns in optimal conditions.

Keywords System time difference · BDS · GPS · GLONASS · BIPM

G. Sun (✉) · H. Lu · M. Guo · Y. Mo · Y. Hong
Xi'an Aerors Data Technology Co. Ltd, Xi'an 710054, China
e-mail: sunguang216@gmail.com

L. Shen
School of Aerospace Science and Technology, Xidian University, Xi'an 710071, China

X. Jia
Xi'an Research Institute of Surveying and Mapping, Xi'an 710054, China

35.1 Introduction

China's BeiDou system (BDS), which can provide services to the Asia-Pacific, is compatible and interoperable with GPS, the EU's Galileo system and Russia's GLONASS. At present, Multi-mode satellite navigation has become a research hotspot. Multi-mode satellite navigation system is the navigation systems which utilize many navigation systems for joint navigation, it can realizes the advantageous complementarities and can improve the navigation precision to some extent [1].

Different navigation system has its own time. GPST is the time of GPS system, which can be traced back to the Universal Time Coordinated [UTC (USNO)] of the United States Naval Observatory. GLONASST is the time of GLONASS system, which can be traced back to the Universal Time Coordinated [UTC (SU)] of Russia's. GST is the time of Galileo system [2–5], which is the integrated computation result of several UTC time in different laboratory of European. BDT is the time of BDS, this time system is produced by the clocks of the BDS ground control station, which can be traced back to UTC (NTSC) [6]. There will be time deviation between different systems, the basis of the multi-mode satellite navigation is unified on different system time.

Due to the system time deviation will affect the precision of positioning, velocity determining and timing, and can lead to large navigation error in a multi-mode navigation. So it is necessary to do the research of time difference systems. There are two main types of method for the system time deviation monitoring. They are the system level method and the user level method, respectively [7].

In this paper, use the system level monitoring method for the system time deviation monitoring. Firstly, receive pseudo range by multimode receiver, and then calculate the time difference between local time with BDS, GLONASS and GPS system. At last, the performance of the time difference of the three systems is evaluated.

35.2 GNSS Time Difference Monitoring Principle

In this paper, the method of spatial signal was used to monitor the system time difference through the single-mode or multimode receiver in a single station to get the space signal of the GNSS system, then using the navigation positioning principle for the calculating of the system time deviation. Monitoring principle is shown in Fig. 35.1, and relevant algorithm can be used to calculate the time difference of the three systems [7–9].

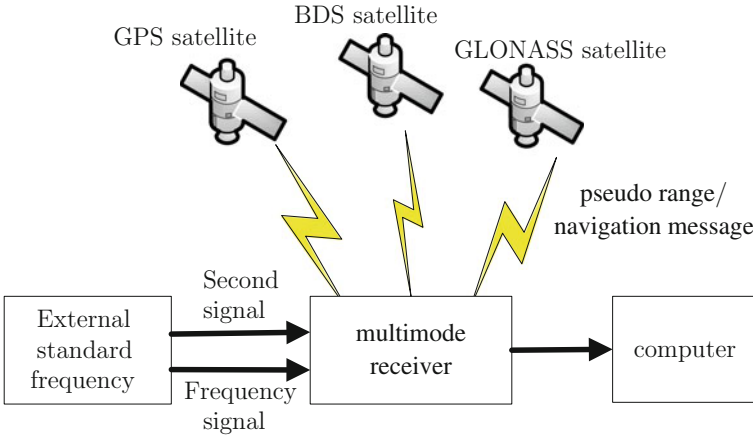


Fig. 35.1 Principle of GNSS time difference monitoring

Pseudo range observation equation is as follows.

$$P = \rho + cdt_r - cdt_s + d_{trop} + d_{ion} + d_{mult} + d_{cha} + d_{ode} + \varepsilon \quad (35.1)$$

P is the observation quantity of pseudo range, ρ is the distance of the station and the satellite, cdt_r is clock correction of the receiver, cdt_s is the clock correction of the satellite, d_{orb} is the orbit biases of the satellite, d_{trop} is the biases of the troposphere, d_{ion} is the biases of the ionosphere, d_{mult} is the time delay of the multipath, d_{cha} is the time delay of the receiver channel, d_{ode} is time delay of the cable refer, ε is the measurement noise of the pseudo range. Therefore, the receiver time difference can be calculated by using the pseudo range observation quantity and navigation message, as shown in formula 35.2.

$$dt_r = \frac{p - \rho}{c} + dt_s - \frac{d_{trop} + d_{ion}}{c} - \frac{d_{cha} + d_{ode}}{c} - \frac{d_{mult} + d_{orb} + \varepsilon}{c} \quad (35.2)$$

The clock error of the receiver can be determined through the navigation message and pseudo range observation quantity. Assume the clock error of GPS is $REFG$, the clock error of GLONASS is $REFR$, and the clock error of BDS is $REFC$, using formula 35.3, the Time difference between two systems can be determined.

$$\begin{cases} T_{CG} = REFG - REFC \\ T_{CR} = REFR - REFC \\ T_{RG} = REFG - REFR \end{cases} \quad (35.3)$$

T_{CG} is the time difference of BDT and GPST, T_{CR} is the time difference of BDT and GLONASST, T_{RG} is the time difference of GLONASST and GPST.

35.3 Error Model of the GNSS Time Difference Monitoring

The main error terms of the GNSS system time difference monitoring include satellite clock error, satellite ephemeris error, relativistic effects, the earth's rotation effect, ionosphere delay, tropospheric delay, cable delay and multipath effect.

- Satellite clock error

As for BDS, satellite clock error at the time of t in BDT [6] can be determined in formula 35.4.

$$\Delta t_{sv} = a_0 + a_1(t - t_{oc}) + a_2(t - t_{oc})^2 \quad (35.4)$$

a_0 is the clock phase deviation of satellite at the initial time, a_1 is the clock frequency deviation of satellite at the initial time, a_2 is the clock frequency drift of satellite, t is the BDT at the time of signal emission, t_{oc} is the reference time of the satellite clock data.

The BII signal also needs to be a further correction before it can be used. The correction can be conducted as shown in formula 35.5.

$$(\Delta t_{sv})_{BII} = \Delta t_{sv} - T_{gd1} \quad (35.5)$$

T_{gd1} is the time delay between the star equipment's and can be obtained by satellite navigation message.

- Satellite ephemeris error

The correction of satellite ephemeris error is the same as the correction of the satellite clock error. Firstly, fitting the optimal estimation of the satellite position estimated by the master control station, and then uploaded the correct satellite ephemeris to the satellite. Finally, broadcast the correct satellite ephemeris to the user in the form of navigation message [7].

- Relativistic effects

Because of the motion state and the stress of the satellite clock and the ground clock's is different, there will be a relative deviation between the satellite clock and the ground clock. The relativistic effects can be corrected by the formula 35.6 shown as follows.

$$\Delta t_r = F \times e \times \sqrt{A} \times \sin E \quad (35.6)$$

e is the satellite orbital eccentricity, A is the satellite orbital semi-major axis. E is the eccentric anomaly calculated by the ephemeris parameter data, F is a constant, $F = -4.442807633 \times 10^{-10} \text{ s}/\sqrt{\text{m}}$.

- The earth's rotation effect

The rotation of the earth will lead to a relativity error of the satellite position at the moment of emission and receiving time, consider the Sagnac effect. Formula 35.7 can be used to correct the satellite coordinates.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos(\omega\tau) & \sin(\omega\tau) & 0 \\ -\sin(\omega\tau) & \cos(\omega\tau) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{sat} \\ y_{sat} \\ z_{sat} \end{bmatrix} \quad (35.7)$$

$[x, y, z]$ is the revised satellite location, ω is the angular velocity of earth spinning, τ is the travel time of the satellite signals in the space.

- Ionosphere delay

The combination of dual-frequency pseudo range observation can be used to eliminate the ionosphere delay, as shown in formula 35.8.

$$\rho = \frac{f_1^2 p_1 - f_2^2 p_2}{f_1^2 - f_2^2} \quad (35.8)$$

p_1, p_2 are the pseudo range observation, f_1, f_2 are the frequency

- Tropospheric delay

The tropospheric delay can be estimated with the model which includes the meteorological parameters. The commonly used models are the Hopfield model and the Saastamoinen model. The Hopfield model is adopted in this paper to correct the tropospheric delay.

35.4 The Experiment Results and Discussions

35.4.1 The Time Differences Result of Measured Data

In this paper, the test data comes from Trimble R9 multimode receiver. It can receive related data of the four systems. Multimode receiver located at the monitoring station connects the 1 PPS/10 MHZ external signal of the reference master clock, select the related data of the BDS, GPS and GLONASS received in the eight months of 2014, the above error models are adopted to correct the errors, using pseudo range observation and navigation message to obtain the time differences of the GPST, the GLONASST and the BDT, the measured time interval is 15 min, eliminate the gross error of the time difference data with the median method, and use vondrak to filtering the data [10, 11]. Figures 35.2, 35.3 and 35.4 shows the GNSS navigation system time difference data waveform over the March to November of 2014. Figure 35.2 is the time difference waveform of the BDT with GPST, Fig. 35.3 is the time difference waveform of the GLONASST with BDT, Fig. 35.4 is the time difference waveform of the GLONASST and GPST, the red line is the filtered graph.

Figure 35.2 shows that the time difference of the BDT and GPST changes between 40 and 110 ns. In Figs. 35.3 and 35.4, the time difference of the GLONASST and GPST, GLONASST and BDT changes between 450 and 520 ns over March to August of 2014, waveform floating range become large over September to November of 2014.

Fig. 35.2 Time difference between BDT and GPST in 2014.3–2014.11

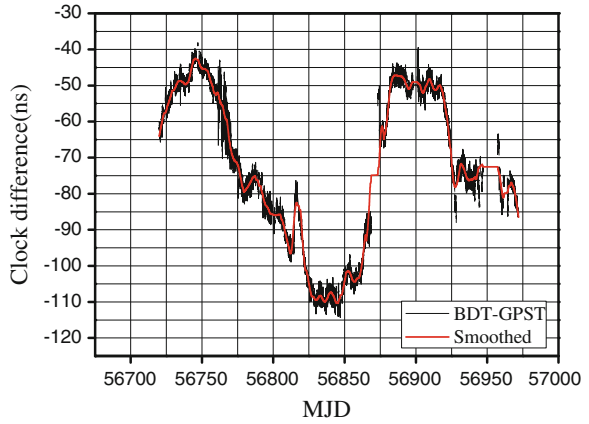


Fig. 35.3 Time difference between BDT and GLONASST in 2014.3–2014.11

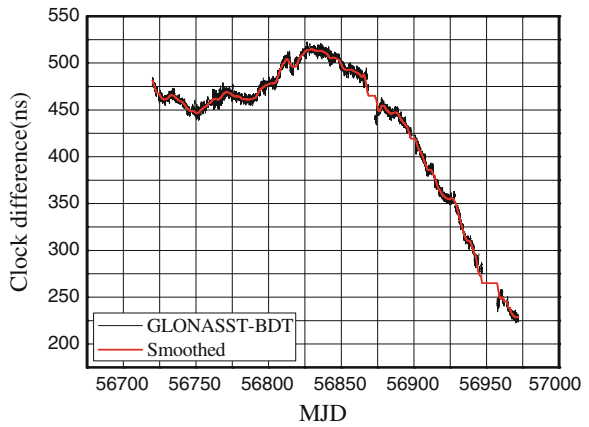
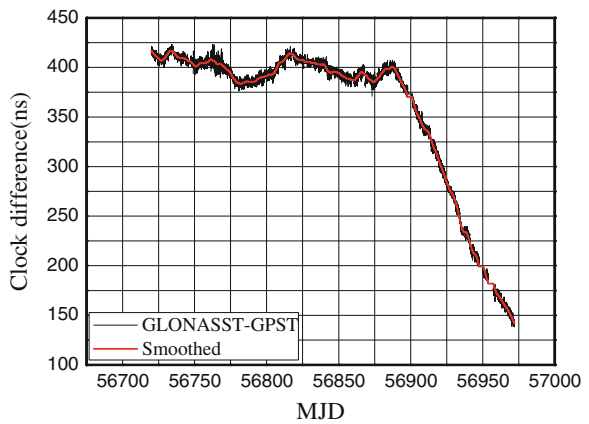


Fig. 35.4 Time difference between GLONASST and GPST in 2014.3–2014.11



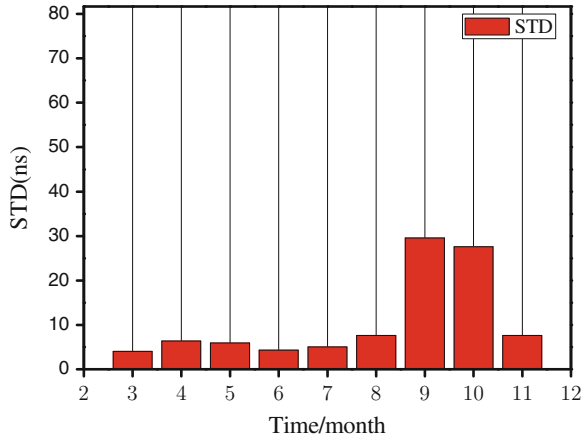
35.4.2 Statistics Analysis of the Measured Data

The statistical results of test data is shown in Table 35.1, RG stands for GLONASST–GPST, CG stands for BDT–GPST, RC stands for GLONASST–BDT, the maximum, minimum, average, standard deviation, root mean square error and 95 % confidence level of the two systems' time difference are calculated over the March to November of 2014 for BDT, GLONASST, GPST.

Table 35.1 The statistical result of the time difference between GLONASST and GPST in 2014.3–2014.11

Time	Time difference	Max (ns)	Min (ns)	Mean (ns)	STD (ns)	RMS (ns)	% 95 upper (ns)
2014/3	RG	423.25	399.94	411.02	4.04	411.04	418.08
	CG	-38.16	-65.64	-50.71	5.75	51.04	-42.21
	RC	484.03	443.37	461.62	7.71	461.68	476.03
2014/4	RG	422.98	378.77	401.78	6.36	401.84	411.19
	CG	-40.48	-81.22	-57.73	10.83	58.74	-44.24
	RC	478.96	439.1	459.34	8.18	459.41	470.7
2014/5	RG	405.92	376.3	389.54	5.92	389.58	402.02
	CG	-71.3	-95.41	-82.09	5.13	82.25	-74.6
	RC	498.23	454.37	471.56	10.27	471.68	491.17
2014/6	RG	420.98	394.56	407.10	4.3	407.13	414.7
	CG	-76.37	-113.4	100.58	9.58	101.03	-80.69
	RC	522.35	486.79	507.50	7.37	507.55	516.64
2014/7	RG	410.04	380.33	393.23	5.04	393.26	402.07
	CG	-84.85	114.32	102.03	6.74	102.27	-88.43
	RC	516.86	470.93	495.38	9.06	495.47	510.13
2014/8	RG	407.31	370.19	390.80	7.6	390.87	402.17
	CG	-43.72	-71.41	-52.81	6.5	53.21	-46.04
	RC	463.22	420.32	443.96	8.8	444.04	456.26
2014/9	RG	373.6	262.51	317.34	29.59	318.72	361.04
	CG	-39.51	-88.03	-57.25	10.57	58.21	-47.24
	RC	422.96	337.41	374.82	21.51	375.43	410.96
2014/10	RG	269.17	163.61	213.13	27.6	214.91	257.01
	CG	-63.49	-85.83	-75.18	3.85	75.28	-68.53
	RC	343.82	234.65	294.05	28.38	295.42	331.5
2014/11	RG	172.37	138.34	155.61	7.59	155.80	166.62
	CG	-73.58	-86.51	-79.81	3.18	79.88	-75.63
	RC	251.03	222.71	232.66	5.49	232.72	244.38
2014/3-11	RG	423.25	138.34	357.83	75.61	365.73	413.23
	CG	-38.16	-114.3	-72.55	21.03	75.54	-45.9
	RC	522.4	222.71	435.72	72.88	441.77	511.8

Fig. 35.5 STD of time difference between GLONASST and GPST in 2014.3–2014.11



The test data monthly variation of the navigation system is described using histogram. Y-axis is the standard deviation of the test data. Figures 35.5, 35.6 and 35.7 show that the standard deviation is better than 10 ns of the GLONASST relative to the GPST and the BDT over March to August. Since the GLONASST data has a downward trend changes in September, the standard deviation increases to 30 ns of the GLONASST relative to the GPST and BDT. In November, because the data quantity is become less, so the standard deviation is smaller. Figure 35.6 shows that the standard deviation of test data concluded BDT and GPST are superior to 15 ns, and the standard deviation can reach 5 ns.

Fig. 35.6 STD of time difference between BDT and GLONASST in 2014.3–2014.11

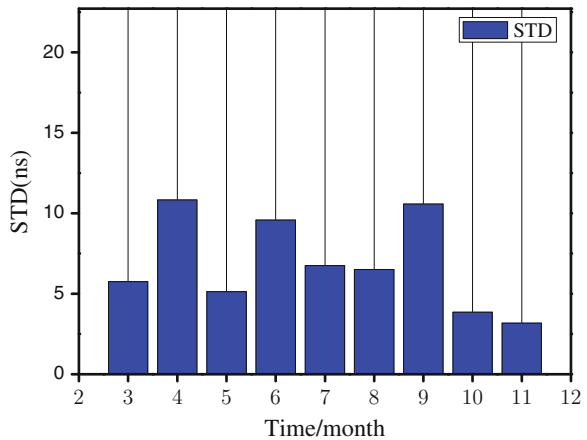
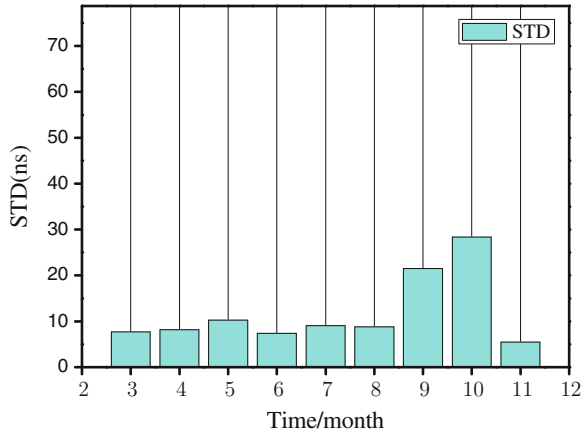


Fig. 35.7 STD of time difference between GLONASST and GPST in 2014.3–2014.11



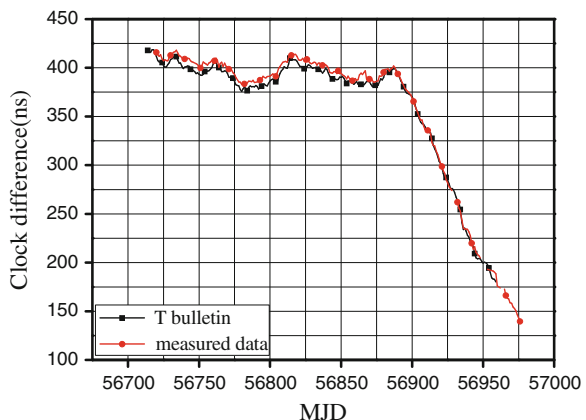
35.4.3 Compare the Measured Data with the T Bulletin Data in BIPM

T bulletin data in 2014 is downloaded in BIPM web site, a set of UTC–GNSST data is included every day. At present, T bulletin without the BDT data, this paper only compare the time difference of the GLONASST–GPST between T bulletin data and the measured data, the operation steps are shown as follows.

- Use the difference of the UTC–GLONASST and UTC–GPST in the fifth part of T bulletin, GLONASST–GPST can be obtained.
- Calculate the biases of time difference from T bulletin and the time difference from the measured data, evaluate the measured GLONASST–GPST with the true value of T bulletin.

Figure 35.8 shows the time difference of the GLONASST–GPST between the true value of the T bulletin and the measured data over March to November of 2014.

Fig. 35.8 Time difference of the measured data and the data from BIPM T bulletin in 2014.3–2014.11



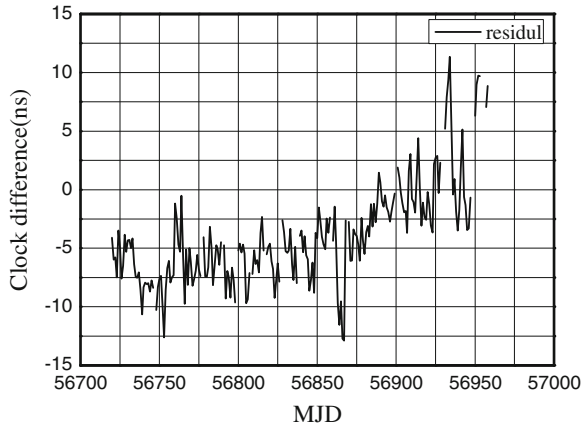


Fig. 35.9 The time difference residual between the measured data and the data from BIPM T bulletin in 2014.3–2014.11

Table 35.2 The statistical result of the residual between the measured data and the data from BIPM T bulletin in 2014.3–2014.11

Statistical result	Max (ns)	Min (ns)	Mean (ns)	STD (ns)
residual	11.35	-12.9	-3.98	4.28

The red square line with triangle symbol show the measured values, the black line with square symbol shows the true value of the T bulletin, Fig. 35.9 shows the residual between the T bulletin true value and the measured data, Table 35.2 shows the parameters of residual error. From the Table 35.2 we can see that the monitoring results of the GLONASST–GPST over 310 days. Compared with the results of T bulletin, the maximum absolute residual value is less than 13 ns, the standard deviation is better than that of 5 ns.

35.5 Conclusions

This paper introduces the principle of GNSS time difference monitoring and the error correction model, analyzed the various measures of the BDT, GPST GLONASST three time difference over March to November of 2014, the analysis results show that BDT and GPST GLONASST annual standard deviation is better than that of 5 ns, 10 ns respectively. Using the T bulletin time difference to assess the results of the measured GLONASST–GPST time difference, the evaluation results show that compared with the results of T bulletin, the maximum absolute residual value of the measured data is less than 13 ns, and the standard deviation is

better than that of 5 ns. The result verified the effectiveness and the high precision of the proposed method.

Time monitoring is a very meaningful work, it play an important role in multi-mode navigation. In this article, the time difference algorithm is based on pseudo range monitoring, in order to improve the accuracy of time difference monitoring results, the smooth pseudo distance of carrier phase or un-differenced phase data can be used.

Acknowledgments The author would like to thank the Monitoring and Assessment Center (MAC) of International GNSS Monitoring and Assessment System (iGMAS) for their help in acquiring data.

References

1. Wu HT, Li XH et al (2011) Time base of the satellite navigation system. Science press, Beijing
2. Hahn J, Powers E (2007) A report on GPS and galileo time offset coordination efforts. In: European frequency and time forum, Switzerland, May 2007
3. Galluzzo G, Mudrak A et al (2010) GGTO and UTC dissemination results in the GIOVE-mission. In: European frequency and time forum, Netherlands, 13–16 Apr 2010
4. Allan DW, Thomas C (1993) Technical directives for standardization of GPS time receiver software. Metrologia, USA 31:69–79
5. Hahn J, Powers E (2005) Implementation of the GPS to galileo time offset. In: Proceedings of the 2005 joint IEEE international frequency control symposium and precise time and time interval systems and applications meeting. Vancouver, Canada, pp 33–212
6. Zhu L (2013) Research on one-station and multi-station combined monitoring method of GNSS system time offset. In: National time service center, Chinese Academy of Sciences, Xi'an, May 2013
7. Li XR (2013) Study on time difference monitoring and forecasting between GPST and GLONASST with single monitoring station. In: National time service center, Chinese Academy of Sciences, Xi'an, May 2013
8. Li XR, Yuan HB, Dong SW (2013) Study on the time difference between GPST and GLONASST with single monitoring station. J Astronaut Metrol Meas 32(6):61–64
9. Zhang HJ, Li XH, Xu LX (2010) Time difference monitoring and forecasting of GNSS system. In: Proceedings of the 1st China satellite navigation conference (CSNC), Beijing, China, May 2010
10. Ding YR (1990) Astronomical data processing method. Nanjing University Press, Nanjing
11. Tian L, Chen JP et al (2012) GNSS time offset and effects on multi-GNSS positioning performance. Bull Surv Mapp 2012(10):45–47

Chapter 36

Prediction of Navigation Satellite Clock Bias by Gaussian Process Regression

Yu Lei, Danning Zhao, Zhaopeng Hu and Hongbing Cai

Abstract Many studies have been carried out in the past for forecasting satellite clock bias utilizing models such as the grey model, linear model, quadratic polynomial model, etc., but the accuracy of these models has not met the requirements for real-time applications. One reason for the fact is that onboard atomic clocks can be easily affected by various factors such as environment and temperature and this leads to complex aspects like periodic and stochastic variations, which are not sufficiently described by conventional models. A hybrid prediction model is thus developed in this work in order to be used particularly in describing the stochastic variation behavior satisfactorily. The proposed hybrid prediction model for satellite clock bias combines the quadratic model plus harmonic model to overcome the linear and periodic effects, and Gaussian process regression (GPR), whose input is reconstructed by the delay coordinate embedding to access linear or nonlinear coupling characteristics. The simulation results have demonstrated that the prediction accuracy of the proposed model is better than that of the IGS ultra-predicted (IGU-P) solutions at least on a daily basis.

Keywords Satellite clock bias · Prediction model · Gaussian process regression (GPR) · Phase space reconstruction

36.1 Introduction

Satellite clock bias is one of main error sources to be eliminated in GNSS positioning. Taking GPS for example, the accuracy of the final GPS clock products provided by the International GNSS Service (IGS) has been shown to be roughly

Y. Lei (✉) · D. Zhao · Z. Hu · H. Cai
National Time Service Center, Chinese Academy of Sciences, Xi'an 710600, China
e-mail: leiyu@ntsc.ac.cn

Y. Lei · D. Zhao · Z. Hu
University of Chinese Academy of Sciences, Beijing 100049, China

75 ps root-mean-squares (RMS), which can achieve centimetre precision in real-time precise point positioning (RTPPP). While the final clock products are of very high-quality, they are not available for users in real-time GPS application because they are made available about 13 days after the end of each GPS week. The performance of clock prediction based on the broadcast ephemeris and the IGS ultra-rapid predicted (IGU-P) products is about 5 and 3 ns RMS, respectively. Although the two clock solutions are available in real time, they have not shown acceptably high-quality prediction performance. Therefore, it is greatly crucial and significant to enhance the accuracy of clock prediction [1–5].

It is difficult to model and predict the satellite clock behaviour accurately. One reason for the fact is that satellite clocks in flight can be easily affected by various factors like environment and temperature and this leads to complicated aspects such as periodic and stochastic variations, which are not sufficiently described by traditional models. Many prediction methods and techniques were used in the past to improve the prediction accuracy of satellite clock bias, e.g. the linear model, quadratic polynomial model and grey model. These models are appropriate to describe the variations of a stationary clock rather than a non-stationary clock. A polynomial model with periodic items was proposed so as to absorb the cyclic effects, such as by Huang et al. [4]. At present, this method is adopted by the IGS to produce the IGU-P products. Nevertheless, all these models take into account the linear trend and periodic terms of satellite clock bias, but ignore random items [5]. In the consideration of the characteristics of satellite clocks, a hybrid model for clock bias prediction is proposed in this work, which combines the polynomial model with a few harmonic components and Gaussian process regression (GPR). The proposed method not only takes into consideration the linear trend and periodic variations of satellite clocks, but also the stochastic variation behaviour. The experimental results have shown that the developed method can forecast satellite clock bias with high accuracy and efficiency.

A Gaussian process (GP) for machine learning is a generic supervised learning algorithm primarily designed to solve regression problems [6, 7]. GPR is a kind of non-parametric modelling method based on Bayesian learning and it has strong capacity to handle stochastic uncertainty and non-stationary processes. A GPR model can be utilized to formulate a Bayesian regression framework that is ideal for prediction of stochastic and non-stationary processes such as satellite clock bias. Therefore, it is theoretically feasible to apply GPR to prediction of satellite clock bias. In recent years, GPR has been successfully applied to various domains including the navigation fields [8, 9]. In this contribution, the GPR technique is employed for prediction of the random part of satellite clock bias.

36.2 The Physical Characterization of Satellite Clocks

The output of an atomic clock can be expressed as $V(t) = [V_0 + \varepsilon(t)] \sin(2\pi v_0 t + \phi(t))$, where v_0 and V_0 are the nominal frequency and amplitude of the output signal, and $\varepsilon(t)$ and $\phi(t)$ are the random fluctuations of the amplitude and phase, respectively. The normalized phase bias $clk(t) = \phi(t)/(2\pi v_0)$ evolves over time in terms of deterministic and random laws. For example, the phase bias, namely clock bias, is often modelled as the following quadratic polynomial plus harmonic model.

$$clk(t) = a_0 + b_0 t + \frac{1}{2} c t^2 + \sum_{i=1}^n A_i \sin(\omega_i t + \varphi_i) + \psi(t) \quad (36.1)$$

where a_0 and b_0 denote the phase and frequency bias at the initial time t_0 , and c is the linear rate of the frequency bias, referred to as frequency drift or aging usually. In practice, the value of c can be designated as 0 for clocks of the GPS Block IIR, IIR-M or IIR satellites [4]. A_i , ω_i and φ_i are the amplitude, frequency and phase shift of the sinusoidal variations, which can be together estimated by the least-squares (LS) solution with a_0 , b_0 and c , and $\psi(t)$ is a generic stochastic noise process. Note that $\psi(t)$ is always ignored by traditional prediction models.

36.3 Methodology

36.3.1 Main Idea

Satellite clock bias is non-linear and non-stationary discrete time-series. The variations in clock bias are a comprehensive reflection of the interaction with some factors, i.e. temperature. Xu et al. [1, 2] characterized satellite clock bias from the perspective of chaotic analysis, and then established a prediction model for satellite clock bias according to the chaotic theory for the first time. Their study showed that satellite clock bias can be viewed as chaotic time-series, and thus predicted by chaotic methodology.

On the basis of the references [1, 2], this paper proposes a hybrid model which integrates the quadratic polynomial model adding cyclic terms and GPR model. The combined model uses a polynomial model with cyclic terms to extract the linear trend and periodic terms of clock bias, and then employs GPR to model the residuals. Since GPR possesses the wonderful ability of nonlinear mapping and parallel processing, it can be utilized to forecast the non-linear and non-stationary time-series. On the other hand, the differences between the polynomial model and actual clock bias, namely the residuals, have definite internal regularity. This type of information processing method is just what GPR is good at, while it is arduous for conventional analytical algorithms. Wherefore, GPR is applied to model and forecast the residuals.

36.3.2 Gaussian Process Regression

A GP is a collection of random variables, any finite number of which have a joint Gaussian distribution [6, 7]. Given a training set $\{(\mathbf{x}_i, y_i)\}_{i=1}^l$, where $\mathbf{x}_i \in \mathbf{R}^m$ and $y_i \in \mathbf{R}$ denote the input vector and observed output, respectively, each observed value y_i can be thought of as related to an underlying function $f(\mathbf{x}_i)$ through a Gaussian noise model.

$$y_i = f(\mathbf{x}_i) + \varepsilon \quad (36.2)$$

where y_i differs from the function value $f(\mathbf{x}_i)$ by additive Gaussian noise ε with zero mean and variance σ_n^2 . Conditioning on the training set and a test input \mathbf{x}_* , the GP results in a Gaussian predictive distribution over the corresponding output y_* .

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} = N\left(0, \begin{bmatrix} \mathbf{K} + \sigma_n^2 \mathbf{I} & \mathbf{K}_*^T \\ \mathbf{K}_* & \mathbf{K}_{**} \end{bmatrix}\right) \quad (36.3)$$

where \mathbf{I} is the unit matrix; $\mathbf{y} = [y_1, y_2, \dots, y_l]^T$;

$$\mathbf{K} = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & k(\mathbf{x}_1, \mathbf{x}_2) & \cdots & k(\mathbf{x}_1, \mathbf{x}_l) \\ k(\mathbf{x}_2, \mathbf{x}_1) & k(\mathbf{x}_2, \mathbf{x}_2) & \cdots & k(\mathbf{x}_2, \mathbf{x}_l) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{x}_l, \mathbf{x}_1) & k(\mathbf{x}_l, \mathbf{x}_2) & \cdots & k(\mathbf{x}_l, \mathbf{x}_l) \end{bmatrix}; \quad \mathbf{K}_* = [k(\mathbf{x}_*, \mathbf{x}_1), k(\mathbf{x}_*, \mathbf{x}_2), \dots, k(\mathbf{x}_*, \mathbf{x}_l)];$$

and $\mathbf{K}_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$, where $k(\mathbf{x}, \mathbf{x}')$ is a covariance function.

The best estimate for y_* is the mean of this distribution.

$$\bar{y}_* = \mathbf{K}_* \mathbf{K}^{-1} \mathbf{y} \quad (36.4)$$

The uncertainty in the estimate relies on both the process noise and the correlation between the training set and given input. The most widely used covariance function is the squared exponential covariance function with the additive noise [10].

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left[-\frac{1}{2} \left(\frac{\mathbf{x} - \mathbf{x}'}{\boldsymbol{\theta}}\right)^2\right] + \sigma_n^2 \delta \quad (36.5)$$

where σ_f^2 is the signal variance, which tunes up the prediction uncertainty in areas of training data density; $\boldsymbol{\theta}$ are the length scales of the process, reflecting the relative smoothness of the process along the different input dimensions; σ_n^2 regulates the global noise of the process; and δ is the Kronecker function, $\delta = \begin{cases} 1 & \mathbf{x} = \mathbf{x}' \\ 0 & \mathbf{x} \neq \mathbf{x}' \end{cases}$. The parameters $\Theta = \{\sigma_f, \boldsymbol{\theta}, \sigma_n\}$ are referred to as hyper-parameters of the GP. They can be determined by maximizing the log marginal likelihood of the given training set [9].

36.3.3 Generation of Training Patterns for GPR

36.3.3.1 Empirical Mode Decomposition de-Noising

In order to mitigate the influences of measurement noise on the chaotic property, empirical mode decomposition (EMD) proposed by Huang et al. [11] is employed to extract the noise that exists in the residuals.

EMD is a data-driven adaptive approach for signal decomposition, which can separate different frequency components within the signal. Several intrinsic mode functions (IMF) can be obtained after the decomposition. In general, as for a noisy signal some IMF with higher frequencies can be regarded as the noise. The signal can be expressed as follows after the noise is mitigated.

$$\psi'(t) = \psi(t) - \sum_{i=1}^k IMF_i, \quad t = 1, 2, \dots, N \quad (36.6)$$

where $\psi'(t)$ represents the noise-free signal and k is the number of noisy components, which can be determined according to the correction coefficient between the original signal $\psi(t)$ and de-noised signal $\psi'(t)$.

$$\rho_{\psi\psi'} = \frac{\sum_{t=1}^N \psi(t)\psi'(t)}{\sqrt{\sum_{t=1}^N \psi^2(t)}\sqrt{\sum_{t=1}^N \psi'^2(t)}} \quad (1.7)$$

Starting from $k = 1$, if $\rho_{\psi\psi'} > c$, let $k = k + 1$ and then recalculate $\rho_{\psi\psi'}$. Stop this process and then let $k = k - 1$ until $\rho_{\psi\psi'} < c$. In this case, k is designated as the number of noisy components. Note that c is an empirical value. In this paper, c is set to 0.8.

36.3.3.2 Phase Space Reconstruction

The Taken theorem proposed by Taken [12] indicates that the phase space of time-series can be reconstructed by using the delay coordinate method. The reconstructed system is equivalent to the original one in the topological sense, if the embedding dimension m satisfies $m \geq 2D + 1$, where D is the dynamics dimension of the system.

In past yeas, many studies have demonstrated that the reconstruction quality not only relies on the individual selection of embedding dimension m and delay time τ , but the determination of m and τ together. In current work, the C-C algorithm is used to determine m and τ together [13]. This method first determine the τ and

embedding window width τ_w . Then the m can be derived from the equation $\tau_w = (m - 1)\tau$. The reconstructed phase space can be written as follows.

$$\begin{cases} \mathbf{x}_1 = [\psi'(1), \psi'(1 + \tau), \dots, \psi'(1 + (m - 1)\tau)] \\ \mathbf{x}_2 = [\psi'(2), \psi'(2 + \tau), \dots, \psi'(2 + (m - 1)\tau)] \\ \vdots \\ \mathbf{x}_i = [\psi'(i), \psi'(i + \tau), \dots, \psi'(i + (m - 1)\tau)] \\ \vdots \\ \mathbf{x}_l = [\psi'(l), \psi'(l + \tau), \dots, \psi'(l + (m - 1)\tau)] \end{cases} \quad (36.8)$$

36.3.3.3 Chaos Identification

Only if the time-series of clock bias were chaotic, could the chaotic analysis be applied to prediction of clock bias. Hence it is necessary to identify whether the series of clock bias are chaotic or not in advance. Lyapunov exponent is a very key index which reflects the property of chaotic systems. It measures the divergence degree of the neighboring points in phase space. If Lyapunov exponent is positive, the system can be viewed at the chaotic state. The small-size dataset algorithm have high data-utilizing efficiency and reliable for small data set [14]. Thus, the algorithm is used to calculate the largest Lyapunov exponent of the residuals of clock bias.

36.3.4 Modelling the Residuals of Clock Bias Using GPR

The residuals of clock bias can exhibit the complexly dynamical behaviour that the conventional modelling techniques are able to capture effectively. Herein, GPR is adopted to prediction of the residuals, since it provides the possibility of assessing linear or nonlinear characteristics of the coupling between variables, and more importantly, it is one of the fundamental tools in nonlinear unknown system modelling. Given the residuals de-noised by the EMD, $\{\psi'(t), t = 1, 2, \dots, N\}$, the time evolution of the system can be caught by mapping past output items to future outputs $y_i = f(\mathbf{x}_i)$, where $\mathbf{x}_i = [\psi'(i), \psi'(i + \tau), \dots, \psi'(i + (m - 1)\tau)]$ and $y_i = \psi'(i + (m - 1)\tau + 1)$ is an observable output. f can be yielded from the predictive mapping $f: \mathbf{R}^m \rightarrow \mathbf{R}$ as described in Sect. 1.3.2, which is the centerpiece of modelling.

In the process of modelling, GPR analyzes the training data set, which results in a GPR model built. During the prediction stage, let $\mathbf{x}_l = [\psi'(l), \psi'(l + \tau), \dots, \psi'(l + (m - 1)\tau)]$ as the input of the GPR model, then we can get the prediction output $\hat{\psi}'(l + (m - 1)\tau + 1)$. Let $\psi'(l + (m - 1)\tau + 1) = \hat{\psi}'(l + (m - 1)\tau + 1)$, then we can obtain the new time-series $\{\psi'(l + \tau), \psi'(l + 2\tau), \dots, \psi'(l + (m - 1)\tau + 1)\}$ as

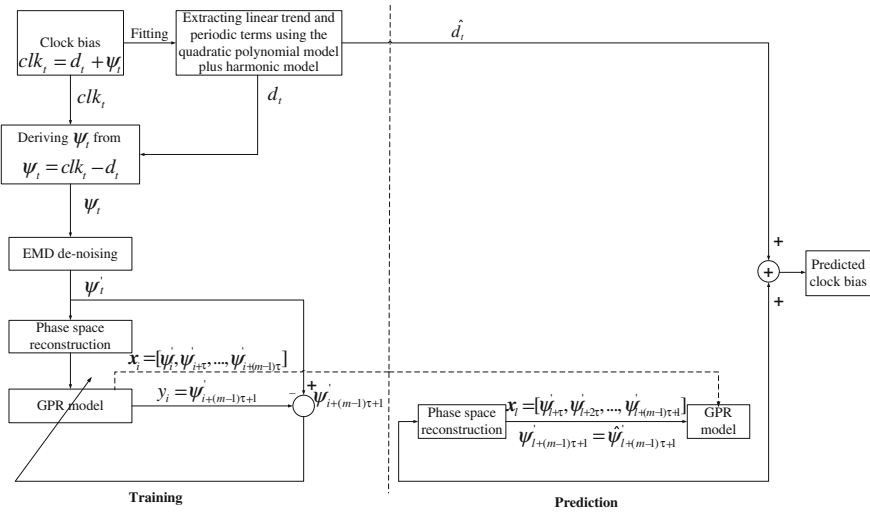


Fig. 36.1 Proposed satellite clock bias modelling strategy

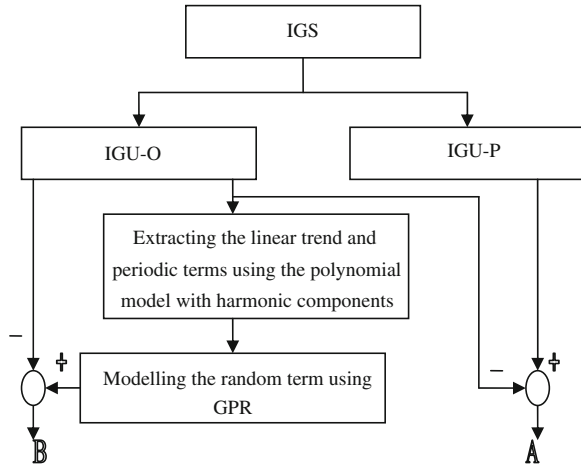
the new input for next moment. In this event, the multi-step ahead prediction for clock bias can be achieved. The developed satellite clock bias modelling strategy is shown in Fig. 36.1.

36.4 Examples

36.4.1 Dataset Selection and Experiment Flow

The IGS ultra-rapid observed (IGU-O) clock bias of 29 Oct 2011 is used as data base to predict the satellite clock bias over the next 24 h, while the IGU-O and IGU-P clock bias of 30 Oct 2011 to check the forecasting capability of the proposed model based the evaluation criteria. The sample interval of these products is 15 min. Five GPS satellites in space are randomly selected for the separate tests, whose clock types include Block IIA rubidium (Rb), IIR rubidium and IIR-M rubidium. The PRN number of the selected satellites is PRN06, PRN19, PRN22, PRN23 and PRN31, respectively. Four separate tests to evaluate the proposed model for clock prediction have been carried out, which are the 3-, 6-, 12-, 24 h-ahead prediction tests. The schematic diagram of the performance test and evaluation is illustrated in Fig. 36.2, where A and B represent the prediction error of IGU-P and the developed algorithm, respectively.

Fig. 36.2 Schematic diagram of the performance test and evaluation



36.4.2 Experiment Procedures

The proposed algorithm is implemented in a robust software code (in MATLAB language) and several separate tests have been performed to validate the presented model for clock bias prediction, which are the 3-, 6-, 12-, 24 h-ahead prediction tests. Integrity check is first carried out in order to exclude the anomalies existing in the time-series of clock bias.

After the data preprocessing, the linear trend and cyclic items are removed from the original clock bias by using the quadratic polynomial plus harmonic model, notable components of which are detected by utilizing the spectrum analysis (the first three items are taken). The PRN19 satellite is taken for example, the residuals after extracting the linear trend and cyclic terms and the de-noised residual series are shown in Fig. 36.3. What can be seen from Fig. 36.3 is that the residuals vary rapidly. This leads to difficult prediction. We compute the optimal delay time τ and embedding dimension m , and then reconstruct the phase space of the de-noised residuals. The $\Delta\bar{S}(\tau) \sim \tau$ and $S_{\text{cor}}(\tau) \sim \tau$ curves corresponding to the PRN19 satellite are illustrated in Fig. 36.4. The value of optimal delay time τ corresponds to the first local minimum point of $\Delta\bar{S}(\tau)$. Also, the optimal embedding window width $\tau_w = (m - 1)\tau$ can be found, which corresponds to the minimum point of $S_{\text{cor}}(\tau)$. From Fig. 36.4, it can be concluded that the optimal delay time τ and embedding dimension m are 3 and 5, respectively. On the basis, the largest Lyapunov exponent 0.26 is gained, indicating that the residual series are at chaotic state.

Once the chaotic characteristics of the residuals are validated, the training patterns can be formed in terms of the reconstructed phase space, and then can be used to training GPR. In this paper, the squared exponential covariance function is selected as the covariance function of GPR, the hyper-parameters of which are determined by the scaled conjugate gradient descent [9]. The subsequently predicted residuals are then added to the polynomial model with cyclic terms in order

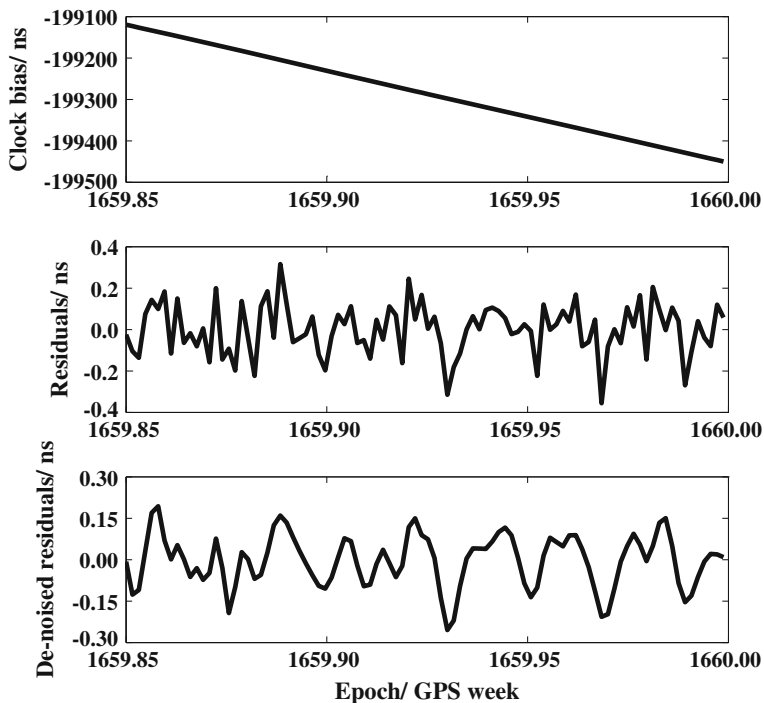
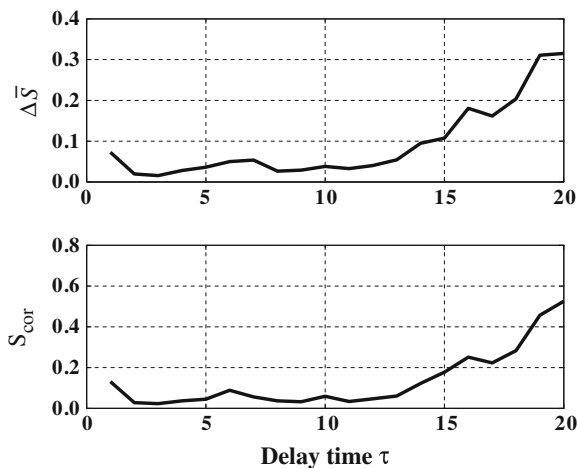


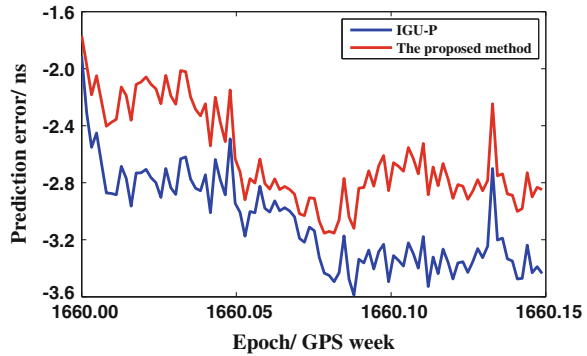
Fig. 36.3 Residuals extraction and EMD de-noising for the PRN19' clock bias

Fig. 36.4 Phase space reconstruction diagram of the residuals of the PRN19' clock bias



to obtain the predicted values of the clock bias. The differences between the predicted results and IGU-O solutions for the PRN19' clock bias are given in Fig. 36.5, along with the differences between the IGU-P and IGU-O products.

Fig. 36.5 Comparison between the prediction error of the proposed method and IGU-P solutions for the PRN19' clock bias



Similar to what has been performed for prediction of the PRN19' clock bias, we can forecast the clock bias of the other four satellites by using the same method. The prediction error of the other four clocks is summarized in Fig. 36.6 and Table 36.1, where the RMS error has been listed for the 3-, 6-, 12- and 24 h-ahead prediction, respectively.

36.4.3 Analysis and Discussion

What can be concluded from Figs. 36.5, 36.6 and Table 36.1 is summarized as follows.

1. The prediction error curves of the five GPS satellite clocks converge within 5 ns, and the prediction precision of the proposed model is within 3.5 ns RMS.
2. As can be observed from Figs. 36.5 and 36.6, the 3- and 6 h-ahead prediction consequences are very similar between the proposed model and IGU-P products, but the IGU-P starts to diverge afterwards.
3. As can be seen in Table 36.1, the 3 and 6 h-ahead prediction accuracy of the developed model is comparable with that of the IGU-P products. However, the presented model can give better performance than the IGU-P solutions for the 12- and 24 h-ahead prediction, and the most significant improvement is 81.29 % as far as the RMS is concerned. This demonstrates that the stochastic residuals can be accurately modelled and predicted to some extent. Note that the improvement extent is different among the five satellite clocks. The possible reason is that the characteristics the atomic clocks in flight are different, such as the stochastic variation behavior.

Thanks to the high computational efficiency of GPR, it takes the proposed method less than six minutes to train the GPR model and predict the clock bias, which can make the proposed algorithm available for real-time prediction of clock

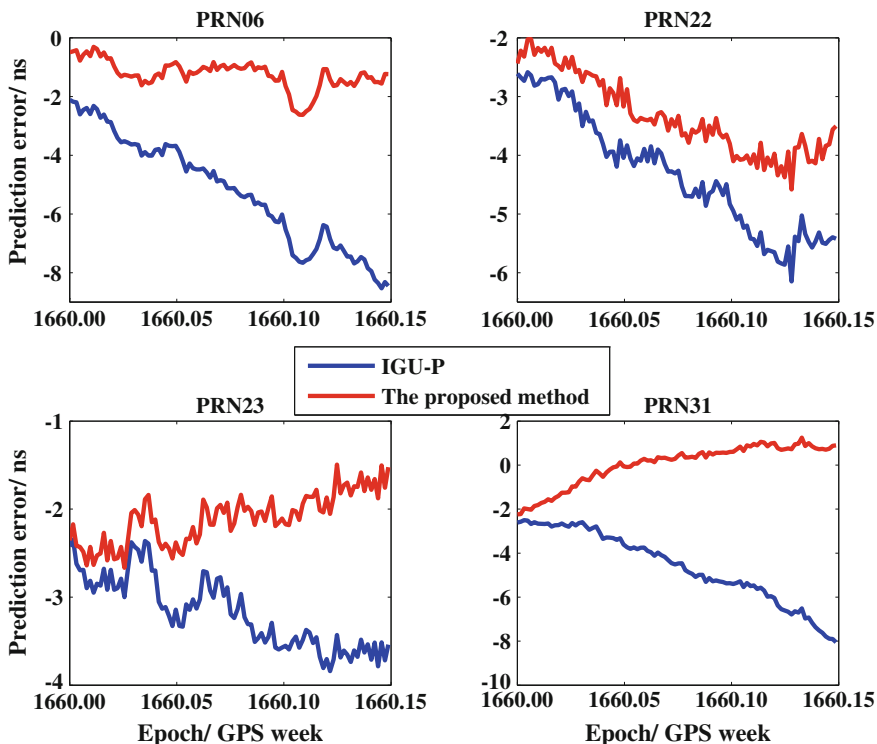


Fig. 36.6 Comparison between the prediction error of the proposed method and IGU-P solutions for the PRN06', PRN22', PRN23' and PRN31' clock bias, respectively

Table 36.1 Comparison between the prediction accuracy of the proposed algorithm and IGU-P solutions for the five GPS satellite clocks (A: IGU-P, B: the proposed algorithm)

PRN	Clock type	3 h		6 h		12 h		24 h	
		A	B	A	B	A	B	A	B
		RMS/ns		RMS/ns		RMS/ns		RMS/ns	
06	IIR Rb	2.47	0.55	3.08	0.99	3.77	1.06	5.65	1.34
19	IIR Rb	2.65	2.18	2.71	2.16	2.84	2.45	3.11	2.64
22	IIR Rb	2.70	2.24	2.96	2.40	3.54	2.84	4.46	3.40
23	IIR Rb	2.74	2.46	2.69	2.36	2.86	2.31	3.22	2.11
31	IIR-M Rb	2.64	1.86	2.69	1.48	3.31	1.06	4.97	0.93

bias. Moreover, the developed prediction model has some advantages, e.g. low complexity, easy use and adaptive computation. These advantages can avoid the influences of human subjectivity on predicted results, and thus enhance the prediction reliability.

36.5 Conclusions

Many studies have been performed in the past for predicting satellite clock bias using models such as the grey model, linear model, quadratic polynomial model, etc. However, the accuracy of these models has not met the requirements for real-time applications. The aim of this work is, accordingly, to develop a satellite clock bias prediction algorithm which results in higher accuracy than any conventional prediction solutions available for at least 24 h-ahead prediction. The developed prediction strategy integrates the quadratic polynomial model plus harmonic model so as to overcome the linear and periodic effects, and GPR in order to capture stochastic variation behavior, whose input is reconstructed by the delay coordinate embedding to access linear or non-linear coupling characteristics. The proposed prediction scheme generally takes into considerations both the physical characteristics of atomic clocks in space and the objective laws which are computed from stochastic residuals. Therefore, it can avoid the human subjectivity and enhance the prediction quality. The experimental results have shown that the prediction accuracy of the presented method is better than that of the IGU-P solutions at a daily basis, demonstrating that, in particular, the stochastic activities from satellite clocks can be captured. Further study will be focused on the improvement of the proposed prediction model, specially, the stochastic variation behavior modeling for the enhancement of long-term prediction.

References

1. Xu B, Wang Y, Yang XH (2013) A hybrid model for navigation satellite clock error prediction. *Comput Intell* 465:307–316
2. Xu B, Wang Y, Yang XH (2013) Navigation satellite clock error prediction based on functional network. *Neural Process Lett* 38(2):305–320
3. Li XY, Dong XR, Zheng K et al (2013) Research of satellite clock error prediction based on RBF neural network and ARMA model. In: Chinese satellite navigation conference (CSNC) 2013 Proceedings: lecture notes in electrical engineering, vol 245, pp 325–334
4. Huang GW, Zhang Q, Xu GC (2014) Real-time clock offset prediction with an improved model. *GPS Solut* 18(1):95–104
5. Lei Y, Hu ZP, Zhao DN (2014) A novel method for navigation satellite clock bias prediction considering stochastic variation behavior. In: Chinese satellite navigation conference (CSNC) 2014 Proceedings: lecture notes in electrical engineering, vol 305, pp 369–379
6. Seeger M (2004) Gaussian processes for machine learning. *Int J Neural Syst* 14(2):69–106
7. Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. MIT Press, Cambridge
8. Atia MM, Noureldin A, Korenberg M (2012) Enhanced Kalman filter for RISS/GPS integrated navigation using Gaussian process regression. In: Proceedings of the 2012 international technical meeting of the institute of navigation, pp 1148–1156
9. Chen HM, Cheng XH, Wang HP et al (2014) Dealing with observation outages within navigation data using Gaussian process regression. *J Navig* 67:603–615

10. Ko J, Klein D, Fox D, et al (2007) GP-UKF: unscented Kalman filters with Gaussian process prediction and observation models. In Proceedings of the IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 1901–1907
11. Huang NE, Shen Z, Long SR (1998) The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc A* 454:903–995
12. Takens F (1981) Detecting strange attractors in turbulence. *Lect Notes Math* 898:366–381
13. Kim HS, Eykholt R, Salas JD (1999) Nonlinear dynamics, delay times, and embedding windows. *Phys D: Nonlinear Phenom* 27(1):48–60
14. Rosenstein MT, Collins JJ, De Luca CJ (1993) A practical method for calculating largest Lyapunov exponents from small data sets. *Phys D: Nonlinear Phenom* 65:117–134

Part III
PNT System and New Technologies
of Navigation

Chapter 37

Distributed GNSS Collaborative Positioning Algorithms and Performance Analysis

Bin Huang, Zheng Yao, Xiaowei Cui and Mingquan Lu

Abstract In global navigation satellite system (GNSS) collaborative positioning, each user shares its processed data with the neighboring users and conducts joint data processing. The study of the performance for distributed GNSS collaborative positioning algorithms is essential for the actual applications. In this paper, the Distributed Least Squares (DLS), Distributed Extended Kalman Filter (DEKF) and Distributed Unscented Kalman Filter (DUKF) algorithms are introduced in detail. After that, the simulation tests are carried out to analyze the performance in the outdoor and partially blocked scenarios respectively. The results show that the performance of three distributed algorithms is all better than GNSS standalone Least Squares algorithm, especially in the partially blocked scenario, the performance improvements are significant. Compared with the other two algorithms, the DEKF algorithm is pretty good in terms of performance and computational complexity. Besides, considering that the computational complexity of users is limited, the results also provide useful information of how to select collaborative users for a better positioning accuracy in different scenarios.

Keywords Distributed · GNSS · Collaborative positioning · Kalman filter

37.1 Introduction

With the development of wireless communication and the increase of the number of global navigation satellite system (GNSS) users in recent years, the GNSS collaborative positioning has been introduced for localization and navigation to improve the accuracy and reliability of position information. By exchanging processed data, each user is helped by its neighbouring users to estimate its position [1]. Collaborative positioning is particularly suitable for harsh GNSS signal

B. Huang · Z. Yao (✉) · X. Cui · M. Lu
Tsinghua University, 1103 Wei Qing Building, Beijing 100084, China
e-mail: yaozheng@tsinghua.edu.cn

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_37

environments. For instance, a motorcade drives through urban canyons, overpasses and tunnels, people locate their positions in a shopping centre and etc.

Distributed collaborative positioning has been widely used in multi-robot localization and wireless sensor network [2, 3]. However, it is difficult to manually set a large network of base stations, and as a result, the coverage area is very limited. For GNSS collaborative positioning, since the satellites are all-weather mobile stations around the world, the issue of positioning coverage which exists in wireless sensor networks has been resolved. The collaboration between GNSS users is first proposed in [4]. However, the algorithms in [4, 5] are centralized, which need a processing centre to combine position-related data among a group of participating entities. It is difficult to meet the demands for most consumer electronics applications. In [6, 7], the distributed collaborative algorithms are presented based on Particle Filter (PF) and Sum Product are introduced. The high computational complexity of the algorithms limits their applications.

In this paper, the observation model and state model are given first. After that, three distributed GNSS collaborative positioning algorithms are introduced in detail, which are Distributed Least Squares (DLS), Distributed Extended Kalman Filter (DEKF) and Distributed Unscented Kalman Filter (DUKF). In addition, simulation tests are carried out with several users in outdoor and partially blocked scenarios respectively. The results show that, compared with GNSS standalone Least Squares algorithm, there is a performance increase for all the three collaborative algorithms. Especially in the partially blocked scenario, the performance improvements are significant. Meanwhile, by contrast with other two algorithms, the DEKF algorithm is pretty good in terms of performance and computational complexity. For the case of limited computational complexity, the results also provide useful information of how to select collaborative users for a better positioning accuracy in different scenarios.

The remainder of this paper is divided into four Sections. The Sect. 37.2 gives the signal and system model of GNSS collaborative positioning. The Sect. 37.3 presents three distributed collaborative positioning algorithms which based on DLS, DEKF and DUKF. The experiment results and discussions are presented in the Sect. 37.4. Finally the conclusions are shown in Sect. 37.5.

37.2 Signal and System Model

A typical example of GNSS multi-user collaborative positioning in challenging environment is presented in Fig. 37.1. Many actual scenarios can be expressed as similar model, such as a motorcade drives through urban canyons, overpasses and tunnels. It can be seen that, the three users are all unable to complete standalone positioning for the number of visible satellites is less than four. However, if observation data can be exchanged between users, GNSS collaborative positioning methods can be used to improve the positioning availability and accuracy.

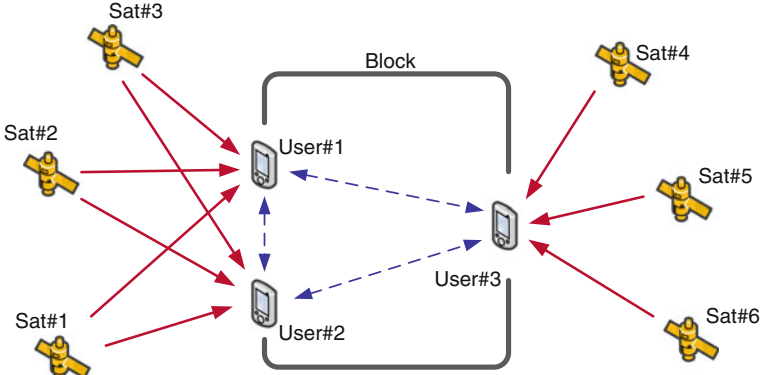


Fig. 37.1 Example of GNSS multi-user collaboration

Consider the user group with N users and M total visible GNSS satellites. The position of satellites and users are indicated by \mathbf{p}_m^s and \mathbf{p}_n^u respectively. For user# n , the clock bias is denoted as $b_n^u = c\delta_n^u$ where c the speed of light, denote by \mathcal{M}_n the set of visible GNSS satellites and by $\mathbf{X}_n^s \triangleq \{\mathbf{p}_m^s\}$ all the visible satellites positions and by \mathcal{K}_n the set of users that can collaborate with and by $\mathbf{X}_n^u \triangleq \{\mathbf{p}_k^u\}$ the users positions for all $k \in \mathcal{K}_n$.

Observations data contain pseudo-range and relative distance measurements:

$$\begin{aligned} \rho_{m,n} &= \|\mathbf{p}_m^s - \mathbf{p}_n^u\| + b_n^u + w_{m,n}^s, \quad m \in \mathcal{M}_n \\ r_{k,n} &= \|\mathbf{p}_k^u - \mathbf{p}_n^u\| + w_{k,n}^u, \quad k \in \mathcal{K}_n \end{aligned} \tag{37.1}$$

where $w_{m,n}^s$ and $w_{k,n}^u$ are the measurement noise. Denote by $\mathbf{z}_n \triangleq \left\{ \left\{ \rho_{m,n}^s \right\}, \left\{ r_{k,n}^u \right\} \right\}$ the observation measurements and by $\mathbf{w}_n \triangleq \left\{ \left\{ w_{m,n}^s \right\}, \left\{ w_{k,n}^u \right\} \right\}$ the observation noise vector. It is assumed that the covariance matrix of \mathbf{w}_n is \mathbf{R}_n . All the observation equation can be written as a compact form with a observation function $h(\cdot)$:

$$\mathbf{z}_n = h(\mathbf{x}_n, \mathbf{X}_n^s, \mathbf{X}_n^u, \mathbf{w}_n) \tag{37.2}$$

The state variable of user# n is defined as $\mathbf{x}_n^{(t)} = f(\mathbf{x}_n^{(t-1)}, \mathbf{v}_n^{(t)})$ where $f(\cdot)$ is the state transition function, and $\mathbf{v}_n^{(t)}$ is the process noise which is assumed as distributed with zero mean and covariance matrix $\mathbf{Q}_n^{(t)}$. In the case of the users with very low mobility, such as pedestrian, the PT model is given by its position and the clock bias $\mathbf{x}_n \triangleq [\mathbf{p}_n^u \quad b_n^u]^T$. Then, it can be modelled that

$$\mathbf{x}_n^{(t)} = f(\mathbf{x}_n^{(t-1)}, \mathbf{v}_n^{(t)}) = \mathbf{F}_n^{(t)} \mathbf{x}_n^{(t-1)} + \mathbf{V}_n^{(t)} \mathbf{v}_n^{(t)} = \mathbf{I}_4 \mathbf{x}_n^{(t-1)} + \Delta t \mathbf{I}_4 \mathbf{v}_n^{(t)} \tag{37.3}$$

where \mathbf{I}_4 is the identity matrix of size 4×4 , Δt is the time elapsed. In the case of the users with very low accelerations, such as motorcycle and car, the PVT model is given by its position, clock bias, velocity, and clock drift $\mathbf{x}_n \triangleq [\mathbf{p}_n^u \ b_n^u \ \dot{\mathbf{p}}_n^u \ \dot{b}_n^u]^T$. Then, it can be derived that

$$\mathbf{x}_n^{(t)} = f(\mathbf{x}_n^{(t-1)}, \mathbf{v}_n^{(t)}) = \mathbf{F}_n^{(t)} \mathbf{x}_n^{(t-1)} + \mathbf{V}_n^{(t)} \mathbf{v}_n^{(t)} = \begin{bmatrix} \mathbf{I}_4 & \Delta t \mathbf{I}_4 \\ \mathbf{0} & \mathbf{I}_4 \end{bmatrix} \mathbf{x}_n^{(t-1)} + \begin{bmatrix} \frac{\Delta t^2}{2} \mathbf{I}_4 \\ \Delta t \mathbf{I}_4 \end{bmatrix} \mathbf{v}_n^{(t)} \quad (37.4)$$

For the users with high dynamic, the state model can be further extended by adding high-order parameters.

37.3 Distributed GNSS Collaborative Positioning Methods

In this Section, three distributed algorithms are presented, which are DLS, DEKF and DUKF. Since the distributed algorithms have stronger flexibility for that different users can choose different algorithms according to their computing capacity. Besides, The PT model is used while the algorithms based on PVT model can be obtained in the same way.

37.3.1 Distributed Least Squares

The DLS algorithm is divided into four steps:

1. If there is no priori information, $\mathbf{x}_n = [0, 0, 0, 0]^T$ is used as the initial position and clock bias. Otherwise, the priori information of user position and clock bias would be used as the initial values.
2. Ignore the effects of measurement noises, the receive data Eq. (37.1) are linearized by using the first order Taylor expansion. Thus, the following equations can be obtained after linearization: $\mathbf{G}_n \Delta \mathbf{x}_n \approx \delta_n$ where

$$\mathbf{G}_n = \begin{bmatrix} \mathbf{G}_n^s \\ \mathbf{G}_n^u \end{bmatrix}, \quad \mathbf{G}_n^s = \begin{bmatrix} \frac{\partial \rho_{m,n}}{\partial \mathbf{x}_n} & 1 \\ \vdots & \vdots \end{bmatrix}, \quad \delta_n^s = \begin{bmatrix} \Delta \rho_{m,n} \\ \vdots \end{bmatrix}, \quad m \in \mathcal{M}_n \quad (37.5)$$

$$\delta_n^u = \begin{bmatrix} \delta_n^s \\ \delta_n^u \end{bmatrix}, \quad \mathbf{G}_n^u = \begin{bmatrix} \frac{\partial r_{k,n}}{\partial \mathbf{x}_n} & 0 \\ \vdots & \vdots \end{bmatrix}, \quad \delta_n^u = \begin{bmatrix} \Delta r_{k,n} \\ \vdots \end{bmatrix}, \quad k \in \mathcal{K}_n$$

3. The weighted least squares algorithm is used to solve the linear equations. In normal conditions, the weight values are decided by the standard deviation of measurement errors. Thus, the algorithm can be written as

$$\Delta \mathbf{x}_n = (\mathbf{G}_n \mathbf{R}_n \mathbf{G}_n)^{-1} \mathbf{G}_n^T \mathbf{R}_n \delta_n \quad (37.6)$$

4. The unknowns $\Delta \mathbf{x}_n$ is used to update the results. Besides, $\|\Delta \mathbf{x}_n\|$ is calculated and compared to a pre-set threshold to determine whether the convergence of Newton iteration algorithm. If $\|\Delta \mathbf{x}_n\|$ is less than the threshold, which means the iterative process has converged and the results can be returned. If not, the process should return to the second step and repeat the iterative calculation.

37.3.2 Distributed Extended Kalman Filter

Compare with DLS, the DEKF linearizes both observation and state models and conducts a global optimal recursive state estimation. The DEKF based on first order Taylor expansion is used here, and the DEKF based on higher order expansion can also be used for higher order accuracy. The filter contains two steps.

1. Predict. The state function $\mathbf{x}_n^{(t|t-1)} = f(\mathbf{x}_n^{(t-1)}, \mathbf{v}_n^{(t)})$ is used to predict the state $\mathbf{x}_n^{(t|t-1)}$ and covariance $\mathbf{P}_n^{(t|t-1)} = (\mathbf{F}_n^{(t)}) \mathbf{P}_n^{(t-1)} (\mathbf{F}_n^{(t)})^T + (\mathbf{V}_n^{(t)}) \mathbf{Q}_n^{(t)} (\mathbf{V}_n^{(t)})^T$, where $\mathbf{F}_n^{(t)} = \left. \frac{\partial f}{\partial \mathbf{x}_n} \right|_{\mathbf{x}_n = \mathbf{x}_n^{(t|t-1)}}$ and $\mathbf{V}_n^{(t)} = \left. \frac{\partial f}{\partial \mathbf{v}_n} \right|_{\mathbf{x}_n = \mathbf{x}_n^{(t|t-1)}}$ are the Jacobin matrices of state and process noise vectors respectively.
2. Update. The observation function $h(\cdot)$ is used to calculate the measurements residual $\Delta \mathbf{z}_n^{(t)} = \mathbf{z}_n^{(t)} - h(\mathbf{x}_n^{(t|t-1)}, (\mathbf{X}_n^s)^{(t)}, (\mathbf{X}_n^u)^{(t)})$ and their covariance matrix:

$$\mathbf{S}_n^{(t)} = (\mathbf{H}_n^{(t)}) \mathbf{P}_n^{(t|t-1)} (\mathbf{H}_n^{(t)})^T + (\mathbf{W}_n^{(t)}) \mathbf{R}_n^{(t)} (\mathbf{W}_n^{(t)})^T \quad (37.7)$$

where $\mathbf{H}_n^{(t)} = \left. \frac{\partial h}{\partial \mathbf{x}_n} \right|_{\mathbf{x}_n = \mathbf{x}_n^{(t|t-1)}}$ and $\mathbf{W}_n^{(t)} = \left. \frac{\partial h}{\partial \mathbf{w}_n} \right|_{\mathbf{x}_n = \mathbf{x}_n^{(t|t-1)}}$ are the Jacobin matrices of state and measurement noise vectors respectively. Thus, the Kalman gain are computed by $\mathbf{K}_n^{(t)} = \mathbf{P}_n^{(t|t-1)} (\mathbf{H}_n^{(t)})^T (\mathbf{S}_n^{(t)})^{-1}$, and the state vector and covariance matrix can be corrected by $\mathbf{x}_n^{(t)} = \mathbf{x}_n^{(t|t-1)} + \mathbf{K}_n^{(t)} \Delta \mathbf{z}_n^{(t)}$ and $\mathbf{P}_n^{(t)} = \mathbf{P}_n^{(t|t-1)} - (\mathbf{K}_n^{(t)}) \mathbf{S}_n^{(t)} (\mathbf{K}_n^{(t)})^T$.

37.3.3 Distributed Unscented Kalman Filter

The unscented Kalman filter uses a deterministic sampling technique known as the unscented transform to pick a minimal set of sample points (called sigma points) around the mean [8]. The filter also contains the two steps.

1. Predict. The estimated state $\mathbf{x}_{n,a}^{(t-1)}$ and covariance matrix $\mathbf{P}_{n,a}^{(t-1)}$ are augmented with the two types of noise. A set of sigma points is derived by

$$\begin{cases} \boldsymbol{\chi}_{n,a,0}^{(t-1)} = \mathbf{x}_{n,a}^{(t-1)} \\ \boldsymbol{\chi}_{n,a,i}^{(t-1)} = \mathbf{x}_{n,a}^{(t-1)} + \left(\sqrt{(L + \lambda)\mathbf{P}_{n,a}^{(t-1)}} \right)_i, i = 1, 2, \dots, L \\ \boldsymbol{\chi}_{n,a,i+L}^{(t-1)} = \mathbf{x}_{n,a}^{(t-1)} - \left(\sqrt{(L + \lambda)\mathbf{P}_{n,a}^{(t-1)}} \right)_i, i = 1, 2, \dots, L \end{cases} \quad (37.8)$$

where L is the length of the state vector, and $(\sqrt{(\cdot)})_i$ is the i th column of the Cholesky decomposition. The sigma points are propagated through the transition function $\boldsymbol{\chi}_{n,a,i}^{(t|t-1)} = f(\boldsymbol{\chi}_{n,a,i}^{(t-1)})$. The weighted sigma points are recombined to produce the predicted state $\mathbf{x}_{n,a}^{(t|t-1)} = \sum_{i=0}^{2L} \tilde{\varepsilon}_n^i \boldsymbol{\chi}_{n,a,i}^{(t|t-1)}$ and the covariance can be derived:

$$\mathbf{P}_{n,a}^{(t|t-1)} = \sum_{i=0}^{2L} \tilde{\varepsilon}_n^i \left[\boldsymbol{\chi}_{n,a,i}^{(t|t-1)} - \mathbf{x}_{n,a}^{(t|t-1)} \right] \left[\boldsymbol{\chi}_{n,a,i}^{(t|t-1)} - \mathbf{x}_{n,a}^{(t|t-1)} \right]^T \quad (37.9)$$

The weights for the state and covariance are given by:

$$\begin{aligned} \tilde{\varepsilon}_n^0 &= \frac{\lambda}{L + \lambda}, & \tilde{\varepsilon}_n^0 &= \frac{\lambda}{L + \lambda} + (1 - \alpha^2 + \beta) \\ \tilde{\varepsilon}_n^i &= \tilde{\varepsilon}_n^i = \frac{\lambda}{2(L + \lambda)}, & i &= 1, 2, \dots, 2L \end{aligned} \quad (37.10)$$

where $\lambda = \alpha^2(L + \kappa) - L$ and α, β, κ are tunable.

2. Update. The sigma points are projected through the observation function $\boldsymbol{\gamma}_{n,a,i}^{(t)} = h(\boldsymbol{\chi}_{n,a,i}^{(t|t-1)})$. The weighted sigma points are recombined to produce the predicted measurement and predicted measurement covariance.

$$\boldsymbol{\gamma}_n^{(t)} = \sum_{i=0}^{2L} \tilde{\varepsilon}_n^i \boldsymbol{\gamma}_{n,a,i}^{(t)}, \quad \mathbf{S}_n^{(t)} = \sum_{i=0}^{2L} \tilde{\varepsilon}_n^i \left[\boldsymbol{\gamma}_{n,a,i}^{(t)} - \boldsymbol{\gamma}_n^{(t)} \right] \left[\boldsymbol{\gamma}_{n,a,i}^{(t)} - \boldsymbol{\gamma}_n^{(t)} \right]^T \quad (37.11)$$

The state-measurement cross-covariance matrix and Kalman gain are computed

$$\mathbf{P}_{\mathbf{z}|\mathbf{x}}^{(t)} = \sum_{i=0}^{2L} \tilde{\varepsilon}_n^i \left[\boldsymbol{\chi}_{n,a,i}^{(t|t-1)} - \mathbf{x}_{n,a}^{(t|t-1)} \right] \left[\boldsymbol{\gamma}_{n,a,i}^{(t)} - \boldsymbol{\gamma}_n^{(t)} \right]^T, \quad \mathbf{K}_n^{(t)} = \mathbf{P}_{\mathbf{z}|\mathbf{x}}^{(t)} \left(\mathbf{S}_n^{(t)} \right)^{-1} \quad (37.12)$$

Finally, the state vector and covariance matrix can be corrected by

$$\mathbf{x}_n^{(t)} = \mathbf{x}_{n,a}^{(t|t-1)} + \mathbf{K}_n^{(t)} \left(\mathbf{z}_n^{(t)} - \gamma_n^{(t)} \right), \quad \mathbf{P}_n^{(t)} = \mathbf{P}_{n,a}^{(t|t-1)} - \left(\mathbf{K}_n^{(t)} \right) \mathbf{S}_n^{(t)} \left(\mathbf{K}_n^{(t)} \right)^T \quad (37.13)$$

For the three algorithms, the uncertainty on users' position estimates are hard to combine with the uncertainty on relative distances. A worst approximation solution is to adapt the new variance as the relative distance measurements variance plus the neighbouring users' covariance matrix trace [7]: $\tilde{\sigma}_{k,n}^u = \sigma_{k,n}^u + \text{tr}(\mathbf{P}_k^{(t)})$.

37.4 Experiment Results

37.4.1 Outdoor Scenario

Table 37.1 gives the users position in east-north-under (ENU) coordinates, and Fig. 37.2 shows the simulation scenario. The pseudo-range and relative distance measurement noises are corrupted by additive white Gaussian noise with standard deviation 5 and 1 m respectively, and the initial positions of each users are set 1 km

Table 37.1 The users' ENU coordinates

User number	1	2	3	4
East (m)	0	0	20	0
North (m)	0	20	0	0
Up (m)	0	0	0	20

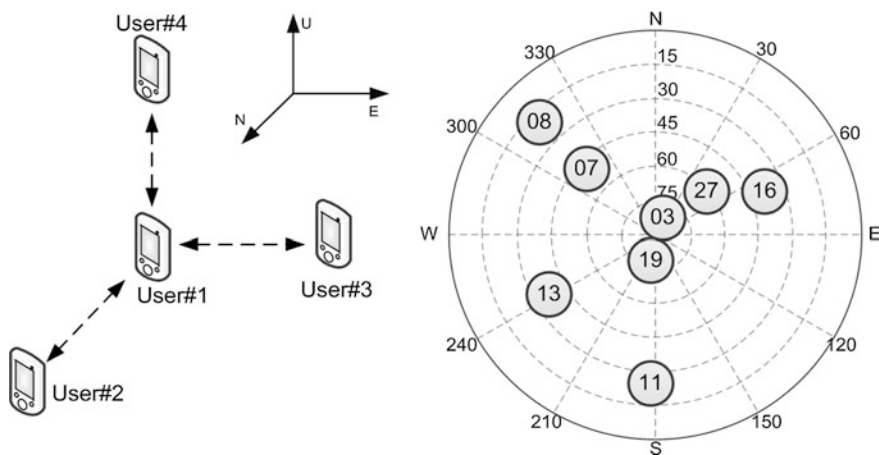


Fig. 37.2 User distribution and skyplot of outdoor experiment scenario

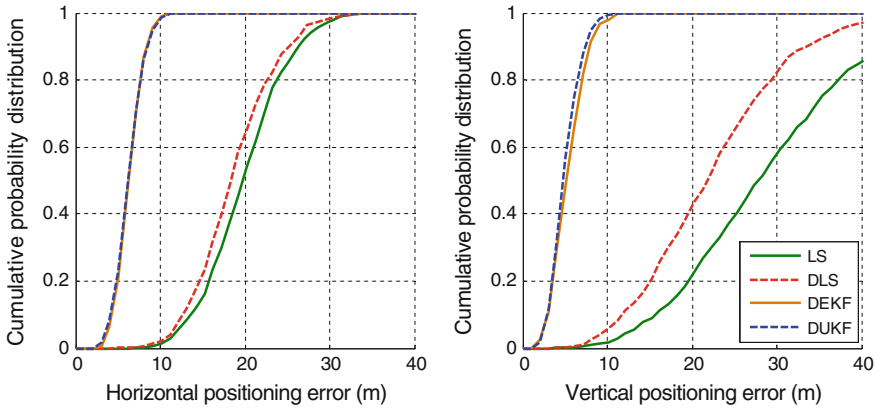


Fig. 37.3 The cumulative probability distribution of *horizontal* and *vertical* positioning errors (over all users) in outdoor scenario

away of the actual value. The Monte Carlo method over 1000 tests is implemented in the following simulations.

Figure 37.3 shows horizontal and vertical positioning performance of three algorithms and GNSS standalone least squares algorithm in terms of error cumulative distribution function. As it can be seen that, the collaborative algorithms have better performance. Besides, the performance improvements of DEKF and DUKF are almost the same, and compared with DLS, the improvements are significant.

For GNSS standalone positioning, if processing capability is limited there is restriction on the number of measurements which can be obtained and processed [9]. It can be seen that the computational complexity of collaborative algorithms increase with the increase of the amount of users. Therefore, the study of how to select a small number of users to collaborate in order to weigh the amount of calculation and performance has a great significance for the actual applications. Figure 37.4 shows the performance of different user collaboration with DLS and DUKF. The performance of DEKF is similar to DUKF, for the sake of more concise figure, the results of DEKF are not given in Fig. 37.4. It can be clearly seen that the collaboration between user#1 and user#4 obtain the optimal performance for both algorithms. The main difference is that the collaborative link between user#1 and user#4 is vertical, which means the vertical collaborative links can obtain better performance for users in outdoor scenarios.

37.4.2 Partially Blocked Scenario

The following analysis focus on the partially occlusion environments where each user can only see part of the satellites. The coordinates of four users remain the same as Table 37.1. Figure 37.5 shows the partially blocked scenario. Compared with Fig. 37.2, the number of visible satellites for each user is reduced three.

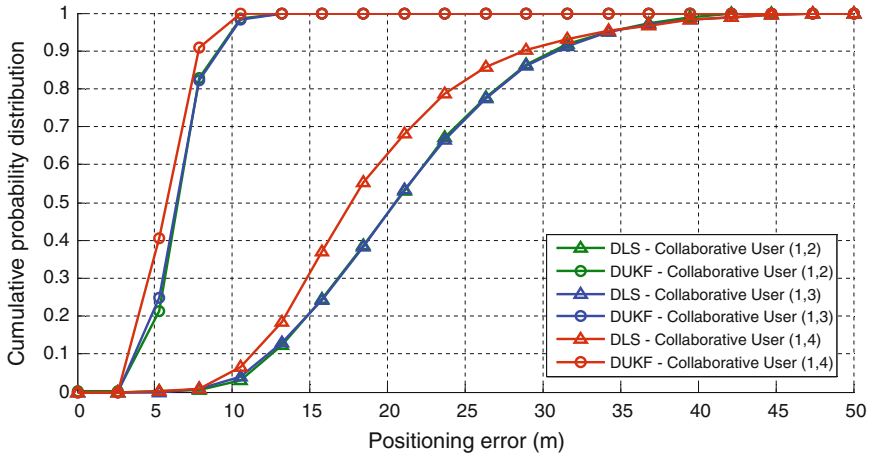


Fig. 37.4 The cumulative probability distribution of positioning errors (over tow users) in outdoor scenario

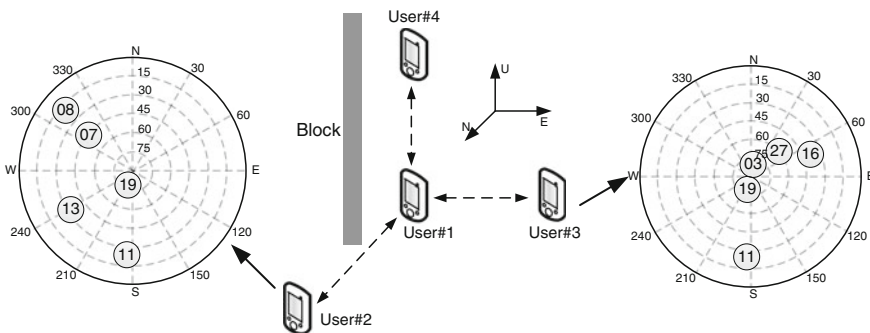


Fig. 37.5 Simulation scenario in which users are blocked

Under this blocked scenario, the distributed collaborative positioning simulation are carried out and the results of positioning errors (over all users) are shown in Fig. 37.6. It can be clearly seen that, because the some satellite signals are blocked, the GNSS standalone positioning accuracy reaches hundred meters and can not meet the command of most applications. However, by collaboration, the positioning performance has been significantly improved, especially for DEKF and DUKF. As the comparison with the results in Fig. 37.3, the advantages of collaborative algorithms are more obvious in such challenging signal environments.

Similar to the simulation in Figs. 37.4 and 37.7 shows the performance of different user collaboration with DLS and DUKF under this blocked scenario. It can be seen that, the collaboration between user#1 and user#2 gains the biggest performance improvement which is different from the results in Fig. 37.4. Moreover,

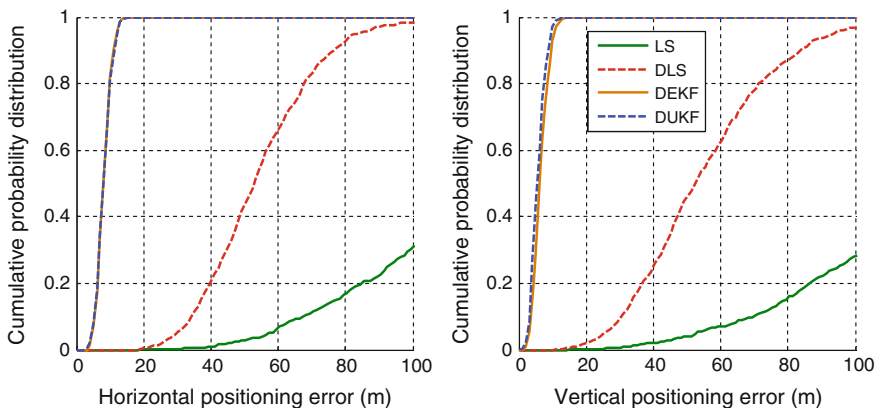


Fig. 37.6 The cumulative probability distribution of *horizontal* and *vertical* positioning errors (over all users) in blocked scenario

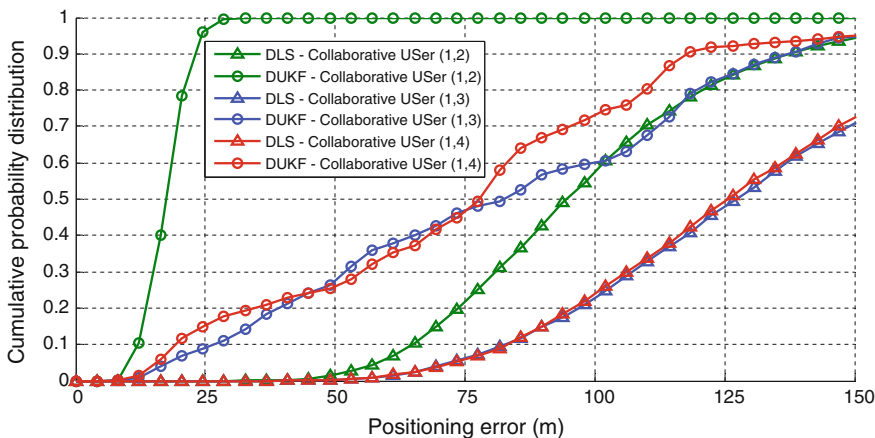


Fig. 37.7 The cumulative probability distribution of positioning errors (over tow users) in blocked scenario

the results vary wildly between the collaboration with user#2 and user#3. The main reason of the above different conclusions is that the distribution of visible satellites for user#2 is better than the others. Therefore, for the case of limited computational complexity in blocked scenarios, the users which can get better standalone positioning accuracy should be most preferred to collaborate with.

37.5 Conclusions

In this paper, three distributed GNSS collaborative positioning algorithms are introduced in detail, and the performance of these algorithms are analysed in outdoor and partially blocked scenarios respectively. The experiment results show that the positioning accuracy of all three algorithms is higher than that of GNSS standalone positioning. Especially in the partially blocked scenarios, the performance improvements by collaboration are significant. For the three distributed algorithms, the DEKF algorithm is pretty good in terms of performance and computational complexity. The distributed algorithms have stronger flexibility for that different users can choose different algorithms according to their computing capacity. For the case of limited computational complexity, the issues that how to select collaborative users are also discussed in outdoor and partially blocked scenarios respectively. The results can provide some useful information for the design and testing of the actual collaboration applications.

Furthermore, more collaborative experiments will be carried out in the next step, such as real-time dynamic collaborative positioning of multiple vehicle navigators in urban and indoor environments.

Acknowledgments This paper is supported by the National Key Scientific Instrument and Equipment Development Project (2012YQ140026).

References

1. Garello R, Samson J, Spirito M et al (2012) Peer-to-peer cooperative positioning. *Inside GNSS*
2. Win MZ, Conti A, Mazuelas S et al (2011) Network localization and navigation via cooperation. *IEEE Commun Mag* 49:56–62
3. Roumeliotis SI, Bekey GA (2002) Distributed multirobot localization. *IEEE Trans Robot Autom* 18:781–795
4. Heinrichs G, Mulassano P, Dovis F (2004) A hybrid positioning algorithm for cellular radio networks by using a common rake receiver architecture. In: *IEEE international symposium on PIMRC 2004*, pp 2347–2351
5. Huang B, Yao Z, Cui XW et al (2014) GNSS collaborative positioning and performance analysis. *ION GNSS + 2014*, pp 1920–1930
6. Caceres MA, Penna F, Wymeersch H et al (2011) Hybrid cooperative positioning based on distributed belief propagation. *IEEE J Sel Areas Commun* 29:1948–1958
7. Sottile F, Wymeersch H, Caceres MA et al (2011) Hybrid GNSS-terrestrial cooperative positioning based on particle filter. *IEEE Globecom 2011*:1–5
8. Dan S (ed) (2006) *Optimal state estimation Kalman H and nonlinear approaches*. Wiley Interscience Publication, New York
9. Phatak MS (2001) Recursive method for optimum GPS satellite selection. *IEEE Trans Aerosp Electron Syst* 37:751–754

Chapter 38

Indoor Location Fingerprinting System Using DTMB Signal

Qiqi Wang, Shuai Li, Yingxin Zhao, Kun Chen, Bing Liu,
Xiyang Xu and Hong Wu

Abstract GNSS signal can be widely used in outdoor environment. However, due to the barrier of walls, GNSS signal becomes too weak to be used in indoor environment. This paper analyzed the Wi-Fi based indoor positioning method, which is mostly widely researched, introduced a DTMB based indoor positioning method and confirmed its feasibility. By using the KNN algorithm and the DTMB signal's location fingerprint, the indoor positioning system proposed by this paper can locate any unknown nodes inside the room. And also some tests were designed to measure the positioning accuracy. Test results showed that the accuracy of the DTMB based indoor positioning system is similar to the Wi-Fi based positioning system. The new DTMB based indoor positioning method can be a useful complement to the existing indoor positioning methods.

Keywords Indoor positioning · DTMB · Signal fingerprint · KNN

38.1 Introduction

GPS can effectively deal with outdoor positioning in the most cases. However, GPS signal becomes so weak in the indoor environment that the precise positioning cannot be got simply by using GPS. With the increasing demand of indoor posi-

This work was supported by Ph.D. Programs Foundation of Ministry of Education of China (20110031110028), Tianjin Research Program of Application Foundation and Advanced Technology (13JCZDJC26000, 13JCQNJC01000).

Q. Wang · S. Li · Y. Zhao · K. Chen · B. Liu · X. Xu · H. Wu (✉)
College of Electronic and Optical Engineering, Nankai University, Tianjin, China
e-mail: wuhong@nankai.edu.cn

Q. Wang · S. Li · Y. Zhao · K. Chen · B. Liu · X. Xu · H. Wu
Tianjin Key Laboratory of Photonics Materials and Technology for Information Science,
Nankai University, Tianjin, China

tioning, the research on new positioning mode has become more and more important. Currently, there have been many achievements in terms of indoor positioning methods. However due to the limitation of the application conditions, these indoor positioning systems are still not widely used. For example, Wi-Fi based positioning system covers only small areas; Hardware requirement for the RFID based positioning system is very high; Cellular network based positioning system covers large area, while its precision is low.

Digital Television Terrestrial Multimedia Broadcasting (DTMB) can cover larger area and can ensure the indoor signal strength. At present, there are few articles researching on the DTMB based indoor positioning. Most of DTMB based researches are still remain in the outdoor positioning field. Article [1] built the hybrid positioning system, using GPS signal and DTMB signal at the same time. The result shows that DTMB signal can improve the accuracy of the existing positioning system. This paper discusses the feasibility of DTMB based indoor positioning.

Currently, Wi-Fi signal is generally used in the indoor positioning system [2]. The reasons are as follows:

- (a) Existing Wi-Fi network has higher coverage rate, which has been installed in some environments, such as office and family, so there is no need to re-install in the process of locating.
- (b) Most mobile devices, such as mobile phone, have been equipped with the Wi-Fi signal, so there is no need to design additional receiver.

From above two points, Wi-Fi based positioning system can decrease the cost of hardware and offer higher positioning accuracy, but there are also some following disadvantages:

- (a) The coverage is small. Although office and family are covered by the Wi-Fi network, there are still many sparsely populated and underdeveloped areas that cannot be covered.
- (b) There is interference from other signals. The frequency range of Wi-Fi signal is 2.4 GHz, which is also shared by Bluetooth and ZigBee. Therefore, it is susceptible.
- (c) The positioning accuracy is largely influenced by the obstacles. When the obstacle size is equal to the wavelength, its effect on the propagation of electromagnetic wave cannot be ignored. Due to the high frequency and short wavelength, Wi-Fi signal is more easily influenced by the surrounding obstacles, thus the positioning accuracy is affected.

In contrast, the DTMB signal has the following advantages [3]:

- (a) The transmitting power of DTMB signal is high, and the coverage can be up to hundreds of kilometers.
- (b) DTMB signal has special frequency band, so the interference from other emission sources is small.

- (c) The frequency of DTMB signal is low and the penetrating ability is strong. The wavelength of DTMB signal is longer than that of Wi-Fi signal, so the positioning accuracy can be less affected by indoor obstacles.
- (d) The weather conditions have little influence on DTMB signal. Compared with Wi-Fi signal affected by the vapor concentration, the effect of weather conditions on DTMB signal can be neglected due to its lower frequency.

38.2 DTMB Based Fingerprint Positioning Method

Compared with outdoor positioning, the environment has more effect on the indoor positioning accuracy. The indoor positioning environment is more complex than the outdoor, so the tiny change of obstacles will lead to the change of signal. Furthermore, it will make a large different between indoor and outdoor environments, when the positioning accuracy is under of 5 m. The requirement of the indoor positioning accuracy is higher.

In addition, many algorithms commonly used in the outdoor positioning are not suitable for the indoor positioning. The outdoor positioning uses GPS signal. Mobile terminal utilizes time of arrival (TOA) algorithm to get the position information of the unknown target. TOA positioning algorithm [4] strictly requires time synchronization between the mobile terminal and the transmitting node. That is to say, both the mobile terminal and the transmitting node are required to install specialized hardware. For example, the atomic clock is installed on the GPS satellite to keep timing accurately. There is no doubt that the cost will be increased. Due to the influence of the obstacle, multipath effect and non-line of sight error, indoor positioning are more obvious than outdoor localization. This restricts the application of TOA positioning algorithm.

The outdoor positioning also uses the angle of arrival (AOA) between the mobile terminal and the transmitting node to locate. AOA algorithm can calculate the location of the mobile terminal accurately by only two transmitting node, but this algorithm requires that both the transmitting node and the mobile terminal are installed with directional antenna or antenna array, which not only increases the system cost, but also makes the system size bigger. Likewise, due to multipath effect, non-line of sight error, and reflection of the wall or other obstacles, the arrival direction of the signal will change. All these factors will seriously affect the indoor positioning accuracy of AOA algorithm.

Currently, RSS fingerprint technology is generally used in the indoor positioning system. This technology can be applied to the complex indoor environment, without considering environments and propagation paths [5] and obtaining the position of the transmitting node in advance. Therefore, it can well deal with reflection, refraction, non-line of sight error and multipath effect. In order to ensure the positioning accuracy, reference points should have different signal features even if they are very close. In the RSS fingerprint technology, a variety of signal's

features can be used as the fingerprint data, in which the most widely feature is received signal strength indicator (RSSI). RSSI indicates the amplitude information of the received signal, related to propagation distance in theory [6], and it can be expressed as:

$$RSSI = P_{send} + P_{amplify} - PL(d) \quad (38.1)$$

where RSSI is the received signal power, P_{send} is transmitting power, $P_{amplify}$ is antenna gain, $PL(d)$ is path loss. $PL(d)$ satisfies the log normal distribution model:

$$PL(d) = P(d_0) - 10n \lg \frac{d}{d_0} - \delta \quad (38.2)$$

where $PL(d_0)$ is path loss in the signal transmission distance d_0 , δ is random noise coefficient. But in practice, RSSI is associated with more factors, such as environments, transmitting node power and signal processing methods of the receiving node.

RSS fingerprint technology can estimate the location of the mobile terminal by off-line training stage and online positioning stage [7]. In the off-line training stage, RSS fingerprints at reference label points will be collected and these messages will be stored into a database. In the online positioning stage, the system receives RSS fingerprints collected from mobile terminal, and uses K Nearest Neighbor (KNN) algorithm to calculate the position of the terminal.

In the process of online positioning, KNN algorithm can compare testing points with the reference label points in the database to find the testing points' positions. KNN is a simple and effective matching algorithm [8], so it is widely used in the fingerprint based indoor position technology.

The specific steps of the KNN algorithm are as followed:

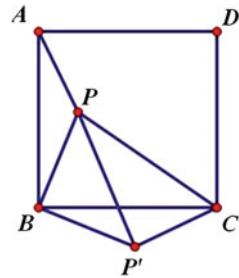
- (a) Calculate the distances (dist) between unknown object and all the reference points in the database and save them in the set A.
- (b) Find K minimum values in the set A, and save the reference points ID in the set B.
- (c) Output the average of the position coordinates of all elements in the set B, as the final position of the unknown object.

KNN algorithm is expressed as

$$(\hat{x}, \hat{y}) = \frac{1}{K} \sum_{r=1}^K (x_r, y_r) \quad (38.3)$$

In the formula, (x_r, y_r) is the location coordinates of the reference points in the set B. K is the number of the reference points, which is based on different positioning environments. Usually K is not more than the square root of the total number of the reference points.

Fig. 38.1 KNN-based positioning diagram



To determine the value of *dist* in KNN algorithm, the Euclidean Distance and the Manhattan Distance are highly used in the most researches. Euclidean distance can be expressed as follows:

$$d(a, b) = \sqrt{(x_{a1} - x_{b1})^2 + (x_{a2} - x_{b2})^2 + \dots + (x_{an} - x_{bn})^2} \tag{38.4}$$

where $x_a(x_{a1}, x_{a2}, \dots, x_{an}), x_b(x_{b1}, x_{b2}, \dots, x_{bn})$ are n-dimensional vectors. Manhattan distance can be expressed as follows:

$$d(a, b) = |x_{a1} - x_{b1}| + |x_{a2} - x_{b2}| + \dots + |x_{an} - x_{bn}| \tag{38.5}$$

A situation may be existed in KNN algorithm, which is shown in Fig. 38.1:

In the figure, point P' is the location of the unknown object. If $k = 4$, the reference points position are respectively the ABCD. Clearly in the picture, point P' and point C are the closest in the distance, while point P' and point A are the farthest. The distance between the reference points and the unknown object is not taken into consider when using KNN algorithm to get the averages of the position coordinates. In order to reduce this calculation error, the distance between the reference points and the unknown object is used as the weight in KNN algorithm, which is called weighted K nearest neighbor (WKNN) algorithm:

$$(\hat{x}, \hat{y}) = \frac{1}{K} \sum_{r=1}^K \omega_r(x_r, y_r); \quad \sum_{r=1}^K \omega_r = 1 \tag{38.6}$$

where ω_r is the weight of the r th reference point. It is inversely proportional to the distance between the reference point and the unknown object. The farther the distance is, the smaller the weight becomes. This algorithm distributes the bigger weight for the nearer distance, avoiding the error caused by the value of some too large reference point [9].

KNN algorithm only needs information of reference points which are saved in the database during offline training stage, without estimating parameters, so the matching speed can be fast. Meanwhile, KNN algorithm doesn't have to consider the signal propagation path or other factors, so it can provide higher positioning accuracy even in the environment with many obstacles. Above two points, KNN algorithm can be widely used in RSS fingerprint technology.

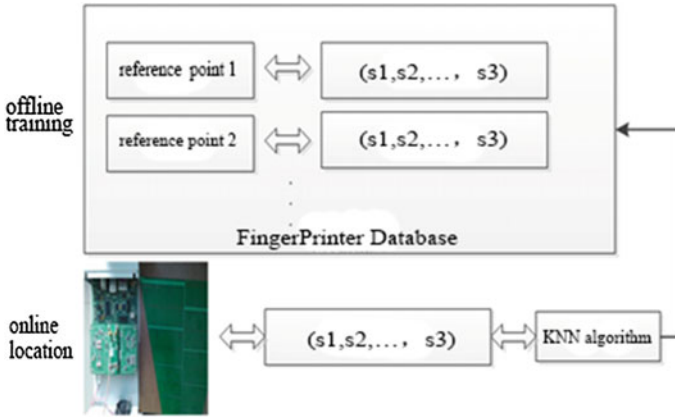


Fig. 38.2 Positioning process based on fingerprint method

38.3 System Workflow

Figure 38.2 shows the overall workflow of DTMB based fingerprint indoor positioning method.

The figure shows that in the off-line training stage, the signal strength at each reference label point is collected and these messages are stored into a database, which can store new data, update and delete data. In the online positioning stage, the mobile terminal measures and records the signal strength at the unknown node. Then calculate the distance between it and signal strength in the database. Finally use KNN algorithm to calculate the position of the terminal.

38.4 Design and Analysis of the Experiment

In this system, data was collected by USRP2. The left part of Fig. 38.3 is the hardware of USRP2, and the right one is the antenna which is used to receive DTMB signal.

Firstly, this paper designed the experiment to prove that the RSSI, which was applied as fingerprints, can be used in the process of DTMB based indoor positioning.

The size of the lab is 4 m × 4 m. Figure 38.4 shows the floor plan of the lab. Marks from 1 to 9 are reference label points, where each point’s signal strength is gotten. See Fig. 38.5.

As shown in the figure, each point’s signal strength changes with the label reference points. The line connecting the points of signal strength is not smooth. It is because of the multipath effect which is caused by the indoor obstacles.

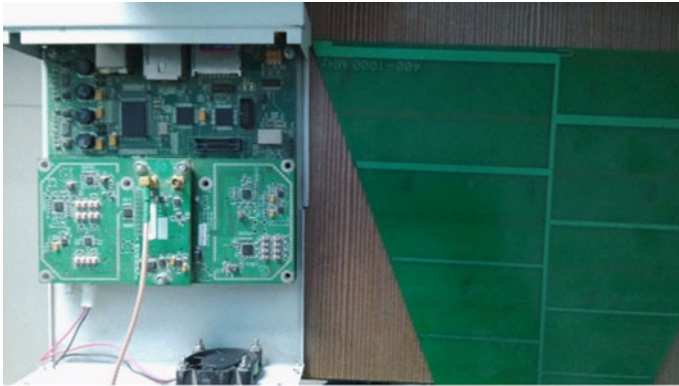


Fig. 38.3 Hardware for indoor positioning

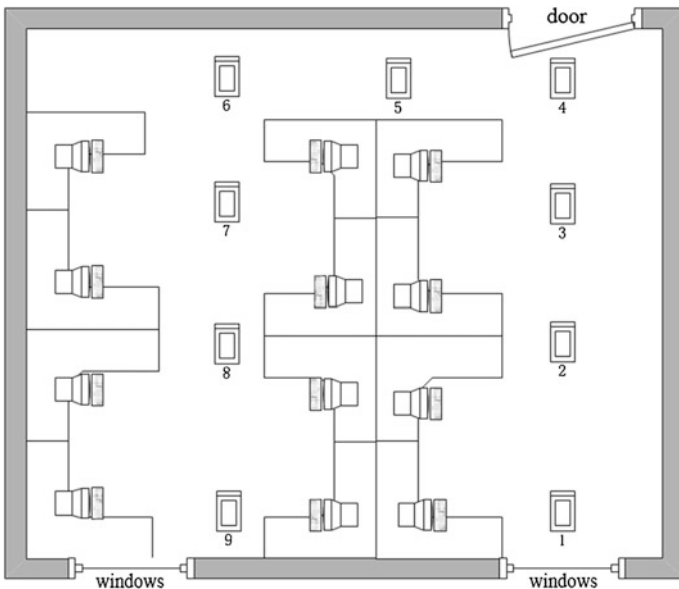


Fig. 38.4 Map of the laboratory

The result shows that RSSI, as a message of RSS fingerprint, can be applied in the process of DTMB based indoor positioning.

Figure 38.6 shows the signal strength of DTMB in different weather conditions (770 MHz).

The solid line indicates different DTMB signal strengths in sunny days, and the dashed line in cloudy. The gap between signal strengths in two weather conditions can be about 0.5 dB.

Fig. 38.5 RSSI of DTMB signals at reference point

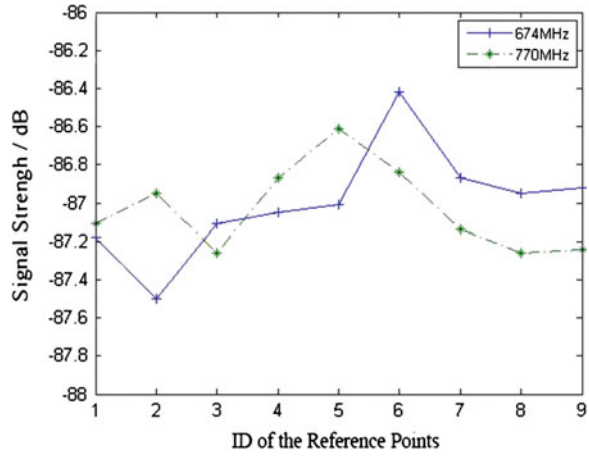


Fig. 38.6 RSSI of DTMB signal under different weather conditions

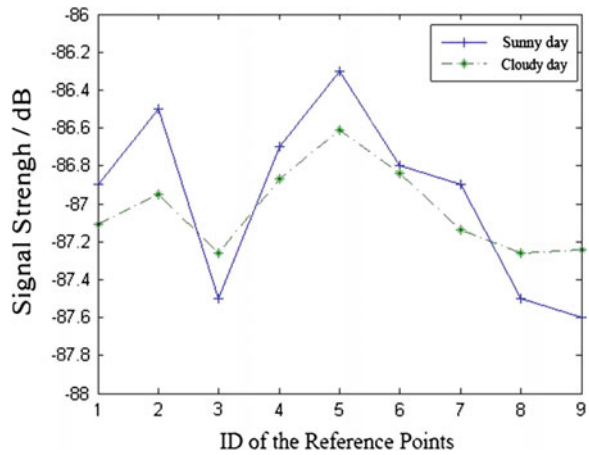


Table 38.1 Indoor positioning error based on RSSI

Cumulative probability	50 %	90 %
location error (m)	0.3	1

In subsequent experiments, 100 testing points were randomly selected in the space of the lab. Then by using the method mentioned in this paper, 100 testing points were located. Set $K = 2$, the number of the reference points, in KNN algorithm. Table 38.1 shows the location error. In comparison with the Literature [10], the indoor positioning accuracy of the method in this paper is as good as Wi-Fi based.

38.5 Conclusion

So far, the Wi-Fi based indoor positioning method still has some defects, such as small signal coverage, unspecialized frequency band and susceptible positioning accuracy. At the same time, the attenuation of Wi-Fi is weak in the short distance, which leads to tiny differences between Wi-Fi signal strengths at different preference points if in narrow indoor environment. If only the Wi-Fi signal is used to locate, that different Wi-Fi signals lead to the different positioning accuracy has to be considered. The positioning accuracy of the method mentioned in this paper is as good as that of Wi-Fi based. Furthermore, DTMB has larger coverage and specialized frequency band. The indoor obstacles also have less effect on the positioning accuracy. So this method can be a useful complement to the existing indoor GNSS.

References

1. Li W, Wu H, Zhang N, Jin Y (2010) Positioning system using digital television terrestrial multimedia broadcasting signals and GPS signals, *Video Engineering*, (6)
2. Lu H, Liu X, Zhang C et al (2010) Comparison of Wi-Fi positioning based on triangle and position fingerprint recognition algorithm. *Mobile Commun* 34(10):72–76
3. Fu X (2010) Research on the method of combined positioning based on DTMB-GNSS. Xi'an University of Science and Technology, Xi'an
4. So HC, Shiu EMK (2003) Performance of TOA-AOA hybrid mobile location. *IEICE Trans Fundam Electron Commun Comput Sci* 86(8):2136–2138
5. Zanca G, Zorzi F, Zanella A et al (2008) Experimental comparison of RSSI-based localization algorithms for indoor wireless sensor networks. In: *Proceedings of the workshop on real-world wireless sensor networks*. ACM, pp 1–5
6. Liu Y, Jin M, Cui C (2010) Improve research on node localization algorithm based on RSSI in wireless sensor networks. *Chinese J Sens Actuators* 23(5):717–721
7. Niu J, Liu Y, Lu B et al (2013) An in-building localization algorithm based on Wi-Fi signal fingerprint. *J Comput Res Dev* 50(3):568–577
8. Zhang Z-Y, Huang Y-L, Wang H-H (2008) A new KNN classification approach. *Comput Sci* 35(3):170–172
9. Liu W, Xu L, Li Z et al (2014) Location algorithm based on distance domain grid matching. *J Huazhong Univ Sci Technol (Nat Sci Ed)* 42(3):19–22
10. Popleteev A (2011) Indoor positioning using FM radio signals. University of Trento, Trento

Chapter 39

Simulation Analysis and Research on Key Technology for a Full-Time-Running Stellar Refraction Autonomous Navigation

Qin Lin, Zhi Li, Huaifeng Li, Liang Song and Bo Meng

Abstract Navigation based on stellar refraction refers to a method that uses large visual field star sensor to detect several stars refracted and not refracted by earth atmosphere, and then finishes autonomous navigation based on star map recognition, with merits such as simple system structure, low cost and high navigation accuracy. However, in order to achieve full-time-running high accuracy, three key technologies, namely strong background star map acquisition, refracted star extraction under strong background and atmospheric model should be solved for such method. By constructing observation policy model and detection limit model under the strong background, this paper realizes high SNR (Signal to Noise Ratio) star map acquisition under strong background through anti-blooming function of detector, and carries out theoretical analysis and outfield test; realizes recognition of dim refracted star and high-accuracy centroid positioning of refracted star under strong background via specific star pick-up algorithm, and performs outfield test on it; studies high-accuracy atmospheric model optimization method and proposes an applicable stellar refraction atmospheric model based on atmospheric refractivity change and identical with stellar refraction rule. Finally, on the basis of key technological solution study, it constructs stellar refraction navigation simulation system based on stellar refraction navigation realization approach and analyzes the influences of key error sources on navigation accuracy through simulation, so as to obtain the result that the accuracy of full-time-running navigation based on stellar refraction is superior to 1.2 km.

Keywords Stellar refraction · Autonomous navigation · Celestial navigation

Q. Lin (✉) · Z. Li · L. Song · B. Meng
Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology,
No. 104 Youyi Road, Haidian District, Beijing 100094, China
e-mail: linqin@bao.ac.cn

H. Li
Beijing Institute of Spacecraft System Engineering, Beijing 100094, China

39.1 Introduction

Navigation by stellar refraction is a type of autonomous celestial navigation method that is characterized by high navigation accuracy, high autonomy, simple equipment, etc. [1, 2], therefore many organizations have conducted researches on its realization approach and relevant technologies [3–5]. However, they all ignored the influence of space light environment on this navigation approach, i.e. the key technology to be solved for realizing high-accuracy navigation when the earth and the atmosphere in the field are lightened by the sun. Due to time-dependent nature of satellite observation position and significance of real-time navigation, it is of vital importance to make clear and solve the key technology of navigation by stellar refraction when the observed object is under strong background. The accomplishment of this work will greatly improve the practicability and ensure high efficiency of navigation by stellar refraction, thus realizing full-time-running navigation by stellar refraction.

Full-time-running navigation by stellar refraction proposed in this paper mainly aims at achieving high-accuracy navigation by stellar refraction under strong background. Therefore, Part II first summarizes the realization approach of autonomous navigation by stellar refraction and states the main error sources of influencing the accuracy of navigation by stellar refraction. On the basis of Part I, Part III analyzes the key technology of full-time-running navigation by stellar refraction and puts forward several key technologies requiring urgent solution mainly based on the main factors to be considered for high-accuracy navigation by stellar refraction and by combining the harsh observation environment confronted by full-time-running navigation by stellar refraction; then discusses and analyzes the solution and feasibility of such key technologies. Based on various models and algorithms realized in key technologies, Part IV constructs application simulation system of navigation by stellar refraction, to perform simulation analysis on the accuracy of full-time-running navigation by stellar refraction as well as the mechanism and degree of influences on navigation accuracy imposed by all error sources. Finally, it summarizes the current work and discusses future prospect of next work.

39.2 Autonomous Navigation Approach by Stellar Refraction

Navigation by stellar refraction is an autonomous navigation approach that uses star sensor to detect several stars refracted and not refracted by earth atmosphere, then calculates refraction angles based on star map recognition and finishes satellite positioning via the triangles of three or more refraction stars. Such navigation approach mainly involves four steps (Fig. 39.1):

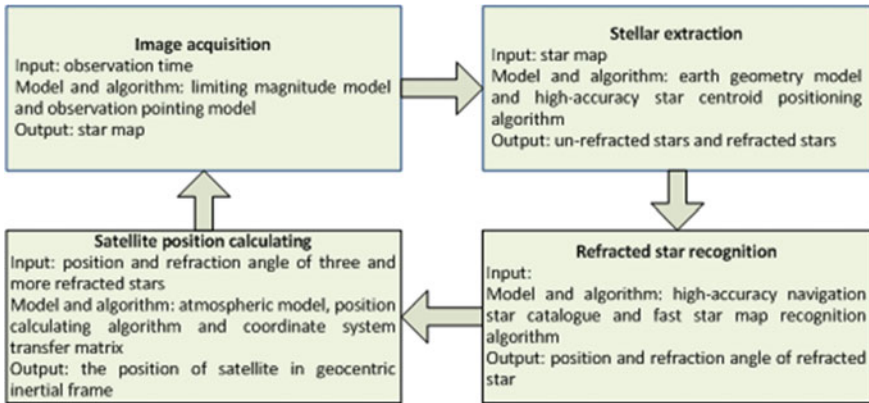


Fig. 39.1 Workflow of autonomous navigation by stellar refraction

(1) Image acquisition

Large visual field stellar refraction sensor carries out directional observation to earth or deflection observation to earth according to observation pointing model and detection limiting magnitude model based on refracted star, to obtain the star map containing the stars refracted and not refracted by atmosphere;

(2) Stellar extraction

First, perform denoising and distortion correction on the images acquired, and then extract outline information of the earth via edge detection algorithm and the earth geometry model, so as to analyze the effective area of causing stellar refraction and classify the stars in star map into refracted star and un-refracted star; finish high-accuracy position calculating of star through star centroid positioning algorithm and perform magnitude calculation, so as to get the luminance of all observed stars and their position relative to image sensor;

(3) Refracted star recognition

Finish fast star map matching by the extracted un-refracted star based on high-accuracy navigation star catalogue and star map recognition algorithm, to obtain the right ascension and declination of each star on star map; then search the corresponding star of each refracted star in navigation star catalogue, and calculate the refraction angle of refracted star according to right ascension and declination of refracted star and its corresponding star;

(4) Satellite position calculating

First, calculate the tangent height of each refracted star according to atmospheric model, then calculate the accurate position of geocenter in satellite coordinate system via position calculating algorithm, and finally calculate the position of satellite in geocentric inertial frame through coordinate transfer matrix.

In position calculating algorithm, the input items of single position calculating combination are: right ascension and declination of three refracted stars and tangent height of each refracted star. In case of several refracted stars in the field, high-accuracy navigation positioning result can be obtained based on data processing algorithm of refracted star of the optimal combination. As a result, quantity and distribution uniformity of refracted stars, position accuracy of refracted star and tangent height error of refracted starlight compose three direct error sources affecting the accuracy of navigation by stellar refraction.

39.3 Key Technology Analyses and Solution

39.3.1 Key Technology Analyses

Based on the four steps of navigation by stellar refraction mentioned above, the influence relation among the error sources of navigation positioning by stellar refraction can be analyzed and shown in Fig. 39.2.

It can be known from the Fig. 39.2 that the quantity and distribution uniformity of refracted stars in star map directly influence the positioning accuracy of navigation by stellar refraction. The position measuring accuracy of star affects the position accuracy of refracted star and measuring accuracy of refraction angle. Measuring accuracy of refraction angle and atmospheric model error jointly influence the tangent height error of refracted star, thus determining the positioning error of navigation by stellar refraction together with position accuracy of refracted star, and quantity and distribution uniformity of refracted stars. It can thus be seen that the initial error source affecting the positioning accuracy of navigation by stellar refraction involves the following three error sources: quantity and distribution uniformity of refracted stars, position measuring accuracy of star and atmospheric model error. In order to achieve the high-accuracy navigation by stellar refraction, it is necessary to realize the optimal state of the above three error sources, viz. minimizing their influence on navigation positioning error.

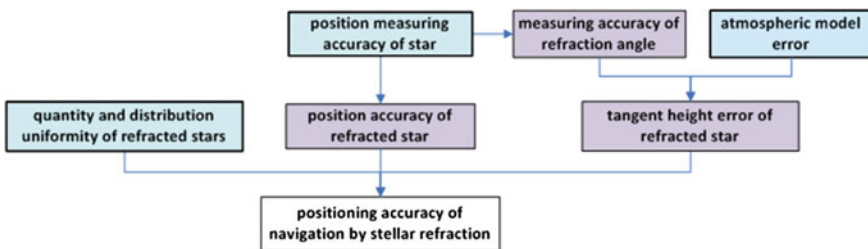


Fig. 39.2 Error sources of navigation positioning by stellar refraction

Full-time-running autonomous navigation of stellar refraction faces such problems as high dynamic range imaging of high-luminance earth and dim star, high SNR (Signal to Noise Ratio) shooting with dim star behind atmosphere lightened by the sun, etc. Therefore, three key technologies, viz. strong background star map acquisition, star extraction under strong background and atmospheric model, are to be solved for full-time-running navigation by stellar refraction under the premise of ensuring high-accuracy navigation. Such three key technologies will achieve full-time-running and high-accuracy navigation by stellar refraction by maximumly improving the quantity and distribution uniformity of refracted stars, position measuring accuracy of star and accuracy of atmospheric model. The realization approach and feasibility of these three key technologies are analyzed and discussed below.

39.3.2 Strong Background Star Map Acquisition

Strong background star map acquisition refers to the phenomenon that both refracted and un-refracted stars under strong background can be acquired simultaneously when the earth and atmosphere lightened by the sun appear in the field. Whether this technology can be realized is of vital importance to improving the quantity and distribution uniformity of refracted stars and developing full-time-running navigation by stellar refraction.

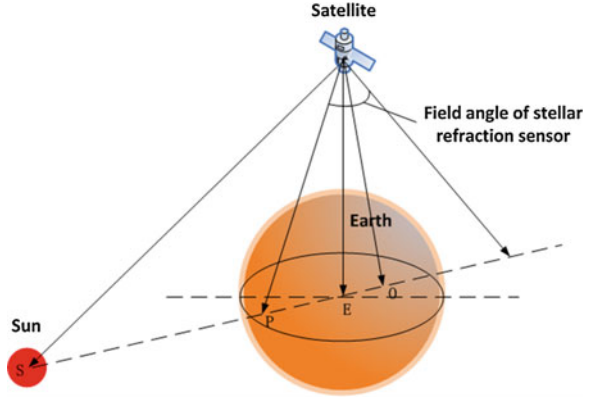
In order to acquire high SNR star map under strong background, we establish corresponding observation pointing model and detection limiting magnitude model based on refracted star, and adopt the detector with anti-blooming function to acquire high SNR star map under strong background.

39.3.2.1 Observation Pointing Model

Optimal satellite attitude planning is required to acquire the highest SNR of image. As the field angle of the earth to high orbit satellite is relatively small, directional observation to the earth should be adopted for stellar refraction sensor of high orbit satellite, while the optimal deflection detection to the earth should be adopted for low orbit stellar refraction sensor according to observation time and the earth's field angle. According to average density of different magnitude stars in the sky, in order to ensure enough un-refracted stars in the field to finish star map matching, attitude planning should be set by following three principles:

- (1) The field of un-refracted stars should not be less than 50 square degrees;
- (2) The minimum sun shading angle: 35° ;
- (3) Areas not shone by the sun shares observation priority.

Fig. 39.3 The schematic diagram of the field of stellar refraction sensor



By adopting attitude planning model and according to the calculated sun direction S , estimated geocenter direction E and the earth's field angle θ , the direction of field edge P and field center O can be calculated based on the above three principles, and finally the optimal satellite attitude at a certain moment can be obtained according to installation position of stellar refraction sensor. It is shown in Fig. 39.3 and the calculating approach is given below.

It can be seen from the principle of giving priority to observing the areas not shone by the sun that field edge P , sun direction S and geocenter direction E are in the same plane; in order to ensure the field of un-refracted star is not less than 50 square degrees, the included angle of vector P and vector E can be determined as $\varphi - \theta/2$, wherein, φ is the maximum earth's field angle allowing the earth to enter the field under the circumstance of ensuring the field for observing un-refracted star is not less than 50 square degrees. As the included angle of vector P and vector S is smaller than that of vector E and vector S , the expression of solving vector P is shown below:

$$\begin{cases} \vec{P} \cdot (\vec{E} \times \vec{S}) = 0 \\ \frac{\vec{P} \cdot \vec{E}}{|\vec{P}| \cdot |\vec{E}|} = \cos(\varphi - \theta/2) \\ (\vec{P} \times \vec{E}) \cdot (\vec{S} \times \vec{E}) > 0 \end{cases} \quad (39.1)$$

Similarly, the field center vector O and vectors S and P are in the same plane, and the included angle of vector O and vector P is $\phi/2$, wherein ϕ refers to the field angle of stellar refraction sensor; the included angle of vector O and vector S is larger than that of vector P and vector S . The expression of solving vector O is shown below:

$$\begin{cases} \vec{O} \cdot (\vec{P} \times \vec{S}) = 0 \\ \frac{\vec{P} \cdot \vec{O}}{|\vec{P}| \cdot |\vec{O}|} = \cos(\phi/2) \\ (\vec{P} \times \vec{O}) \cdot (\vec{S} \times \vec{P}) > 0 \end{cases} \quad (39.2)$$

Based on the calculated vector P and vector O and according to the installation position of stellar refraction sensor, the corresponding attitude angle of satellite can be calculated through coordinate transformation, to obtain the optimal planning attitude.

39.3.2.2 Detection Limiting Magnitude Model Based on Refracted Star

As star intensity increases with detection limiting magnitude, to improve the detection limiting magnitude of stellar refraction sensor will allow obtaining more background stars, thus improving navigation accuracy. The detection limiting magnitude of stellar refraction sensor is related to observing waveband, atmospheric background luminance, the sensor's performance and exposure time. Therefore, we first of all select a best waveband from the Johnson-Morgan-Cousins (UBVRI) metering system, and build a corresponding detection limiting magnitude model based on refracted star so as to provide reference for setting up proper exposure time while ensuring fast update rate of navigation data. The detection limiting magnitude model based on refracted star that we build based on SNR of not less than 5 as required by signal extraction and high-accuracy positioning is shown in Eq. (39.3).

$$\begin{aligned} m_s = & -1.88 - 1.25 \lg L(h, \theta) + 1.25 \lg t + 2.50 \lg \tau_{am}(h) \\ & - 2.50 \lg(\tan \phi) + 1.25 \lg(hc/\lambda) - 1.25 \lg n \\ & + 1.25 \lg(\tau \cdot QE) + 2.50 \lg f_0 + 2.50 \lg D \end{aligned} \quad (39.3)$$

In the above Equation, t is exposure time in unit of s, ϕ is detector's single-pixel aperture angle in degree, hc/λ is single-photon energy in J, n is number of pixels occupied by star, $\tau(\lambda)$ is system efficiency, $QE(\lambda)$ is detector's quantum efficiency, f_0 is flow of 0-magnitude stars in *number of photons/m²/s* in the best observing waveband, m_s is magnitude of refracted star and D is optical caliber of stellar refraction sensor in m.

$\tau_{am}(h)$ represents atmospheric transmittance in the best observing waveband and is obtained through modtran simulation based on the U.S. standard atmospheric model. It's shown in Eq. (39.4) with fitting error of 2×10^{-9} .

$$\tau_{atm}(h) = \sum_{i=1}^{13} p_i h^{13-i} \tag{39.4}$$

where, atmospheric height is $h \in [20, 80]$, in km. $p_1 = 1.58e-18$, $p_2 = -1.01e-15$, $p_3 = 2.92e-13$, $p_4 = -5.06e-11$, $p_5 = 5.84e-9$, $p_6 = -4.72e-7$, $p_7 = 2.74e-5$, $p_8 = -0.01e-1$, $p_9 = 0.03$, $p_{10} = -0.72$, $p_{11} = 9.93$, $p_{12} = -81.45$, $p_{13} = 3.00e+2$.

$L(h, \theta)$ is a relational model between atmospheric radiance and atmospheric height established based on modtran simulation analysis, the U.S. standard atmospheric model, and in consideration of general air quality and 0° solar zenith distance. It's shown in Eq. (39.5) with fitting error of $2 \times 10^{-9} w/m^2/sr$.

$$L(h, \theta) = \sum_{i=1}^{13} b_i h^{13-i} \tag{39.5}$$

where, atmospheric radiance $L(h, \theta)$ is in $w/m^2/sr$, atmospheric height is $h \in [20, 80]$ in km and solar zenith distance θ is in degree. $b_1 = 1.21e-18$, $b_2 = -7.61e-16$, $b_3 = 2.16e-13$, $b_4 = -3.69e-11$, $b_5 = 4.19e-9$, $b_6 = -3.34e-7$, $b_7 = 1.92e-5$, $b_8 = -8.00e-4$, $b_9 = 0.02$, $b_{10} = -0.50$, $b_{11} = 7.05$, $b_{12} = -59.47$, $b_{13} = 2.32e+2$.

39.3.2.3 Anti-blooming Function

The highest luminance of earth exceeds -22 magnitudes in the field, but at least 3 refracted stars are needed to achieve navigation by stellar refraction. Only when the detection limiting magnitude of refracted stars reaches 7.5, can it be ensured probability that 3 refracted stars appear in the effective field when running in the GEO orbit reaches 96.5 % or more. As a result, only using CCD featuring anti-blooming can achieve detecting high-luminance earth and 7.5-magnitude stars in the same field at the same time.

Feasibility of CCD's anti-blooming is analyzed as below.

The equation for calculating the number of photons generated against bright earth background receiving sunlight is:

$$X = \psi_b A q_b t \alpha^2 \tag{39.6}$$

where, ψ_b is background photon flux in $m^{-2}s^{-1}(\prime)^{-2}$ on unit pixel; A is receiving area of telescope in m^{-2} ; q_b is detector's average quantum efficiency against sky background; t is exposure time in S ; α^2 is angle area $(\prime)^2$ on unit pixel.

In order to enlarge effective observation field for navigation by stellar refraction in medium-high orbit, the earth tends to enter the field. It's analyzed that at this time, luminance of the earth receiving sunlight is not more than 4 mag/arc sec² while space-based observation finds that ψ_b , the photon flux generated by the Earth is not more than $7.264 \times 10^9 \times 10^{-4/2.5} m^{-2}s^{-1}arc\ sec^{-2}$. Based on the above

equation and current CCD development, it can be calculated that the number of background photons obtained by using 5 cm stellar refraction sensor and $t = 0.3$ s, a long exposure period, is less than 1.6×10^7 electric charge.

Currently, The full-well charge of CCD can reach 120,000 in general and the index of anti-blooming more than 10,000 times or 1.2×10^9 which is larger than 1.6×10^7 charge caused by bright earth background. Therefore, with anti-blooming, CCD can meet the demand for imaging in super-large dynamic range against bright earth background.

39.3.3 Star Extraction Under Strong Background

Because a large part of images collected in actual observation of navigation by stellar refraction are star maps against strong background, and the star positioning accuracy influences navigation accuracy by influencing error in tangent height, a key technology for stellar image extraction is star extraction under strong earth background, including high-accuracy calculation of stellar position. The method adopted for star extraction under strong earth background is shown as below:

(1) Calibrating star map

Excluding errors in star map due to inhomogeneity of observing system, specific operations include correcting CCD background, rectifying the instrumental pixel-pixel non-uniformity effect, marking bad pixels, etc.

(2) Deducting background

Gray values, which enter computer after CCD's potential well charge gets digitalized, not only come from lights given by the observed targets, but also contributions by skylight background and other ambient lights. In order to detect a faint target, noises from skylight background in different sky areas should be estimated accurately; large-scale background noises may drown local faint signals. In order to estimate skylight background, first of all, traverse the whole frame and deduct bright targets' impact on background estimation, then establish evaluation operators and finally establish a background image for the whole frame through iteration and fitting and deduct it from the whole frame.

(3) Target detection and extraction

The star map with background deducted is identified based on graphic characteristic parameters, and filtered. Meanwhile, SNR threshold is set for star extraction and stars whose SNR is larger than the threshold get extracted; the stars here include those both refracted and not refracted by atmosphere;

(4) Star centroid positioning

Because of pixel matching, diffraction effect and other factors, the observed targets tend to occupy more than one pixel on CCD. First of all, energy of candidate pixels should be extracted to obtain an energy distribution map for the targets. Adjacent pixels get clustered to form energy distribution map for respective targets. This study achieves high-accuracy star positioning through fitting method among methods for star centroid positioning and creates coordinates of targets in the XY coordinate system on CCD target surface.

Based on the above calculation method for obtaining star map under strong background and star extraction under strong earth background, field test on star extraction under strong background has been carried out in Xinglong Observatory of National Astronomical Observatories of China (NAOC), by observing the moon and stars around it. The test has verified that anti-blooming is feasible, and that through processing and analysis of the obtained star map and given the detection limiting magnitude of refracted star consistent with that calculated by Eq. (39.3), the detection limiting magnitude under strong background calculated by Eq. (39.3) can be regarded as valid. Besides, with fitting method, position measuring accuracy of star better than 0.12 pixels has been achieved.

39.3.4 Atmosphere Model

Atmosphere model error causes error in tangent height of refracted star in two ways: one by accuracy of refraction angle measurement acting on atmosphere model; the other by error of atmosphere model itself. This is why optimizing atmosphere model is critical to realizing high-accuracy navigation by stellar refraction.

39.3.4.1 Method of Optimizing Atmosphere Model

Since the angle of stellar refraction is directly related to atmospheric density at the tangential refraction height while the atmospheric density involves density elevation, Earth shape, seasonal variation, diurnal variation and other factors, a rather accurate atmospheric model should take into account impacts by seasonal variation, sunlight variation, Earth shape and other factors apart from density elevation [6].

The research results show that latitude and season are dominant elements for variations in atmospheric density, so density variation in atmosphere model can fall into two parts: system variation determined by latitude and season, and random error. The number of system variations can be obtained based on monthly average atmospheric density in the *USAF Reference Atmosphere*. Statistical properties of random error in atmospheric density can be obtained based on NASA's global reference atmosphere model. In order to further improve accuracy of stellar atmospheric refraction model and control its error range within 1%, we have proposed a

method of using satellite carrying stellar refraction sensor for in-orbit measurement so as to further correct stellar atmospheric refraction model. The equipment to build in-orbit atmospheric model with high accuracy and multiple variables include two high-accuracy stellar refraction sensors and a microwave altimeter.

39.3.4.2 Building Atmospheric Simulation Model

Based on the mechanism of atmospheric model error's impact on refracted star's tangent height error, an atmospheric model that reflects tangent height changing along with refraction angle should be built up during simulation, with atmospheric model errors entered. Through simulation analysis of multiple atmospheric models, an applicable stellar refraction atmosphere model based on atmospheric refractivity change and identical with stellar refraction rule is put forward. Its fitting equation is shown as below:

$$\gamma = 2350.1074e^{-0.10326788h} \quad (39.7)$$

where, h is atmospheric height in km; γ is refraction angle in ".

39.4 Simulation Analyses

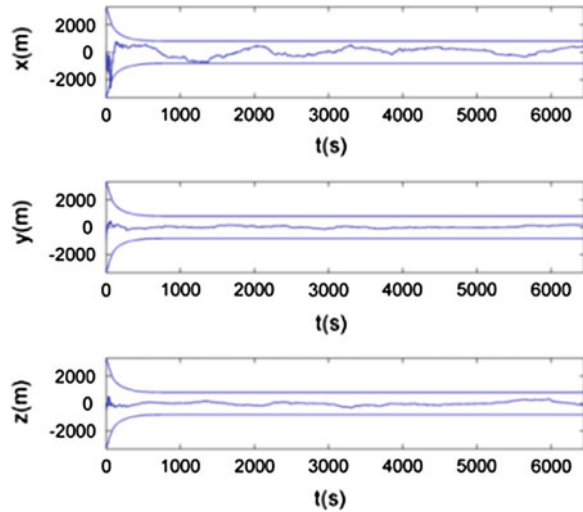
According to the approach to navigation by stellar refraction, observation orientation model, detection limiting magnitude of refracted star, atmospheric model and calculation method for star extraction under strong background realized in key technologies, a simulation system of navigation by stellar refraction has been built up, with simulation analysis of such navigation's accuracy shown as below.

39.4.1 Simulation Condition

Input condition of simulation system is as follows:

- (1) Simulation orbit parameters: $a = 24421.2$ km, $e = 0.726543$, $i = 28.5^\circ$, $\Omega = 139.3^\circ$, $\omega = 75^\circ$.
- (2) Technical indexes of stellar refraction sensor system:
 - FOV : $25^\circ \times 25^\circ$.
 - Detection sensitivity: 7.5 magnitude.
 - Stellar position precision: 3" (3σ).
- (3) Atmosphere model error: 1 %.
- (4) Atmosphere model, observation pointing model and detection limiting magnitude model based on refracted star: see article.

Fig. 39.4 Curve of changes in positioning precision of navigation by stellar refraction along with time



39.4.2 Simulation Results

Figure 39.4 is an estimated error curve of satellite position when the measurement accuracy of stellar refraction sensor is set as $3''$ and error of atmospheric model density as 1% based on the real stars distribution in the sky. It can be seen that the value of position estimation error of X, Y and Z has always been basically within 3σ error bound, suggesting fine consistency in estimation. When the simulation ends, the 3σ error bound for position estimation of X, Y and Z is about 1.01, 0.51 and 0.51 km respectively, with 1.2 km positioning accuracy.

Simulation Result Fig. 39.4 is obtained in full consideration of the influences of three main error sources, viz. quantity and distribution uniformity of refracted stars, position measuring accuracy of star and atmospheric model error, on navigation accuracy, and using 0° solar zenith distance, i.e., the earth's strongest exposure to sunlight. We can learn from it that the accuracy of full-time navigation by stellar refraction is better than 1.2 km.

39.5 Conclusions

Based on discussion on solutions to key technologies of full-time navigation by stellar refraction, and simulation analysis of navigation application system, we will next develop prototype of stellar refraction sensor and carry out HWIL simulation experiment on navigation by stellar refraction in order to lay a foundation for applying satellite's celestial navigation by stellar refraction.

Acknowledgments The authors are grateful for the support of the Xinglong Station, NAOC.

References

1. White RL, Gounley RB (1987) Satellite autonomous navigation with SHAD, AD-A184-988. The Charles Stark Draper Laboratory Inc, Cambridge
2. Xiao-lin N, Jian-cheng F (2007) An autonomous celestial navigation method for LEO satellite based on unscented Kalman filter and information fusion. *Aerosp Sci Technol* 11:222–228
3. White RL, Thurman SW, Barnes FA (1985) In: Proceedings of the forty-first annual meeting, the institute of navigation, Annapolis, MD, pp 83–89, 25–27 June 1985
4. Lair JL, Duchon P (1988) Satellite navigation by stellar refraction. *Acta Astronaut* 17 (10):1069–1079
5. Xiao-lin N, Long-hua W, Xinbei B et al (2013) Autonomous satellite navigation using starlight refraction angle measurements. *Adv Space Res* 51:1761–1772
6. Counley R, White R, Gai E (1984) Autonomous satellite navigation by stellar refraction. *J Guid Control Dyn* 7(2):129–134

Chapter 40

Experimental Investigation on CNS/SINS Integrated Navigation Using Star Tracker

Dangwei Wang and Lei Zhou

Abstract With the characteristics of good observing, simple structure and high reliability and so on, The CNS/SINS Integrated Navigation Using Star Tracker applies to seafaring and large-scale airborne platform in aerosphere. However, the star tracker can track only one star every time and celestial positioning depends on the performance of horizon measurement. Moreover, the horizon measurement of strapdown inertial navigation system is digital and low accurate. This brings on a reduced performance of CNS/SINS integrated navigation. Therefore, a new CNS/SINS method is advanced. With several observations of single star tracker, the attitude determination is realized. Finally, the CNS/SINS integrated navigation system is established. Both the simulation and experimental results demonstrate that the performance of this integrated navigation system is excellent. In this way, the output divergence is restrained and the accurate of positioning and heading is improved. In a word, the CNS/SINS method applies to the long-distance and long-endurance airborne platform.

Keyword Airborne CNS/SINS integrated navigation · Star tracker · SINS

40.1 Introduction

In CNS/INS integrated navigation system, inertial navigation system (INS) with advantages of positioning, velocity and attitude output successively, high precision in short time and total independence is the main navigation system in military application. However, the INS is affected by the error drift in Inertial measurement unit (IMU), which bring about the performance of INS deteriorates by time. So the

D. Wang (✉) · L. Zhou
The 20th Research Institute of CETC, Xian 710068, China
e-mail: wangdw629@163.com

L. Zhou
e-mail: zhoulei101213@qq.com

INS does not apply to large-scale airborne platform. On the contrary, the celestial navigation system (CNS) can output navigation information without error drift, since the star in space is predictable and stable. Moreover, the CNS is the absolutely necessary navigation mode in modern high-tech war. But the CNS is affected easily by the atmosphere for airborne, which result in navigation outputs be disconnected. Therefore, the CNS is an accessorial navigation system and Integrated with the INS [1, 2].

According to the difference of field of view, the CNS contains narrow-field star tracker (such as sextant) and wide-field star sensor. The attitude determination technology based on wide-field star sensor has been applied to aerospace territory. But for the aviation application, the star sensor can't detect star at daytime because of strong noise of sky background and it is still in the theoretical researching stages. However, the star tracker can overcome these limits with narrow-field, single star detected every time, high signal-to-noise and contrast [3–6].

Based on the advantages of small cubage, light weight and stabilization for strapdown inertial navigation system (SINS), the CNS/SINS Integrated Navigation is proposed and develops quickly. However, the horizontal measurement of SINS is digital and virtual. This leads to the horizontal measurement is affected easily by the platform error of SINS. Moreover, the celestial positioning method depends on the performance of horizontal measurement. Therefore, the platform error of SINS must be adjusted for the realization of CNS/SINS Integrated Navigation. Meanwhile, only one star can be detected every time by the star tracker. This does not satisfy the prerequisites of attitude determination, so there is a bottleneck.

A new attitude determination algorithm is presented in this paper aim at the problems above. The keystone of this technology is that the attitude of static body is determined by the asynchronous detection strategy of star tracker. Then, the platform error of SINS is estimate by kalman filter and compensated by closed-loop feedback.

The paper is organized in five sections. The theory of CNS/SINS integrated navigation is presented in Sect. 40.2. The field experiment setup is given in Sect. 40.3. The experimental results and discussions are placed in Sect. 40.4. And Sect. 40.5 gives conclusion.

40.2 Theories

40.2.1 Theories of Attitude Determination Using Star Tracker

According to the characteristics of star tracker, the asynchronous detection strategy for static body is shown as follows:

The star tracker observes a star at time t_j , where t_j is $t_1, t_2, \dots, t_M (M \geq 2)$ with a tap of dt_j , where M stands for the number of stars detected for one calculation. Then

the elevation angle H_j and the azimuth AZ_j angle under body frame b , the right ascension RA_j and the declination Dec_j under inertial frame i of the corresponding star is output by star tracker, the vernal equinox Greenwich hour angle $GAST_j$ of corresponding time t_j is calculated. The Greenwich hour angle GHA_j is obtained:

$$GHA_j = GAST_j + 360^\circ - RA_j \quad (40.1)$$

We can think that the stars goes around the earth with earth rotation angular velocity under earth frame e . Then the GHA_j and the Dec_j can be stand for the star. Because the body keeps still, we can think that a new star is detected for any time t_j under earth coordinate. Therefore, several stars observed by asynchronous detection at any time can be regarded as a known matching combination as long as the body keeps still. In other words, the attitude of the body can be obtained by the way used by star sensor.

Based on the analysis above, the right angle coordinate vector of star under earth frame is calculated:

$$T_j = \begin{cases} -GHA_j & GHA_j < 360^\circ \\ 360^\circ - GHA_j & GHA_j \geq 360^\circ \end{cases} \quad (40.2)$$

then:

$$\begin{cases} x_{Gj} = \cos(Dec_j) \cos(T_j) \\ y_{Gj} = \cos(Dec_j) \sin(T_j) \\ z_{Gj} = \sin(Dec_j) \end{cases} \quad (40.3)$$

where (x_{Gj}, y_{Gj}, z_{Gj}) is considered as a vector $\mathbf{G}_j = (x_{Gj}, y_{Gj}, z_{Gj})$. In the same way, the celestial detection vector $\mathbf{S}_j = (x_{sj}, y_{sj}, z_{sj})$ under body system b is obtained:

$$\begin{cases} x_{sj} = \cos(H_j) \cos(Az_j) \\ y_{sj} = \cos(H_j) \sin(Az_j) \\ z_{sj} = \sin(H_j) \end{cases} \quad (40.4)$$

Define matrix:

$$\begin{cases} F_S = (\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_M)^T \\ F_G = (\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_M)^T \end{cases} \quad (40.5)$$

when $M > 2$, generate direction cosine matrix from body frame b to earth frame e by least square method (LSM):

$$C_b^e = (F_G^T F_G)^{-1} F_G^T F_S \tag{40.6}$$

when the number of stars detected for one calculation is $M = 2$, consider $S_j = (x'_{sj}, y'_{sj}, z'_{sj})$ and $G_j = (x'_{Gj}, y'_{Gj}, z'_{Gj})$. Then:

$$\begin{cases} S'_1 = S_1, & S'_2 = \frac{S_1 \times S_2}{\|S_1 \times S_2\|}, & S'_3 = \frac{S'_1 \times S'_2}{\|S'_1 \times S'_2\|} \\ G'_1 = G_1, & G'_2 = \frac{G_1 \times G_2}{\|G_1 \times G_2\|}, & G'_3 = \frac{G'_1 \times G'_2}{\|G'_1 \times G'_2\|} \end{cases} \tag{40.7}$$

And consider $F_S = (S'_1, S'_2, S'_3)^T$ and $F_G = (G'_1, G'_2, G'_3)^T$. Then C_b^e is obtained:

$$C_b^e = F_G^{-1} F_S \tag{40.8}$$

Meanwhile, get the local position (λ_M, L_M) according to the SINS, then calculate the direction cosine matrix from earth frame e to navigation frame n :

$$C_e^n = \begin{bmatrix} -\sin \lambda_M & \cos \lambda_M & 0 \\ -\sin L_M \cos \lambda_M & -\sin L_M \sin \lambda_M & \cos L_M \\ -\cos L_M \cos \lambda_M & \cos L_M \sin \lambda_M & \sin L_M \end{bmatrix} \tag{40.9}$$

Integrate formula (40.6, 40.8 and 40.9) and calculate the direction cosine matrix from body frame b to navigation frame n

$$C_b^n = C_e^n C_b^e \tag{40.10}$$

Then the Euler angle vector $(\gamma_c \ \theta_c \ \varphi_c)^T$ from body frame b to navigation frame n is obtained according to C_b^n and paper [7], where γ_c is roll, θ_c is pitch, and φ_c is yaw.

40.2.2 CNS/SINS Integrated Navigation

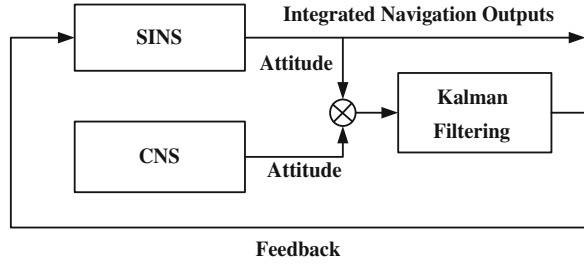
Combining with corresponding $(\gamma_s, \theta_s, \varphi_s)$ SINS outputs, the attitude determination errors between SINS and CNS is calculated as follows:

$$Z(t) = (\gamma_s - \gamma_c, \theta_s - \theta_c, \varphi_s - \varphi_c)^T \tag{40.11}$$

Then, the CNS/INS integrated navigation system is setup, including linear kalman filter and closed-loop feedback [8, 9]. The case of the integrated navigation system is shown in Fig. 40.1.

Where observed quantity of kalman filter is $Z(t)$ given in formula (40.11), the state equation is the error equation of SINS, the navigation frame is the

Fig. 40.1 CNS/SINS integrated navigation system



east- north-up local-level-frame. According to the quality of SINS, the state equation can be written in the form of matrix [10]:

$$\dot{X}(t) = F(t)X(t) + G(t)W(t) \tag{40.12}$$

$$X = (\phi_e, \phi_n, \phi_u, \delta v_e, \delta v_n, \delta v_u, \delta L, \delta \lambda, \delta h, \varepsilon_{bx}, \varepsilon_{by}, \varepsilon_{bz}, \varepsilon_{rx}, \varepsilon_{ry}, \varepsilon_{rz}, \nabla_x, \nabla_y, \nabla_z) \tag{40.13}$$

where ϕ_e, ϕ_n, ϕ_u are the platform errors of SINS, $\delta v_e, \delta v_n, \delta v_u$ are velocity errors, $\delta L, \delta \lambda, \delta h$ are positioning errors, $\varepsilon_{bx}, \varepsilon_{by}, \varepsilon_{bz}, \varepsilon_{rx}, \varepsilon_{ry}, \varepsilon_{rz}$ are errors of gyroscope, and $\nabla_x, \nabla_y, \nabla_z$ are errors of accelerometer. Meanwhile, the observational equation of fixing celestial attitude is given.

$$Z(t) = H(t)X(t) + V(t) \tag{40.14}$$

where $H(t)$ can be written as $H(t) = [C_T^s \quad 0_{3 \times 15}]$, and C_T^s shown in Eq. (40.15) is the transformation matrix from attitude errors to platform errors of SINS, and $V(t)$ standing for observational noise is given in Eq. (40.16).

$$C_T^s = \begin{pmatrix} -\cos \psi & \sin \psi & 0 \\ \cos \theta & \cos \theta & 0 \\ \sin \psi & -\cos \psi & 0 \\ -\cos \psi \operatorname{tg} \theta & -\sin \psi \operatorname{tg} \theta & -1 \end{pmatrix} \tag{40.15}$$

$$V(t) = [N_E \quad N_N \quad N_U]^T \tag{40.16}$$

With desecrating Eqs. (40.12)–(40.14) and obtaining the platform errors after kalman filtering, then the attitude transformation matrix of SINS is compensated:

$$C_b^n = (I + \Phi^n)C_{bM}^{nM} \tag{40.17}$$

where I is the 3×3 unit matrix and Φ^n is shown in Eq. (40.18).

$$\Phi^n = \begin{pmatrix} 1 & -\phi_u & \phi_n \\ \phi_u & 1 & -\phi_e \\ -\phi_n & \phi_e & 1 \end{pmatrix} \tag{40.18}$$

40.3 Field Experimental Setup

The experimental system is setup at an observatory.

As shown in Fig. 40.2, the experimental system contains turntable, SINS, star tracker, GPS receiver and integrated processor. The IMU and star tracker is fixed on the turntable that can adjust and measure attitude of star tracker. Meanwhile, the transformation relationship between IMU and star tracker is determined by high precision installation. The drift error of gyroscope is $0.02^\circ/\text{h}$, and the drift error of accelerometer is $50 \mu\text{g}/\text{h}$.

The work procedures of this experimental system contain that integrated processor receives and processes the outputs of SINS, star tracker and GPS receiver, then the outputs of integrated processor feedback into the SINS.

40.4 Results and Discussion

40.4.1 Results of Celestial Detection

The experiment obtained a great deal of data. One of the clutches of data is an successive observation for 4 h. There were three stars had been detected, which are star No.3 (Arcturus), star No.5 (Vega), and star No.12 (Altair). And the results are shown as follow.

Fig. 40.2 Field experimental system

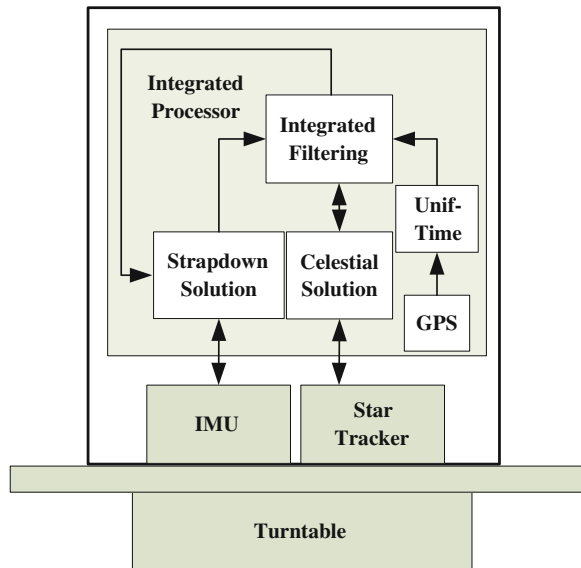
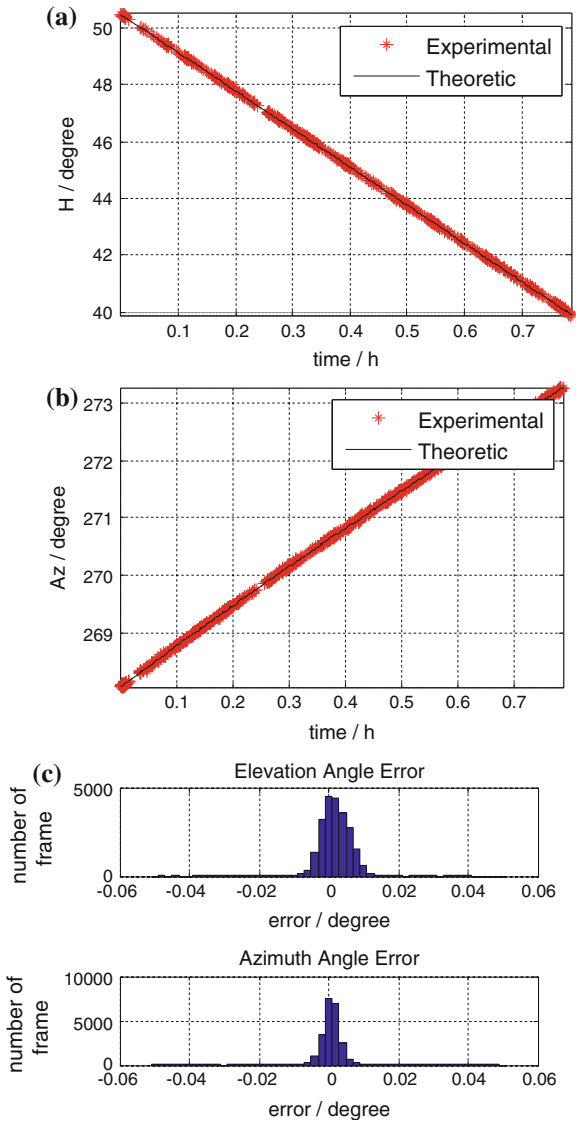


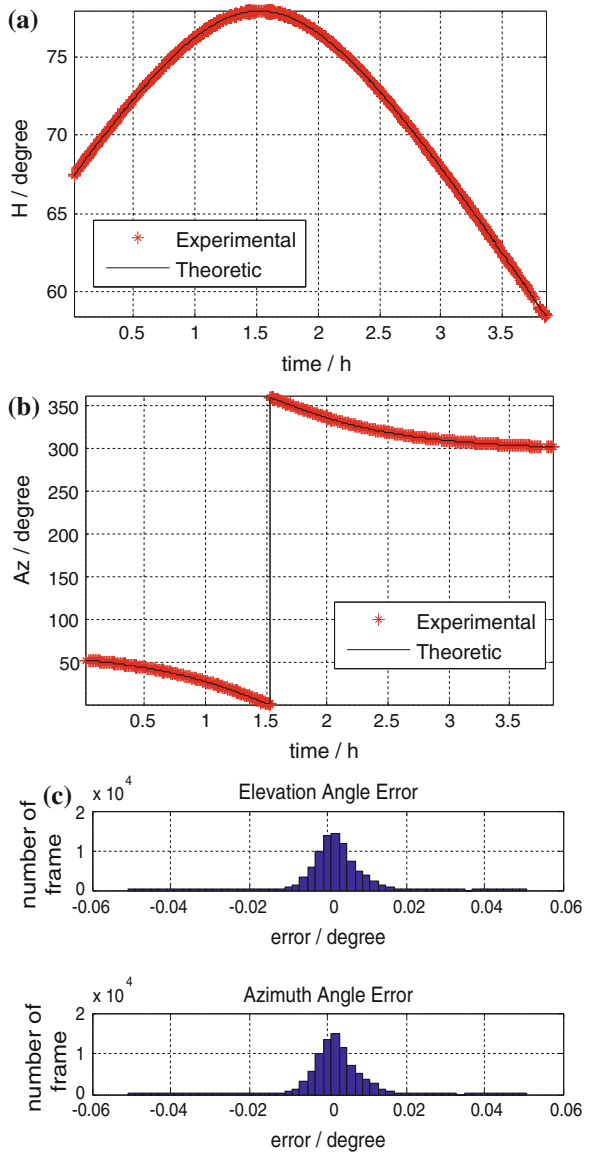
Fig. 40.3 Detected results of Arcturus. **a** Elevation angle, **b** Azimuth angle, **c** detection errors



In Fig. 40.3, the Elevation angle and Azimuth angle output from star tracker coincide with the theory results perfectly. And the detection errors of star No.3 are $RMS_H = 17.43''$, $RMS_{Az} = 15.07''$.

The observation of star No.5 goes through the whole period of 4 h. The detection errors of star No.5 are $RMS_H = 21.92''$, $RMS_{Az} = 20.93''$. It deteriorates a little compared with star No.3.

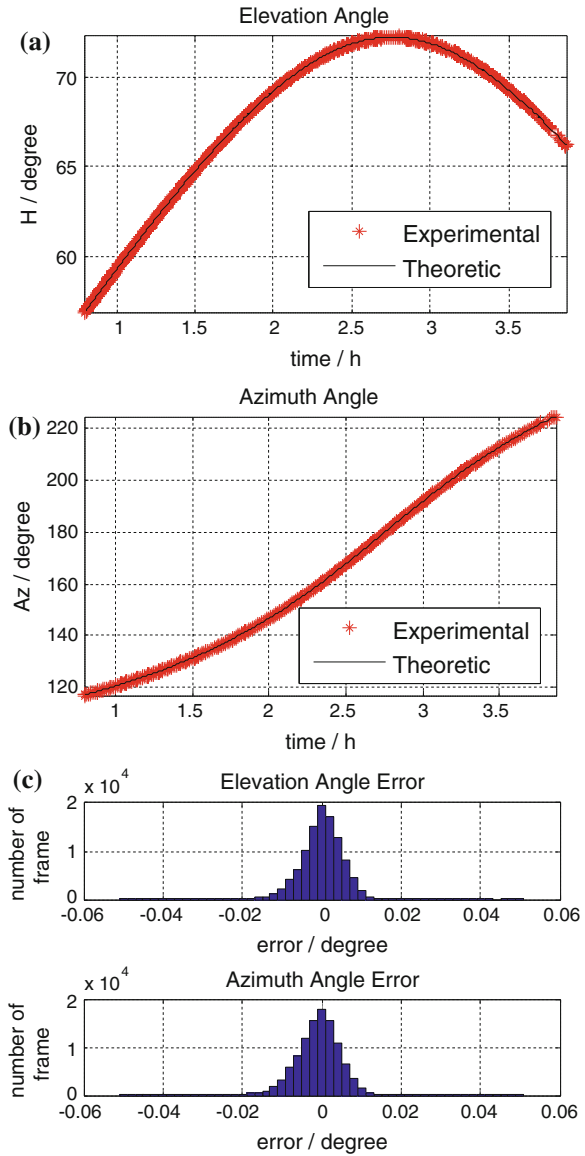
Fig. 40.4 Detected results of Vega. **a** Elevation angle, **b** Azimuth angle, **c** detection errors



The same with the star No.5, The observation of star No.12 goes through the whole period of 4 h. And the detection errors are $RMS_H = 23.98''$, $RMS_{Az} = 25.95''$ (Figs. 40.4 and 40.5).

In a word, the performance of star tracker is excellent.

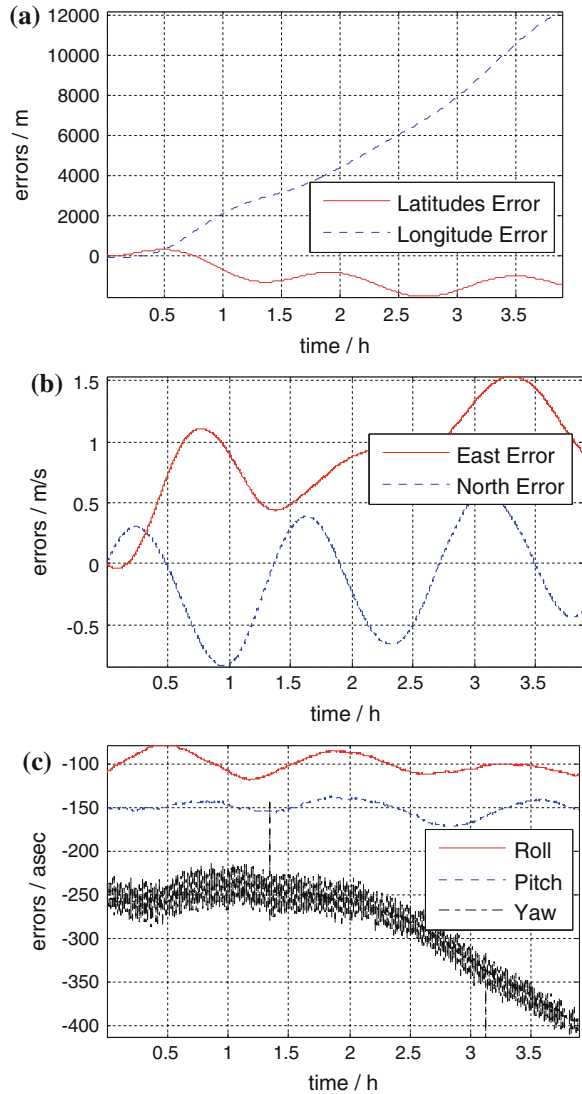
Fig. 40.5 Detected results of Altair. **a** Elevation angle, **b** Azimuth angle, **c** detection errors



40.4.2 Outputs of SINS

According to the outputs of SINS in course of the experiment are shown in Fig. 40.6, the performance of the SINS is ordinary. The circular probable positioning error is $CEP = 1.66$ nmile/h and the maximum value runs up to 6.5 nmile. The attitude error is 0.02° approximately. What's more, the yaw error is diverging and the maximum value runs up to 0.12° in 4 h.

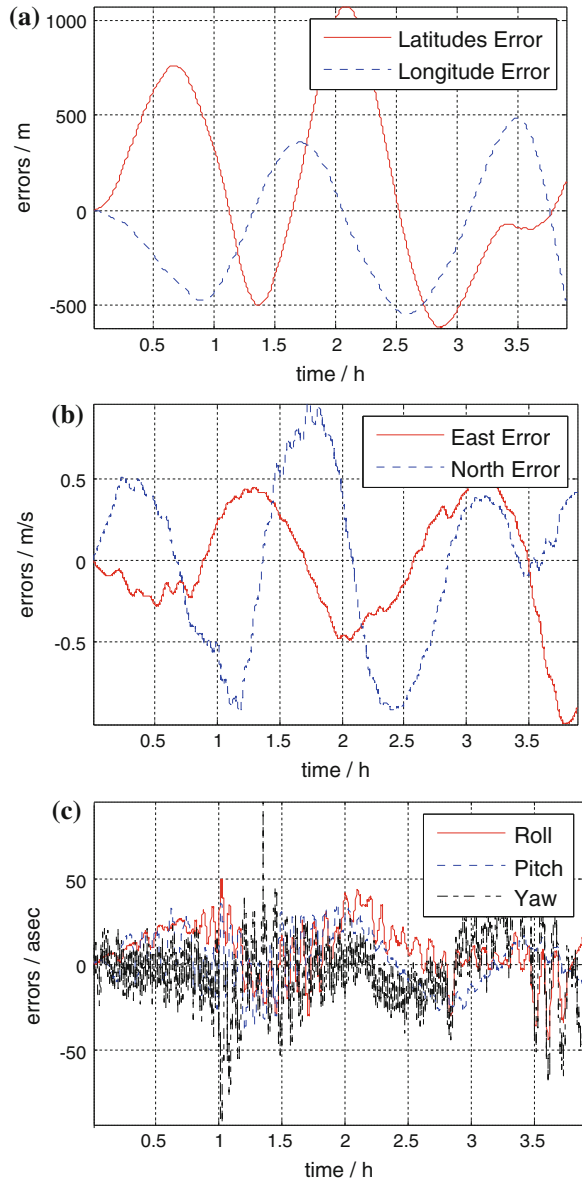
Fig. 40.6 Results of SINS
a Positioning error, **b** velocity errors, **c** attitude errors



40.4.3 Results of CNS/SINS Integrated Navigation

Based on the performance of celestial observation and SINS shown above, the results of CNS/SINS integrated navigation are obtained. As shown in Fig. 40.7, the circular probable positioning error is $CEP = 485.39$ m in 4 h, and the maximum value is 1000 m, the east velocity error is $RMS_{east} = 0.38$ m, the north velocity error is $RMS_{north} = 0.52$ m, and the roll error is $RMS_{\gamma} = 18.7''$, the pitch error is $RMS_{\theta} = 18.9''$, the yaw error is $RMS_{\phi} = 20.64''$. Moreover, in contrast to the

Fig. 40.7 Results of CNS/SINS integrated navigation. **a** Positioning error, **b** velocity errors, **c** attitude errors



performance of SINS, the positioning and yaw error of integrated navigation is convergent. However, there are slight oscillations in yaw error because of the instability of the turntable on the horizontal direction.

In general, the performance of CNS/SINS integrated navigation is excellent, not to mention much better than SINS.

40.5 Conclusion

In this paper, a new CNS/SINS algorithm based on star tracker is proposed. The attitude determination is realized for static body and the platform error of SINS is compensated by a closed-loop after kalman filtering. Moreover, a field experimental system is established. The results demonstrate that the performance of CNS/SINS integrated navigation using star tracker is excellent, and the positioning error is less than 500 m in 4 h. The future work will focus on the dynamic experiments and the improvement of performance of this integrated navigation system.

References

1. Qu Q (2011) Research on key technique of airborne SINS/CNS integrated navigation system. Nanjing University of Aeronautics and Astronautics, The Graduate School, pp 10–28
2. Van Allen JA (2004) Basic principles of celestial navigation. *Am J phys* 72(11):1418–1424
3. Sun HC (2002) Study of producing small and light star tracker. China academy and sciences, Beijing, pp 31–52
4. Gruman N (2006) LN-120G Stellar-Inertial navigation System
5. Ted H, Doug A, Mitch M (1996) The B2 stealth bomber integrated navigation system. In: 23rd joint service data exchange for guidance, navigation and control, pp 1–10
6. Atkinson D, Agnew J, Miller M (1993) The B-2 navigation system. In: Aerospace and electronics conference, Proceeding of the IEEE 1993 nation
7. Fang JC, Ning XL (2006) Theory and application of celestial navigation. Beijing University of Aeronautics and Astronautics, Beijing, pp 69–90
8. Qin YY, Zhang HJ, Wang SH (1998) Kalman filter and integrated navigation system. Northwestern Polytechnical University Press, Xian, pp 33–75
9. Kalman RE (1960) A new approach to linear filtering and prediction theory. *Trans ASME J Basic Eng* 82D:35–46
10. Anderson BD, Moore JB (1979) Optimal filtering. Prentice-Hall, New Jersey

Chapter 41

Simultaneous Ranging and Communication Based on X-Ray Communications

Shibin Song, Luping Xu, Hua Zhang, Yuanjie Bai and Liyan Luo

Abstract To satisfy the increasing requirement for accurate range measurement and high rate data transmission, a new simultaneous ranging and communication system based on X-ray communications (XCOM) is proposed. To achieve simultaneous range measurement and data communication, a specific signal structure is designed. The designed signal structure guarantees the simultaneous transmission of the ranging and communication information. By matching the local sequence number of the signal symbol with the sequence number propagating through the two-way range, the range can be calculated. In this paper, the basic principles of the proposed method is described and the related key techniques are analysed in detail. Besides, the performance of the proposed method is analysed both theoretically and numerically. The analysis covers the data rate, the acquisition probability, the ranging jitter, etc. Compared with the available methods, the proposed method improves the link efficiency and is helpful for system miniaturization and integration, which provides an option for future deep space explorations.

Keywords Simultaneous ranging and communication · X-ray communications · Deep space exploration · Acquisition probability · Bit error rate

S. Song · L. Xu (✉) · H. Zhang · Y. Bai · L. Luo
Xidian University, Xi'an Xifeng Road 266, Xi'an 710126, China
e-mail: lpxu@mail.xidian.edu.cn

S. Song
e-mail: shbsong@yeah.net

H. Zhang
e-mail: zhanghua@mail.xidian.edu.cn

Y. Bai
e-mail: byjck666@163.com

L. Luo
e-mail: xiaoyan12027@163.com

Y. Bai
Xi'an Microelectronics Technology Institute, Xi'an Taiyi Road, Xi'an 710054, China

41.1 Introduction

With the further exploration of the space and the increasing exploration activities, the demand for large volume of data transmission has become a popular issue in deep space explorations [1]. Meanwhile, the range measurement, an essential observable in space explorations, has found its application in various areas, such as navigation [2, 3], satellite constellation configuration [4, 5], etc. In addition, as the circumstances of deep space change rapidly, the duration of available link is limited and highly efficient utilization of the link is of great importance. Simultaneous ranging and communication not only saves the power and bandwidth, but also makes the best use of the link availability.

There has been references concerned with developing the composite system for simultaneous ranging and communication [6–8]. However, present ‘composite systems’ are mainly time sharing systems, which conduct the communication and the ranging separately, resulting in a low link efficiency. In Ref. [9], Andrews et al. proposed a telemetry-based ranging method, which keeps the uplink signal and replaces the downlink signal by a telemetry signal. By analyzing the time information transmitted by the downlink telemetry signal, the range could be obtained. The flaw is that data can only be transmitted along the downlink, which limits its application. Sun et al. presented the simultaneous ranging and communication between the satellite laser ranging (SLR) station at NASA Goddard Space Flight Center (GSFC) and the Lunar Reconnaissance Orbiter (LRO) [10]. The communication was carried out by a 4096-ary pulse position modulation and the ranging measurement was conducted by recording the pulse arrival time and sending it back to the emitter. In the case, the system needs an extra microwave link to feedback the time information, which occupied the link power and bandwidth and requires extra microwave equipment.

Considering the flaws of the available methods, a simultaneous ranging and communication method based on X-ray communications (XCOM) is proposed and analyzed. XCOM is a new concept for deep space communications [11], which utilizes man-made X-ray as the information carrier. Different from the available methods, the proposed method conducts the range measurement and the data communication simultaneously utilizing a unique signal format in both uplink and downlink, which makes better use of the space link. Besides, the equipment related to X-ray emission and reception is small and energy-saving, and the communication and the ranging share the same equipment. Thus, it is possible to miniaturize the system.

41.2 Principles of the Proposed Method

41.2.1 Basic Principle

In the range measurement, the range calculation is achieved by transmitting the ranging information between objectives. Therefore, the range measurement is

actually a ‘communication procedure’. Thus, it is possible to perform the range measurement and the communication simultaneously.

As to the signal structure, the uplink signal symbol consist of two parts: the uplink symbol identifier and the data communication region. The uplink symbol identifier comprises two sections: the synchronization sequence and the sequence number (SN). The synchronization sequence is used for the acquisition of the symbol, which generally has fine autocorrelation property. SN is the unique identifier of the symbol, utilized for the time synchronization for the range measurement. The downlink link has the same signal symbol structure as that of the uplink. By transmitting the SN to the slave station and sending it back to the master station, the transmitted SN is paired up with the local SN to acquire the two-way light travel time, τ . Finally, the two-way range can be calculated

$$R = c\tau/2 \tag{41.1}$$

where c is the speed of light.

41.2.2 System Structure

The signal processing at the master station is presented in Fig. 41.1. The master station and the slave station share some similarities in the signal process. Thus, we take the signal processing of the master station as an example to present the key points of the proposed method.

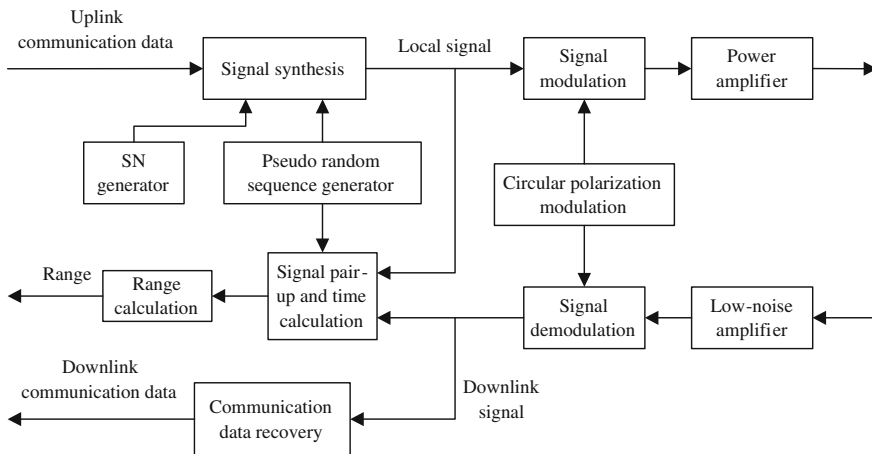


Fig. 41.1 Function block diagram of the master station

41.2.2.1 Modulation Scheme

The circular polarization modulation method is adopted to modulate the signal. Polarization states [12] show fine stability in deep space propagation [13, 14], which makes it suitable for space communication [15, 16]. The principle of the circular polarization modulation is to change a binary sequence into a sequence of circularized states. Each circular polarization state have a time duration of T , which is called a slot. To describe the circularly polarized signal, the Stokes vector is adopted

$$\mathbf{S}(A_x, A_y, \delta) = [A_x^2 + A_y^2 \quad A_x^2 - A_y^2 \quad 2A_xA_y \cos \delta \quad 2A_xA_y \sin \delta]^T \quad (41.2)$$

A_x and A_y are the amplitude of the light vector component and δ is the phase error between the two light vector components, i.e., $\delta = \theta_x - \theta_y$ and δ determines the type of the polarization states. θ_x and θ_y are the phase of the light vector components indicated in Fig. 41.2.

The circularly polarized signal at the side of the transmitter can be presented as

$$\mathbf{S} = \{\mathbf{S}_i(A_x, A_y, \Psi(s(t)))\}, i \in N, i < M \quad (41.3)$$

where $\Psi(\cdot)$ is the phase error and $s(t)$ is the transmitted signal, defined as

$$s(t) = \sum_i d_i g(t - iT) \quad (41.4)$$

$\{d_i\}$ is the signal in the binary form and $g(t)$ is the gate signal defined as

$$g(t) = \begin{cases} 1 & t \in [0, T] \\ 0 & \text{others} \end{cases} \quad (41.5)$$

41.2.2.2 Signal Pair-Up and Time Calculation

By pairing up the SNs between the local SN and the SN travelling through the two-way range, we could achieve the time synchronization between the master station and the slave station. For the i th SN, denoted as SN_i , is generated at the master

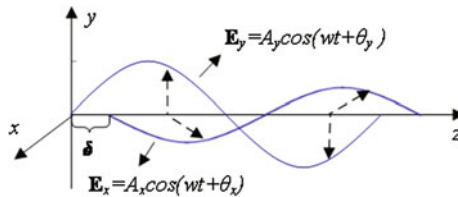


Fig. 41.2 Light vector components of polarized light

station and assigned to an uplink symbol. SN_i satisfies $SN_i \in [0, 2^{l_{SN}} - 1]$ and $SN_i \in N$. To some extent, SN_i can be used as the unique identifier for an uplink symbol. Based on Fig. 41.1, when the symbol denoted by SN_i is transmitted at the master station, its emitting time is recorded. Through space transmission, SN_i is received and extracted from the symbol at the slave station. Then, exactly the same SN_i is reassigned to the downlink symbol through the regeneration process and sent back along with the downlink symbol. At the master station, the arrival time is calculated by correlating the downlink signal with the local synchronization sequence. Then, by pairing up the local SN with the SN extracted from the downlink symbol, we can know the emitting time and arrival time of the specific symbol. Consequently, we can obtain the light travel time, τ , of SN_i over the two-way range.

41.3 Performance Analysis of the Proposed Method

41.3.1 Noise on Signal Synchronization

To acquire the signal symbol, an m-sequence, whose length is l_{SS} , is adopted as the synchronization sequence. The synchronization sequence is defined as

$$s_{ss}(t) = \sum_{i=0}^{l_{ss}} m_i g(t - iT) \quad (41.6)$$

where $\{m_i, i = 1, \dots, l_{SN}\}$ is the binary code of the m-sequence.

Being through the space, $s_{ss}(t)$ is influenced by noise, which has the form

$$s'_{ss}(t) = \sum_{i=0}^{l_{ss}} m_i g(t - iT) + n(t) \quad (41.7)$$

Generally, $n(t)$ is modelled as a zero-mean additive Gaussian white noise with a two-sided power spectral density of $N_0/2$. The acquisition of the signal symbol is achieved by correlating the received synchronization sequence with the local synchronization sequence. The correlation process can be presented as

$$\begin{aligned} R_{corr}(k) &= \int_0^{T_{corr}} s'_{ss}(t) s_{ss}(t - kT) dt \\ &= \int_0^{T_{corr}} s_{ss}(t) s_{ss}(t - kT) dt + \int_0^{T_{corr}} n(t) s_{ss}(t - kT) dt \\ &= \rho_{ss}(k) + \zeta_{ss}(k) \end{aligned} \quad (41.8)$$

where T_{corr} is the correlation time. ζ_{SS} can be regarded as the influence of the noise on the correlation results

$$\zeta_{SS}(k) = \int_0^{T_{corr}} n(t)s_{SS}(t - kT)dt \tag{41.9}$$

$\zeta_{SS}(k)$ is a Gaussian variable with zero mean value. As $n(t)$ has a two-sided power spectral density of $N_0/2$, the variance of $\zeta_{SS}(k)$ is $N_0T_{corr}/2$.

To simulate the effect of noise on the correlation, m-sequences of different orders, denoted by n , are adopted and additive Gaussian noises with different power levels are utilized as the noise source. Then, the noise-interfered m-sequence is correlated with the local sequence. The slot duration is set to $T = 10$ ns. Figure 41.3a shows the average peak values of the correlation results. For each SNR, the peak values of the correlation are averaged over 100 trials. It can be observed in Fig. 41.3a that at low level of SNR, the peak values are high and fluctuant. The high peak value is caused by the high noise power. Figure 41.3b presents the peak ratio of the correlation results under different SNR levels. Here, the peak ratio is defined as $p_r = p_{in-phase}/p_{max}$, where $p_{in-phase}$ is the correlation value when the two sequences are in-phase, and p_{max} is the peak value of the correlation results excluding $p_{in-phase}$. Similarly, the correlation process is repeated for 100 times at each SNR and the peak ratio is calculated. Then, the peak ratios are averaged over the trials. With the increasing SNR, the peak ratio increases accordingly, meaning that with high SNR the peak value can be clearly separated. At low SNR, the peak ratio is small because of the influence of the noise.

Besides of the correlation results, we adopt the acquisition probability, P_{acq} , to evaluate the acquisition performance. P_{acq} is defined as the successful acquisition of the peak value of the correlation results. Thus, P_{acq} can be written as

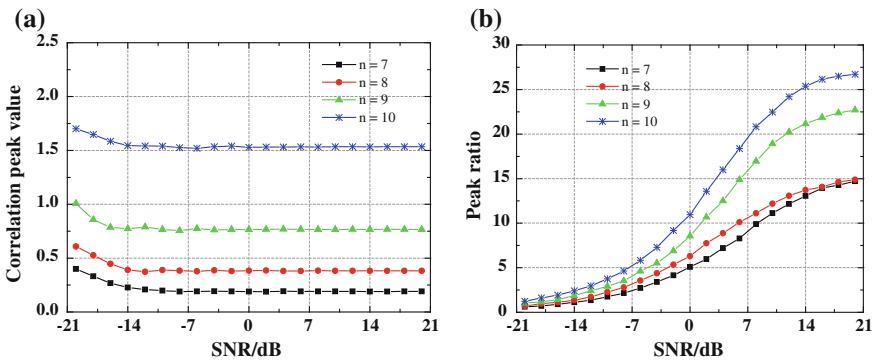


Fig. 41.3 Effects of noise on synchronization. **a** Average correlation peak values. **b** Peak ratio of the correlation results

$$\begin{aligned}
P_{acq} &= P(R(0) > \max\{R(k), 0 < k \leq l_{SN} - 1\}) \\
&= P(R(0) > R(k), 0 < k \leq l_{SN} - 1) \\
&= \prod_{k=1}^{l_{SN}-1} P(R(0) > R(k)) \\
&= \prod_{k=1}^{l_{SN}-1} P(\xi_{SS}(k) - \xi_{SS}(0) < \rho_{SS}(0) - \rho_{SS}(k))
\end{aligned} \tag{41.10}$$

It should be noted that $\xi_{SS}(0)$ and $\xi_{SS}(k)$ are not independent. The reason is that $s_{SS}(t - kT)$ and $s_{SS}(t)$ are coupled because of the time delay. Since the length of the synchronization sequence is limited and the influence of the coupling on the integration can be ignored, we assume that $\xi_{SS}(0)$ and $\xi_{SS}(k)$ are independent. Therefore, $\xi_{SS}(k) - \xi_{SS}(0)$ follows a Gaussian distribution of $N(0, N_0 T_{corr})$, and

$$\begin{aligned}
&P(\xi_{SS}(k) - \xi_{SS}(0) < \rho_{SS}(0) - \rho_{SS}(k)) \\
&= \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{\rho_{SS}(0) - \rho_{SS}(k)}{\sqrt{2N_0 T_{corr}}}\right)
\end{aligned} \tag{41.11}$$

where $\operatorname{erf}(\cdot)$ is the error function. Then, Eq. (41.10) can be rewritten as

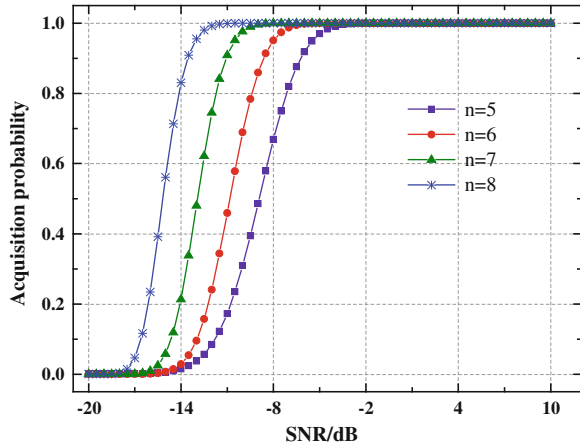
$$\begin{aligned}
P_{acq} &= \prod_{k=1}^{l_{SN}-1} \left\{ \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{\rho_{SS}(0) - \rho_{SS}(k)}{\sqrt{2N_0 T_{corr}}}\right) \right\} \\
&= \frac{1}{2^{l_{SN}-1}} \prod_{k=1}^{l_{SN}-1} \left\{ 1 + \operatorname{erf}\left(\frac{\rho_{SS}(0) - \rho_{SS}(k)}{\sqrt{2N_0 T_{corr}}}\right) \right\}
\end{aligned} \tag{41.12}$$

Figure 41.4 presents the variation of the signal symbol acquisition probability with SNR. As shown in Fig. 41.4, when the SNR is low, the acquisition probability is small because of the noise interference. With the increasing SNR, the acquisition probability arises correspondingly. It can also be seen that longer synchronization sequence shows better performance in acquisition. This is because long synchronization sequence has better correlation performance and anti-noise property.

41.3.2 Ranging Jitter

For the proposed method, the range measurement is realized by matching the SN being through two-way transmission with the local SN. Thus, the synchronization process will cause a ranging jitter because the signal synchronization influenced by the noise will lead to a time synchronization error.

Fig. 41.4 Acquisition probability under different SNR



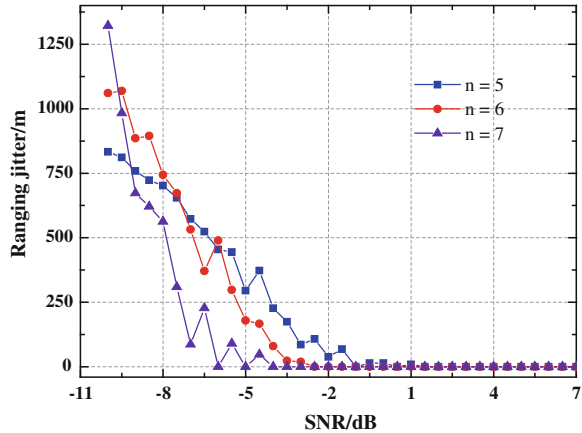
Based on the definitions of Eqs. (41.6) and (41.7), it is clear that the correlation error is caused by the noise from the space transmission. For the synchronization process, the aim is to find k with the maximum value of $R_{corr}(k)$, i.e.,

$$\begin{aligned} \kappa &= \max_k \{R_{corr}(k)\} \\ &= \max_k \{\rho_{SS}(k) + \zeta_{SS}(k)\} \end{aligned} \tag{41.13}$$

κ can be seen as a random variable. The variance of κT , $\sigma_{\kappa T}^2$, contributes to the ranging jitter. The ranging jitter, denoted by σ_R , can be written as

$$\sigma_R = \frac{c\sigma_{\kappa T}}{2} \tag{41.14}$$

Because the analytical form for Eq. (41.13) cannot be obtained, the Monte Carlo simulation method is adopted to analyze the variance of κ and to discuss the impact of SNR on the ranging jitter. For each SNR, 1000 correlation trials expressed by Eq. (41.13) are conducted and the variance of κT is calculated. Then, the ranging jitter is calculated based on Eq. (41.14) and presented in Fig. 41.5. It is shown in the figure that when the SNR is low, the synchronization error causes a large ranging jitter and that the ranging jitter decreases with the increasing SNR. This shows the influence of the noise on the range measurement. It should be noted that at low SNR, the ranging jitter with respect to the high order m-sequence is larger than that of the lower ones. The reason for this is at lower SNR, the noise contributes much to the correlation process. When the synchronization sequences are correlated, the noise has a larger effect on the longer synchronization sequences than to the shorter ones.

Fig. 41.5 Ranging jitter

41.4 Conclusions

The design and analysis of a new simultaneous ranging and communication method has been presented in this paper. Simultaneous ranging and communication could improve the link efficiency under the fickle deep space conditions and was helpful for the system miniaturization and integration. For the proposed method, by transmitting the ranging information together with the communication data with a specific signal symbol structure, the simultaneous operation was realized. The principle of the proposed method was introduced and the performance of the method was analyzed in detail utilizing both mathematical analysis and numerical simulation. From the mathematical analysis and the numerical simulations, we knew that the correlation results and the acquisition probability were influenced by the noise and that better correlation result and higher acquisition probability could be obtained with higher SNR. Besides, the ranging jitter was related to the correlation property of synchronization sequence of the signal symbol and was also influenced by the level of SNR. Based on the detailed analysis, we could conclude that the simultaneous ranging and communication method based on XCOM was feasible and that it could be utilized in deep space application. With the increasing demand for the range measurement and large volume data transmission, simultaneous ranging and communication would find its application in future deep space exploration.

References

1. Cesarone RJ, Abraham DS, Deutsch LJ (2007) Prospects for a next-generation deep-space network. Proc IEEE 95:1902–1915. doi:10.1109/JPROC.2007.905043
2. Park C-W (2001) Precise relative navigation using augmented CDGPS, Citeseer

3. Dewberry B, Petroff A (2013) SwarmNet: a distributed navigation network using ultra wideband ranging and communications. In: Proceedings of microwave conference (EuMC), 2013 European, pp 6–10 Oct 2013
4. Wang S, Zhang Er Y (2002) Inter-satellite radio links and spread-spectrum ranging for satellite formation flying. In: Proceedings of microwave and millimeter wave technology, 2002 proceedings ICMMT 2002, 3rd international conference on, 17–19 Aug 2002. doi: [10.1109/ICMMT.2002.1187678](https://doi.org/10.1109/ICMMT.2002.1187678)
5. Gao P, You Z (2007) Characters of laser ranging on inter-satellite relative position measurement. In: 3rd international conference on recent advances in space technologies RAST '0, pp 14–16. doi: [10.1109/RAST.2007.4284066](https://doi.org/10.1109/RAST.2007.4284066)
6. Yi LF (2007) Research on the estimation model of TDOA in the multiplex system. Harbin Engineering University, Harbin
7. Deng K, Jiang DG, Yao ZS, Yang HY (2008) A novel technology combined with free space optics communication and laser ranging. In: Proceedings of optical transmission, switching, and subsystems VI, 27 Oct 2008, USA. SPIE—The International Society for Optical Engineering
8. Yu GS (2009) Impulse communication and ranging complex system technology research. Graduate School of Chinese Academy of Sciences, Beijing
9. Andrews K, Hamkins J, Shambayati S, Vilnrotter V (2010) Telemetry-based ranging. In: Proceedings of 2010 IEEE aerospace conference, pp 6–13 March 2010. doi: [10.1109/AERO.2010.5446926](https://doi.org/10.1109/AERO.2010.5446926)
10. Sun X, Skillman DR, Hoffman ED, Mao D, McGarry JF, Neumann GA, McIntire L, Zellar RS, Davidson FM, Fong WH (2013) Simultaneous laser ranging and communication from an Earth-based satellite laser ranging station to the lunar reconnaissance orbiter in lunar orbit. In: Proceedings of SPIE LASE, 2013. International society for optics and photonics
11. NASA (2012) NASA Space technology roadmaps and priorities: restoring NASA's technological edge and paving the way for a new era in space, USA, p 376
12. Goldstein D, Goldstein DH (2011) Polarized light, revised and expanded. CRC Press, New York
13. James DF (1994) Change of polarization of light beams on propagation in free space. J Opt Soc Am A 11:1641–1643. doi: [10.1364/JOSAA.11.001641](https://doi.org/10.1364/JOSAA.11.001641)
14. Zhao XH, Yao Y, Sun YX, Xu XC, Tian JJ, Liu C (2010) Condition of keeping polarization property unchanged in the circle polarization shift keying system. IEEE/OSA J Opt Commun Networking 2:570–575. doi: [10.1364/JOCN.2.000570](https://doi.org/10.1364/JOCN.2.000570)
15. ul Abidin Z, Pei X, Amin M, Fusco V (2012) Circular polarization modulation for digital communication systems. In: 8th international symposium on communication systems, networks, and digital signal processing, pp 18–20. doi: [10.1109/CSNDSP.2012.6292758](https://doi.org/10.1109/CSNDSP.2012.6292758)
16. Liu C, Yao Y, Yang YF, Yuan YJ, Zhao YF, Yu BS (2013) Performance of free-space optical communication systems using circle polarization shift keying with spatial diversity receivers. Chin Opt Lett 11:20101. doi: [10.3788/COL201311.S20101](https://doi.org/10.3788/COL201311.S20101)

Chapter 42

Cross Correlation Mitigation Algorithm for Indoor Positioning Receiver

Yang Hu, Zhongliang Deng, Jichao Jiao, Jie Xia,
Yuezhou Hu and Zhuang Yuan

Abstract In recent years, the location based service has been gradually integrated into people's daily life. Many fields such as the public security, emergency rescue, warehousing logistics, and IOT (The Internet of things) applications present extensive application demand on indoor location. This paper discusses the cross correlation interference in the TC-OFDM (time divided CDMA-OFDM superimposed signal system) receiver which is based on ground mobile radio network. In the TC-OFDM system, because of the complexity of the indoor environment, multiple access interference is terrible, which greatly affects the performance of the receiver for signal acquisition and tracking. This paper studies the characteristics of the multiple access interference in TC-OFDM system in indoor environment. The cross correlation value between the strong and weak signal can be estimated by putting the characteristic information of strong signal into the weak signal tracking channel. While this information can be extracted from the strong signal's tracking channel. Therefore, strong signal's effect on the weak signal's autocorrelation function can be reduced. So that the tracking loop can get more accurate phase value for tracking. This method can improve the TC-OFDM receiver's ability of inhibiting cross correlation by 10–13 dB, and effectively enhance the capacity for sustained and stable tracking of weak signal. In addition, in the snapshot replay mode of the weak signal which cannot be tracked continuously, this method helps to estimate the correlation peak more accurate.

Keywords Indoor location · Multiple access interference · Tracking · Strong and weak signal

Y. Hu (✉) · Z. Deng · J. Jiao · J. Xia · Y. Hu · Z. Yuan
School of Electronic Engineering, Beijing University of Posts and Telecommunications,
Beijing, China
e-mail: huyang90@bupt.edu.cn

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015
Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_42

485

42.1 Introduction

Outdoor positioning technology based on satellite is already quite mature, and the precision of GPS, GLONASS, Galileo and Beidou satellite positioning systems have met the needs of daily use. At the same time, in many fields such as the public security, emergency rescue, warehousing logistics, IOT (The Internet of things) applications presents extensive application demand on indoor location. The TC-OFDM (time divided CDMA-OFDM superimposed signal system) system discussed in this paper is designed for high-precision indoor position services. Recently, most of the positioning systems are based on CDMA (code division multiple access) which are the same with TC-OFDM system. In the TC-OFDM system, the CDMA signal is superimposed on a mobile broadcast or communication signal, without affecting the original system. The CDMA system has strong anti-interference ability, high utilization rate of frequency spectrum, and excellent confidentiality. But due to the different codes are not strictly orthogonal, and the effect of channel noise, there are intersymbol interference between different codes. In the TC-OFDM system, on the one hand the distance between the user and base stations are uncertain; on the other hand, because of the complexity of the indoor environment, and the influence of multipath transmission; the signal power fluctuates with user's change of spatial and time. It is easy to appear that one signal is much stronger than the other one. These two aspects will produce seriously cross correlation interference, so that the correlation peak of the weak signal will be submerged in the cross-correlation peaks. Thereby, the weak signal is unable to be captured or tracked. It is important to reduce or eliminate the cross correlation interference for TC-OFDM receiver, which can greatly improve the sensitivity of the TC-OFDM receiver.

Current methods of suppressing intersymbol interference in CDMA system are multiple peak detection method, subspace projection method and series/parallel interference cancellation method [1]. Multiple peak detection method detects intersymbol interference through the different distribution characteristics of the cross-correlation peaks and the correlation peak in the signal acquisition and tracking process. The basic principle of subspace projection method is to construct a new code which is completely orthogonal with the strong signal's GOLD code and and correlative with weak signal's GOLD code. The receiver can use this new code to calculate the correlation value, so that the influence of strong signal can be eliminated. The main idea of the serial/parallel interference cancellation method is to reconstruct the strong signal through the information from strong signal tracking channel. The cross correlation value can be estimated by the correlation operation between the new strong signal and the weak signal GOLD code. This paper mainly discusses the realization of the serial interference cancellation method in TC-OFDM receiver.

42.2 Cross Correlation Mitigation Algorithm

The CDMA signals received by TC-OFDM receiver are sent from N base stations around. The signal has explored channel attenuation and superimposed the corresponding channel noise. In order to facilitate the analysis, we assume that the receiver antenna received mixed signal from two base stations, and the signal strength is not the same, the digital intermediate frequency signal can be expressed as:

$$\begin{aligned} r_{IF}(t) &= S_{SIF}(t) + S_{WIF}(t) \\ &= A_S D_S G_S(t - \tau_S) \exp\{j[2\pi(f_{IF} + f_{dS})(t - \tau_S) + \theta_{IFS}]\} \\ &\quad + A_W D_W G_W(t - \tau_W) \exp\{j[2\pi(f_{IF} + f_{dW})(t - \tau_W) + \theta_{IFW}]\} \end{aligned} \quad (42.1)$$

where A_S and A_W are amplitude of the intermediate frequency signal. D is gold code of the corresponding base station. τ is signal propagation delay. f_{IF} is residual frequency of the intermediate frequency signal. f_d is Doppler shift and θ_{IF} is original Carrier phase. The correlation value can be obtained by formula (42.2) [2]:

$$\begin{aligned} V(n) &= V_{WW}(n) + V_{SW} \\ &\approx A_W D_W R_{WW}(\tau_W^{(e)}) \sin c(f_W^{(e)} T_{coh}) \cos(\phi_W^{(e)}) \\ &\quad + A_S D_S R_{SW}(\tau_W^{(e)} + \tau_{SW}) \sin c((f_W^{(e)} + f_{SW}^{(d)}) T_{coh}) \cos(\phi_{SW} + \phi_W^{(e)}) \end{aligned} \quad (42.2)$$

where $\tau^{(e)}$ is code phase error of the tracking loop, $f^{(e)}$ is frequency discrimination error, $\phi^{(e)}$ is phase discrimination error, T_{coh} is the coherent integration time. $f_{SW}^{(d)}$ is Doppler shift difference between the weak and strong signal, τ_{SW} is code phase difference between the weak and strong signal, ϕ_{SW} is the carrier phase difference. R_{WW} and R_{SW} are the auto correlation function and cross correlation function. The first part in the right side of the equation is the result of weak signal's auto correlation which can be put into the tracking loop for the next step of frequency discrimination and phase discrimination. The second part appears as noise for weak signal. Along with the increase of $A_S R_{SW}/A_W R_{WW}$, the weak signal's frequency discrimination and phase discrimination result will be getting more and more inaccurate, which may lead to weak signal's loss of lock. On the other hand, tracking loop may track the cross-correlation peak instead of the weak signal's auto correlation peak, which will cause large positioning error [3].

Because of the strong signal can be stabilized tracking, strong signal's code phase, carrier frequency, carrier phase, data bits and other information related can be obtained from the strong signal tracking channel. According to these informations, the strong signal can be reconstructed. By calculating the correlation between the reconstructed strong signal and weak signal's gold code, the second part in the right side of Eq. (42.2) can be estimated. Then we can get more accurate weak signal's autocorrelation value. Strong signal's reconstruction can be expressed as:

$$S_{SIF}(t) = A_S D_S G_S(t - \tau_S) \exp\{j[2\pi(f_{IF} + f_{dS})(t - \tau_S) + \theta_{IFS}]\} \tag{42.3}$$

where $D_S, G_S, \tau_S, f_{IF} + f_{dS}, \theta_{IFS}$ can be obtained from the strong signal tracking channel, while the amplitude of the strong signal can be estimated by the formula below, and N_{coh} indicates the length of the gold code.

$$V_S(\max) = A_S N_{coh} = \sqrt{I_S^2 + Q_S^2} \tag{42.4}$$

$$\text{And } A_S = V_S(\max)/N_{coh} = \sqrt{I_S^2 + Q_S^2}/N_{coh} \tag{42.5}$$

42.3 Simulation Test

The algorithm is verified preliminary by simulation in MATLAB. We analyze the 511 gold code used in TC-OFDM signal and have obtained some simulation results. Due to the interference of the strong signal, weak signal's auto correlation peak is no longer ideal symmetric shape, thus giving a phase discrimination error to the tracking loop. In addition, the difference between strong and weak signal code phase is also can cause different influence on weak signal correlation peak. Figure 42.1 shows the effects of strong signal's interference on weak signal correlation peak. Where the solid line shows the strong signal's power is 30 dB higher than the weak signal's; the long dashed line shows the strong signal's code phase is 31 code chips ahead of the weak signal's; the short dashed line shows the weak signal's autocorrelation function.

By using the cross correlation mitigation algorithm, we can correct the distorted correlation peak of weak signal. The correction result is shown in Fig. 42.2:

When faced with the weak signal which cannot be tracked continuously, the TC-OFDM receiver usually get the code phase information of current signal by

Fig. 42.1 Curve: correlation peak before reducing cross correlation

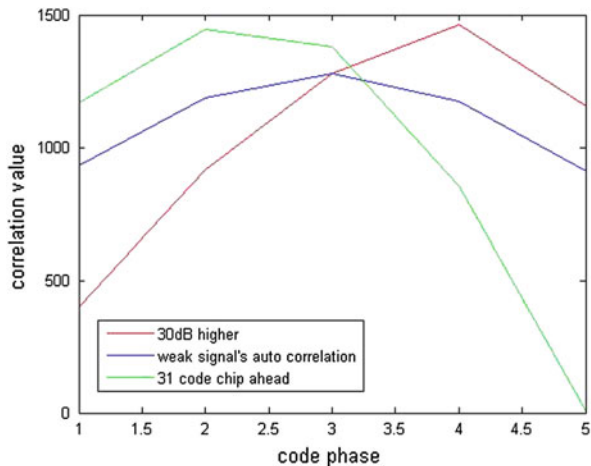
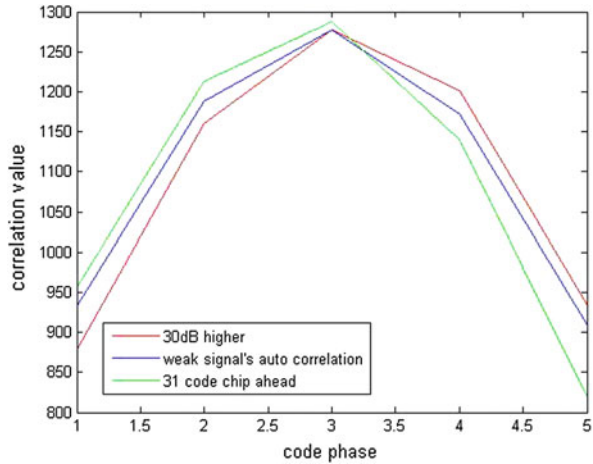


Fig. 42.2 Curve: correlation peak after reducing cross correlation



using snapshot-replay algorithm [4]. The main process of snapshot algorithm is shown below: firstly, a section of the weak signal will be saved in a ram; secondly, numbers of correlation values will be calculated near the correlation peak; thirdly, according to the characteristics of correlation peak curve, the code phase can be estimated accurately. The prerequisite for this algorithm is that the correlation peak curve is not seriously distorted. Therefore, according to the analysis, the cross correlation mitigation algorithm can also improve the code phase estimation accuracy in the snapshot mode.

When a signal's power is much higher than the other one, the weak signal's auto correlation peak may be submerged by the cross-correlation peaks. So that tracking loop may lock the cross correlation peaks as weak signal's auto correlation peak and cause incorrect tracking. According to the cross correlation mitigation algorithm, the cross correlation peak can also be reduced, so that the tracking loop can trace the peak correctly. Figure 42.3 indicates the correlation value of the weak signal, in which strong signal's power is 30 dB higher than the weak signal's. The weak signal's auto correlation peak has been completely invisible.

Compared with Figs. 42.3 and 42.4 clearly shows the the cross correlation has been reduced from correlation value, and it is much easier for the tracking loop to trace the weak signal's auto correlation peak. Through the simulation results can explain the strong signal elimination algorithm is very feasible in improving the TC-OFDM receiver's tracking ability of weak signal.

42.4 Realization of the Algorithm

The algorithm is achieved in the TC-OFDM receiver, and in order to guarantee the resources are enough, an EP3C120 series FPGA of ALTERA is used in the development. In addition, an ARM9 and a RF chip are used for the control of the

Fig. 42.3 Curve: correlation value before reducing cross correlation

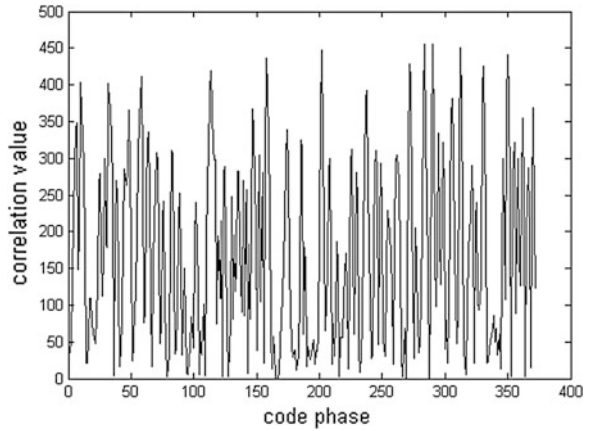


Fig. 42.4 Curve: correlation value after reducing cross correlation

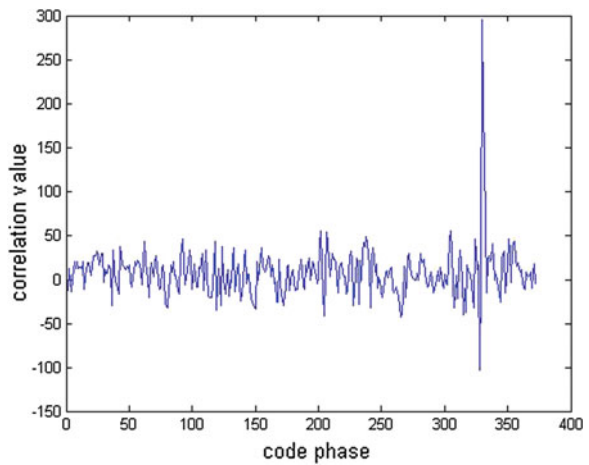
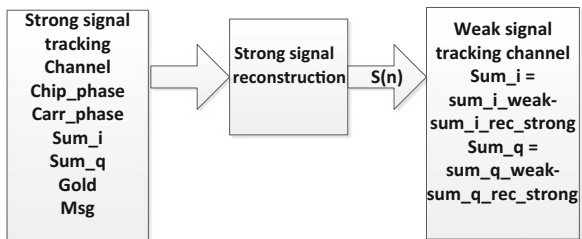
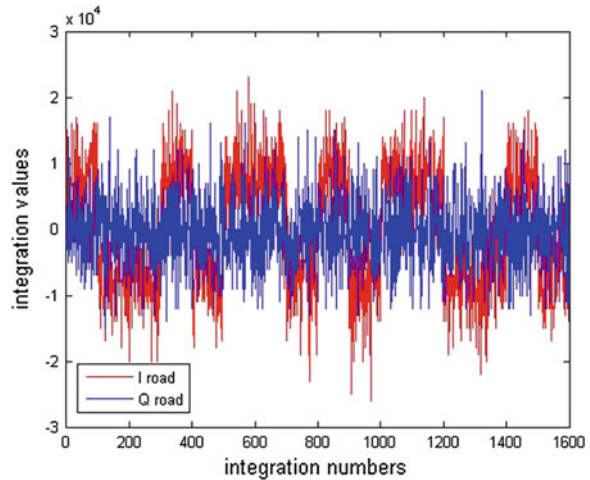


Fig. 42.5 Algorithm block diagram



baseband and receive the signal of TC-OFDM. The core algorithm block diagram is shown in Fig. 42.5. Strong signal reconstruction module acquires the code phase, carrier phase, coherent integral value, gold code and message information of the strong signal from tracking channel and put the reconstructed signal into weak

Fig. 42.6 Curve: I and Q correlation value before reducing cross correlation



signal tracking channel. Tracking channel consists of code tracking loop, carrier tracking loop, integrators and so on. The weak signal tracking channel has to add corresponding integrators to calculate the cross correlation value between the weak signal and the reconstructed signal. The difference of the new added integrators and the original integrators is modified result of weak signal's auto correlation value.

In order to verify that the cross correlation mitigation algorithm is validated in practice, we used two signal sources to generate TC-OFDM signals and connect them to the receiver for tracking, and compare the receiver's performance before and after the algorithm by gradually adjusting the power of the two signal sources.

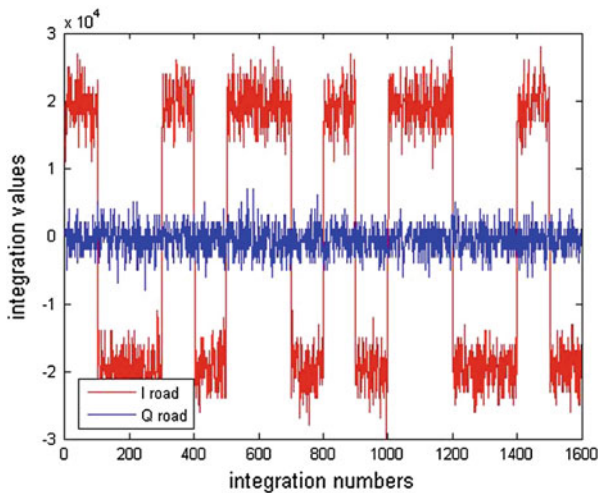


Fig. 42.7 Curve: I and Q correlation value after reducing cross correlation

Figure 42.6 shows the I road and Q road coherent integral value of weak signal without using the algorithm above, where the power of two signals differ by 30 dB. In this case, the jitter of coherent integral value is very large and the separation degree of IQ path is small. After applying the above algorithm, the tracking performance is improved remarkably and the IQ road is completely separated, so that the weak signal can be normal tracked (Fig. 42.7).

42.5 Conclusions

This paper discusses the cross correlation interference on TC-OFDM receiver's tracking capacity of weak signal and the realization of serial interference cancellation method in TC-OFDM receiver. At the same time, the algorithm is implemented on the FPGA. It shows that this method can improve the TC-OFDM receiver's ability of inhibiting cross correlation by 10–13 dB, and effectively enhance the capacity for sustained and stable tracking of weak signal. This paper mainly discusses the two weak and strong signals, but according to Fig. 42.5, the strong signal reconstruction module can be improved to reconstruct several strong signals at the same time and put them into the weak signal tracking channel to reduce the interference of multi-strong signals.

Acknowledgments Foundation project: The National High Technology Research and Development Program (“863” Program) of China (No. 2014AA123103). The National Natural Science Foundation of China (No. 61372110).

References

1. Hao J (2009) Implementation of cross-correlation interference mitigation algorithm for GNSS receiver. *Microelectronics* 39(6):857–860
2. Xie G (2009) Principles of GPS and receiver design. Publishing House of Electronics Industry, Beijing
3. He W, Xu J, Ye T (2011) Detection of GNSS weak signal in cross correlation interference. *Acta Electronica Sinica* 39(2):471–475
4. Hu Y, Deng Z, Yuan Z, Hu Y (2014) A snapshot-replay method of pseudo-code ranging under frequency limited fading channel. *Adv Mater Res* 1016–1022

Chapter 43

An Optimal Data Fusion Algorithm Based on the Triple Integration of PPP-GNSS, INS and Terrestrial Ranging System

Wei Jiang, Yong Li and Chris Rizos

Abstract This paper describes the integration of *Locata*, GNSS and INS technologies within a loosely-coupled triple integration algorithm. The conventional methods for multi-sensor integration can be classified as either centralised filtering or decentralised filtering. Centralised Kalman filtering (CKF) provides globally optimal state estimation by directly combining measurement data. However CKF system has some disadvantages such as a comparatively large computational burden and poor fault detection and isolation ability. Decentralised Kalman filtering (DKF) addresses such defects while aiming to achieve the same accuracy as a centralised filter. On the other hand global optimal filtering (GOF) can achieve a higher accuracy than the traditional CKF because it utilises more information resources than the CKF. In the information space, the information resources that can be used for estimation include the measurements, the local predictions, and the global predictions. In order to evaluate the system performance, a field experiment was conducted on a vehicle with different kinds of maneuvers, including circular motion and accelerated motion. The results indicate that: (1) GOF-based PPP-GNSS/*Locata*/INS integration system can provide better positioning accuracy compared with CKF and federated Kalman filtering; (2) covariance analysis shows that the GOF improves the system estimation covariance; and (3) a comparison of GOF with local filters confirms the superiority of a GOF-based triple integration system.

Keywords GNSS · *Locata* · INS · Multisensor system · Data fusion

W. Jiang (✉) · Y. Li · C. Rizos
Surveying and Geospatial Engineering, School of Civil and Environmental Engineering,
University of New South Wales, Sydney, Australia
e-mail: w.jiang@student.unsw.edu.au

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015*
Proceedings: Volume III, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_43

493

43.1 Introduction

Although Global Navigation Satellite System (GNSS) technology has been widely used to satisfy positioning and navigation requirements in many application fields, the major disadvantage of GNSS still remains: signal blockage due to obstructions, and power attenuation of the signals when operated indoors. The combination of GNSS with other sensors, such as a self-contained inertial navigation system (INS), provides an ideal position and attitude determination solution which can not only mitigate the weakness of GNSS, but also bound the INS error that otherwise would grow with time when the INS operates alone. However, the navigation accuracy provided by GNSS/INS is strongly dependent on the quality and geometry of the GNSS observations, the INS technology used, and the integration model applied.

In general, with limited GNSS availability, the navigation accuracy can only be maintained at the metre-level for no more than a few minutes even if a reasonably high quality INS is used. Many advanced fusion algorithms have been investigated to mitigate this problem however during longer GNSS outages navigation solutions may not meet the accuracy requirement of the application [1, 2].

It is possible to incorporate an auxiliary sensor system to augment GNSS/INS navigation. “*Locata*” is a terrestrial, radio-frequency based technology which uses a local “constellation” of signal transmitters to provide continuously time synchronised ranging signals in difficult GNSS environments, such as where there is poor satellite geometry: signal blockages in tunnels, urban canyons, and deep open-cut mines etc.

In this paper, in order to satisfy accuracy and reliability performance parameters, GNSS, based on the precise point positioning (PPP) data processing methodology, and *Locata* are integrated with an INS to provide robust global navigation solutions.

The fundamental challenge of integrated navigation systems is information fusion. The commonly used method is the centralised Kalman filtering (CKF) [3, 4] which ensures globally optimal state estimation by processing all sensor measurements at a central site. However, the CKF can result in a large computational burden and large data memory demands [5]. From the point of view of fault tolerance, the centralised complementary Kalman filter is not convenient to adaptively detect and isolate sensor faults, thus CKF-based navigation systems have limited reliability [6, 7].

Another approach is based on decentralised Kalman filtering (DKF), where the information from local estimators can yield global optimal or sub-optimal state estimation according to certain information fusion criteria. The advantage of this approach is that the communications load and processing demands are decreased, and the input data rates could be increased due to the use of parallel structures. Moreover, decentralised filtering leads to easy fault detection and isolation. Recently developed multi-sensor data fusion via the information space approach has shown that the global optimal filtering (GOF) can achieve a higher accuracy than the traditional CKF [8]. In order to take advantage of decentralised filtering and to improve the estimation accuracy, the GOF approach is applied in this paper.

According to the information space concept, the optimal data fusion can be expressed as the projection of the state vector to all kinds of information spaces. Thus the more information resources are utilised, the better the system performance. In the information space, the measurements, the local predictions, and the global predictions are resources for the estimation process. The GOF utilises all the information resources, and hence differs from the traditional CKF and the Federated Kalman filter (FKF) that utilise only a portion of the available information resources.

In this paper GOF is applied to implement a PPP-GNSS/*Locata*/INS integrated navigation algorithm. This approach overcomes the disadvantages of conventional decentralised fusion by utilising all of the information resources. A decentralised estimation fusion method is established for individual integrations of PPP-GNSS and *Locata* with INS to independently obtain the local predictions and local estimation. The local and global information is further fused to generate the global optimal state estimation of the triple-integrated navigation system. Experimental results are presented to demonstrate the performance of the proposed method.

43.2 System Model and Conventional Fusion Algorithms

INS is used to establish the inertial navigation model. The system dynamic equation is described by:

$$\dot{X}(t) = F(t)X(t) + W(t) \quad (43.1)$$

where $X(t)$ is the state vector of the system, $F(t)$ is the dynamic matrix of the system, $W(t)$ is the system noise. $X(t)$ is defined as:

$$X(t) = [\psi_N, \psi_E, \psi_D, \delta r_N, \delta r_E, \delta r_D, \delta v_N, \delta v_E, \delta v_D, \varepsilon_x, \varepsilon_y, \varepsilon_z, \nabla_x, \nabla_y, \nabla_z, \delta \eta_{Lx}, \delta \eta_{Ly}, \delta \eta_{Lz}, \delta \eta_{Gx}, \delta \eta_{Gy}, \delta \eta_{Gz}]^T \quad (43.2)$$

where (ψ_N, ψ_E, ψ_D) is the attitude angle error, $(\delta r_N, \delta r_E, \delta r_D)$ is the position error, $(\delta v_N, \delta v_E, \delta v_D)$ is the velocity error, $(\varepsilon_x, \varepsilon_y, \varepsilon_z)$ is the gyroscope's constant bias, $(\nabla_x, \nabla_y, \nabla_z)$ is the accelerator's bias, $(\delta \eta_{Lx}, \delta \eta_{Ly}, \delta \eta_{Lz})$ is the *Locata* lever arm components with respect to the inertial measurement unit (IMU) centre, $(\delta \eta_{Gx}, \delta \eta_{Gy}, \delta \eta_{Gz})$ is the GNSS lever arm estimation with respect to the IMU centre.

The observation information of the integrated PPP-GNSS/*Locata*/INS system includes the position and velocity (PV) information from *Locata* and PPP-GNSS, and PV information from the INS. For centralised filtering, the measurements are described by the stacked observation vector:

$$Z(t) = [Z_L, Z_G]^T = [r_L - r_{INS}, v_L - v_{INS}, r_G - r_{INS}, v_G - v_{INS}]^T \quad (43.3)$$

The system observation equation can be written as:

$$Z(t) = H(t)X(t) + V(t) \quad (43.4)$$

where $H(t)$ and $V(t)$ are the observation matrix and observation noise of the integrated navigation system respectively.

Applying the discrete process, the system Eqs. (43.1) and (43.3) become:

$$x(k) = F(k)x(k-1) + \omega(k) \quad (43.5)$$

$$z(k) = H(k)x(k) + v(k) \quad (43.6)$$

It is assumed that $\omega(k)$, $v(k)$ are zero-mean white sequences uncorrelated with each other, and $E[\omega(k)\omega(k)^T] = Q(k)$, $E[v(k)v(k)^T] = R(k)$. The initial state $x(0)$ is a zero-mean Gaussian random vector, $x(0) \sim N(\bar{x}(0), P(0))$, and is independent of $\omega(k)$ and $v(k)$.

The Kalman filter algorithm is composed of time and measurement updates. The prediction of the state $\bar{x}(k)$, and its covariance matrix $\vec{P}(k)$, are obtained from the time-updating step:

$$\bar{x}(k) = F(k)\hat{x}(k-1) \quad (43.7)$$

$$\vec{P}(k) = F(k)\hat{P}(k-1)F^T(k) + Q(k) \quad (43.8)$$

When the measurements are available the state vector $\hat{x}(k)$ is updated as:

$$K(k) = \vec{P}(k)H^T(k)[H(k)\vec{P}(k)H^T(k) + R(k)]^{-1} \quad (43.9)$$

$$\hat{P}(k) = [I - K(k)H(k)]\vec{P}(k) \quad (43.10)$$

$$\hat{x}(k) = \bar{x}(k) + K(k)[z(k) - H(k)\bar{x}(k)] \quad (43.11)$$

where $K(k)$ is the Kalman gain, and $\hat{P}(k)$ is the covariance matrix of the state.

For a decentralised system, the sub-systems for the *Locata* and GNSS sensors independently observe the output of the INS. The two local filters estimate the state of the INS using the observed data. Since the two local filters share the INS navigation model, the dynamic equations of the two local systems are the same. The model of the i th ($i = 1, 2$) local filter is described by the following equations:

$$x(k) = F(k)x(k-1) + \omega(k) \quad (43.12)$$

$$z_i(k) = H_i(k)x(k) + v_i(k) \quad (43.13)$$

where $z_i(k)$ is the output of the i th subsystem, and $v_i(k)$ is the measurement noise of the local system, which is assumed to be white noise. The covariance of the measurement noise in the local filter is defined by the general accuracy of the local *Locata* and PPP-GNSS solutions, which is set in this evaluation as:

$$\begin{aligned} R_{Locata} &= \text{diag}((0.2)^2 \quad (0.3)^2 \quad (0.1)^2 \quad (0.1)^2 \quad (0.1)^2 \quad (0.1)^2) \\ R_{PPP-GNSS} &= \text{diag}((0.2)^2 \quad (0.1)^2 \quad (0.1)^2 \quad (0.1)^2 \quad (0.1)^2 \quad (0.1)^2) \end{aligned}$$

Each local filter generates the optimal estimation of the local state by using the Kalman filter formulas (43.7)–(43.11).

After completing the computations of the two decentralised local filters, two local optimal state estimations $\hat{x}_1(k)$ and $\hat{x}_2(k)$ can be obtained and further fused using the global filter.

43.3 GOF for PPP-GNSS/*Locata*/INS Integrated Navigation

The multi-sensor optimal estimation is resolved in two approaches, in terms of the random vector space (RVS) approach and the information space approach [8]. In the RVS framework, the sources of information have been identified and used as the bases of the space. The estimation is mathematically described as the procedure for finding the projection of the state vector on the bases of the space. The fusion algorithm therefore describes how the global state estimate is combined by the projections and associated bases.

Differing from the RVS approach, from the point of view of the information space, optimal fusion is implemented by a series of transformations between the information spaces. The transformations map the source information vectors from the measurement information spaces to the estimate information space to produce the fused information vector. The information space approach provides a means by which the accuracies of different algorithms can be compared on a theoretical basis.

As the RVS approach is similar to the widely applied Kalman filtering forms, it is used below to discuss the nature of GOF.

Assume a linear discrete system \hat{x}_i ($i = 1, 2, \dots, n$) is the unbiased estimators of the stochastic vector x . If \hat{x}_i are orthogonal vectors, the estimate of x can be expressed as the sum of the projection of \hat{x}_i in a random vector space (RVS) Θ [8]:

$$\Theta = \left\{ \hat{x} | \hat{x} = \sum_{i=1}^n C_i \hat{x}_i \right\} \quad (43.14)$$

The optimal state estimation and its covariance can be written as:

$$\hat{x} = \hat{P} \left(\sum_{i=1}^n S_i^T \hat{P}_i^{-1} \hat{x}_i \right) \quad (43.15)$$

$$\hat{P} = \left(\sum_{i=1}^n S_i^T \hat{P}_i^{-1} S_i \right)^{-1} \quad (43.16)$$

where S_i is the mapping matrix between \hat{x} and \hat{x}_i , and $\hat{x}_i = S_i \hat{x}$; \hat{P}_i is the corresponding posteriori estimate error covariance of \hat{x}_i , which is calculated epoch by epoch to the convergence status. The initial setting of P_i does not affect the estimate convergence, and it can be set according to the approximate accuracy of initial $x_i(0)$. It can be seen that the more information that is used, represented here as \hat{x}_i , the smaller the trace of \hat{P} that is obtained. Hence the system could have improved performance.

For a better understanding, recall the Kalman filter's measurement updating Eq. (43.11). This prediction form indicates that the state prediction is corrected when the measurement is made available. It can be rewritten in the following fusion form:

$$\hat{x}(k) = [I - K(k)H(k)]\bar{x}(k) + K(k)z(k) \quad (43.17)$$

It also can be regarded as the random vector space below [8]:

$$\Theta = \{\hat{x} | \hat{x} = C_1 \bar{x} + C_2(k)z\} \quad (43.18)$$

where $C_1 = \hat{P}\bar{P}^{-1}$, and $C_2 = \hat{P}H^T R^{-1}$

The estimation \hat{x} can be seen as a point or vector in the space Θ , which is expressed as the projections C_1 and C_2 on the bases of \bar{x} and z .

Similarly, the RVS of the CKF can be expanded by the measurements z_1, z_2 and the global prediction \bar{x} :

$$\Theta_{CKF} = \{\hat{x} | \hat{x} = C\bar{x} + C_1(k)z_1 + C_2(k)z_2\} \quad (43.19)$$

where $C = \hat{P}\bar{P}^{-1}$, $C_1 = \hat{P}H_1^T R_1^{-1}$, $C_2 = \hat{P}H_2^T R_2^{-1}$

An optimal multi-sensor estimation in turn requires information from the bases and the determination of the associated projections. The following random vectors in a multi-sensor system can be used as the bases [8]:

Measurements: $z_i(k)$, $i = 1, 2$

Local predictions: $\bar{x}_i(k)$, $i = 1, 2$

Global predictions: $\bar{x}(k)$

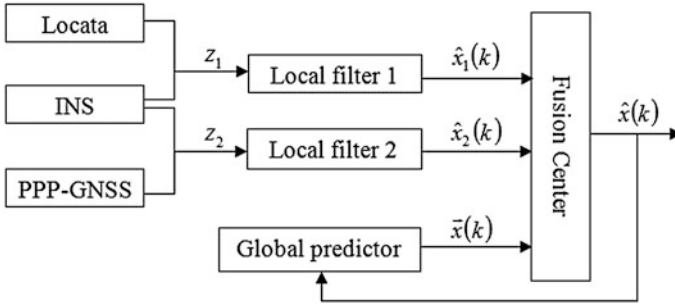


Fig. 43.1 System configuration of GOF

Utilising different sources as the bases of RVS leads to different fusion algorithms. From Eq. (43.19), CKF uses global prediction and all local measurement information, FKF uses local estimates and global prediction. However, in order to ensure the same accuracy of the CKF, the local predictions of the FKF are computed from the global prediction using the so-called information-sharing principle [9, 10].

To further improve the accuracy of decentralised filtering, the GOF uses all the information sources, including the measurements, the local and global predictions to achieve global optimality, which can be represented as:

$$\Theta = \left\{ \hat{x}(k) | \hat{x}(k) = \sum_{i=1}^2 a_i(k) \tilde{x}_i(k) + \sum_{i=1}^2 b_i(k) z_i(k) + c(k) \tilde{x}(k) \right\} \quad (43.20)$$

The measurements and the local predictions support local Kalman filtering fusion, thus they can be replaced by the local state estimates. The GOF configuration is shown in Fig. 43.1. The RVS of GOF in Eq. (43.20) can be rewritten as:

$$\Theta = \{ \hat{x}(k) | \hat{x}(k) = c(k) \tilde{x}(k) + d_1(k) \hat{x}_1(k) + d_2(k) \hat{x}_2(k) \} \quad (43.21)$$

where a_i, b_i, c, d_1 and d_2 represent the projections of the global estimate vector \hat{x} on the bases.

In Fig. 43.1, one can see both *Locata* and PPP-GNSS outputs are combined with INS solution in the local filters, thus the local estimates \hat{x}_1 and \hat{x}_2 are correlated with previous feedback-INS solution, represents the previous state estimate here. On the other hand, the current global prediction is also computed from the previous global estimate. That means the two local estimates and the current global prediction are oblique to each other, not orthogonal. Therefore these three vectors need to be orthogonalised [8]. The reconstructed orthogonal global prediction and local estimates are represented as \tilde{x} and \tilde{x}_i :

$$\tilde{x}_1(k) = \hat{x}_1(k), \tilde{x}_2(k) = \hat{x}_2(k), \tilde{x}(k) = \alpha_1 \hat{x}_1(k) + \alpha_2 \hat{x}_2(k) + \alpha \bar{x}(k) \quad (43.22)$$

The corresponding covariance of reconstructed global prediction is:

$$\tilde{P}(k) = \alpha \bar{P}(k) \alpha^T + \sum_{i=1}^2 \alpha_i \hat{P}_i(k) \alpha_i^T + \sum_{i=1}^2 \alpha \tilde{P}_i(k) \alpha_i^T + \sum_{i=1}^2 \alpha_i \tilde{P}_i^T(k) \alpha^T \quad (43.23)$$

where

$$\tilde{P}_i = \text{cov}(\bar{x}, \tilde{x}_i) = F(k) \hat{P}(k-1) F^T(k) \bar{P}_i^{-1}(k) \hat{P}_i(k) \quad (43.24)$$

As global prediction \bar{x} and local estimates \hat{x}_i are all unbiased estimates, the orthogonalised global prediction \tilde{x} is also unbiased, hence $E(\tilde{x} - x) = 0$, from which one can derive:

$$\alpha_1 + \alpha_2 + \alpha = I \quad (43.25)$$

Since the reconstructed global prediction \tilde{x} and local estimation \tilde{x}_1, \tilde{x}_2 are orthogonal to each other. Thus $\text{cov}(\tilde{x}, \tilde{x}_1) = 0, \text{cov}(\tilde{x}, \tilde{x}_2) = 0$, from which one can obtain:

$$\alpha = (I - \tilde{P}_1 \hat{P}_1^{-1} - \tilde{P}_2 \hat{P}_2^{-1})^{-1}, \alpha_1 = -\alpha \tilde{P}_1 \hat{P}_1^{-1}, \alpha_2 = -\alpha \tilde{P}_2 \hat{P}_2^{-1} \quad (43.26)$$

According to the optimal state estimation Eq. (43.15) and its covariance Eq. (43.16), the optimal fused global estimation is:

$$\hat{x}(k) = \hat{P}(k) [\tilde{P}^{-1}(k) \tilde{x}(k) + \hat{P}_1^{-1}(k) \hat{x}_1(k) + \hat{P}_2^{-1}(k) \hat{x}_2(k)] \quad (43.27)$$

$$\hat{P}(k) = [\tilde{P}^{-1}(k) + \hat{P}_1^{-1}(k) + \hat{P}_2^{-1}(k)]^{-1} \quad (43.28)$$

43.4 Experiment and Analysis

The integrated test was conducted at *Locata's* Numerella Test Facility (NTF), located in a rural area outside of the city of Canberra, Australia. The NTF covers an area of approximately three hundred acres and is ideally suited for real-world navigation system testing area. A number of *Locata* transmitters were set up to cover the NTF area. The devices that were used in the test include two Leica dual-frequency GNSS receivers (one used as the rover receiver, and the other as the base station), one H764 IMU, and one *Locata* rover unit. The GNSS antenna and *Locata* antenna were mounted with the IMU on the top of a truck. The GNSS data rates and *Locata*

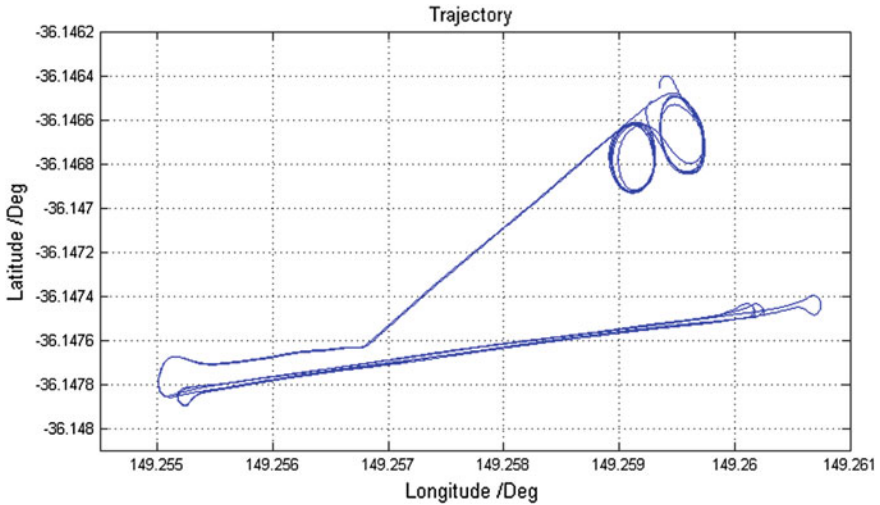


Fig. 43.2 Trajectory from triple-integration system at NTF

data rates were both set to 10 Hz, while the IMU’s data rate was 256 Hz. Both *Locata* and IMU measurements were synchronised with those from GNSS.

The PPP-GNSS and *Locata* solutions were post-processed independently. The initial convergence period of PPP-GNSS was excluded from this triple integration evaluation. The GNSS integer ambiguity-fixed differential carrier phase positioning solution computed by the Leica Geo Office (LGO) software was served as the

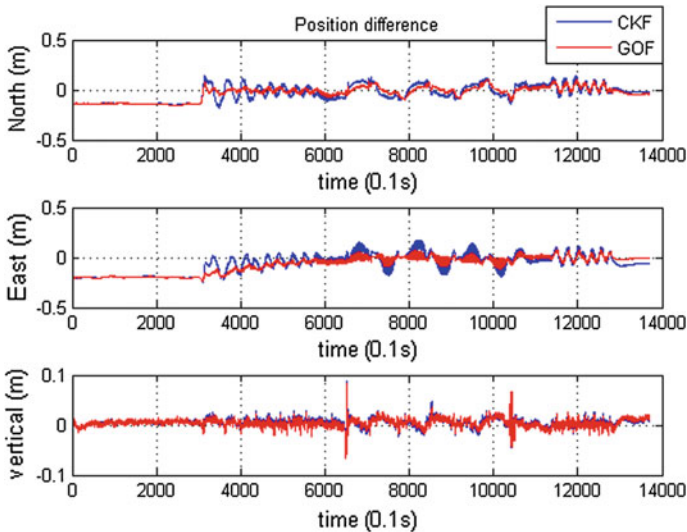


Fig. 43.3 Position comparison between CKF and GOF solutions

Table 43.1 MRSE comparison of triple-integration system and local systems

MRSE (m)	Triple-integration solution			Local system solutions	
	GOF	CKF	FKF	PPP-GNSS/INS	<i>Locata</i> /INS
	0.132	0.141	0.141	0.144	0.192

ground-truth as it had a nominal accuracy of a few centimetres. The trajectory of the field test is shown in the Fig. 43.2.

The field test included circular motion and accelerated motion. As CKF and FKF provide solutions of the same accuracy, Fig. 43.3 shows the comparison only between the GOF and CKF solutions. Since the experiment was conducted across relatively flat terrain, a vertical constraint was applied to account for the poor vertical component observability of *Locata*. It can be seen that the vertical solutions of the two systems are almost the same. However in the horizontal north and east directions the GOF provides a more accurate and a smoother solution. In particular during the circular motion from 250 to 600 s, the GOF solution is obviously better. The Radial Spherical Error (MRSE) of the 3D position error is displayed in Table 43.1. The MRSE of the GOF-based triple-integration solution is 0.132 m,

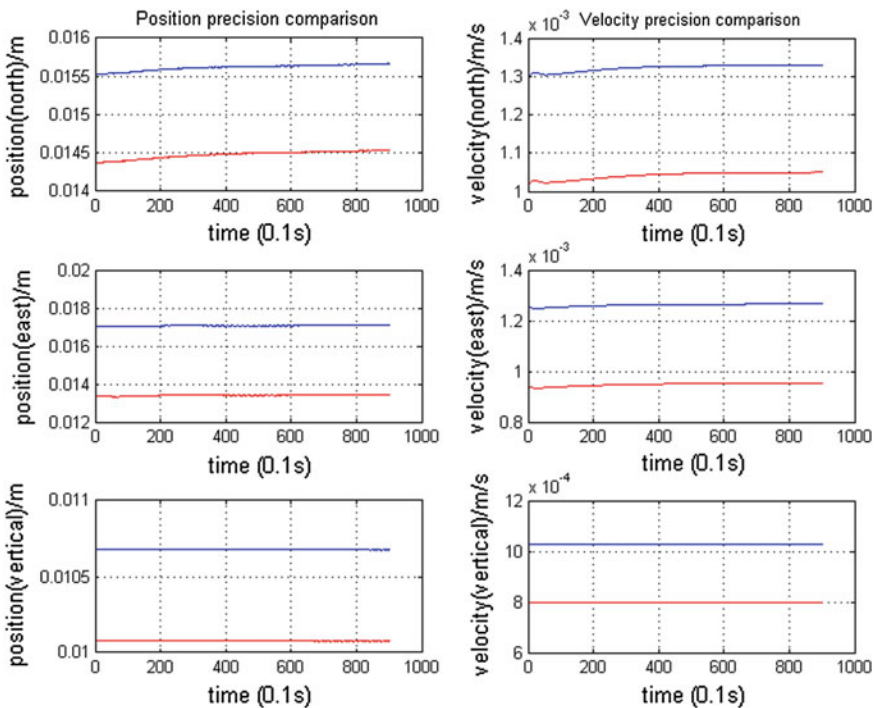


Fig. 43.4 Position and velocity precision comparison ($\sqrt{\hat{P}(k)}$) between CKF and GOF during the last 100 s of the test

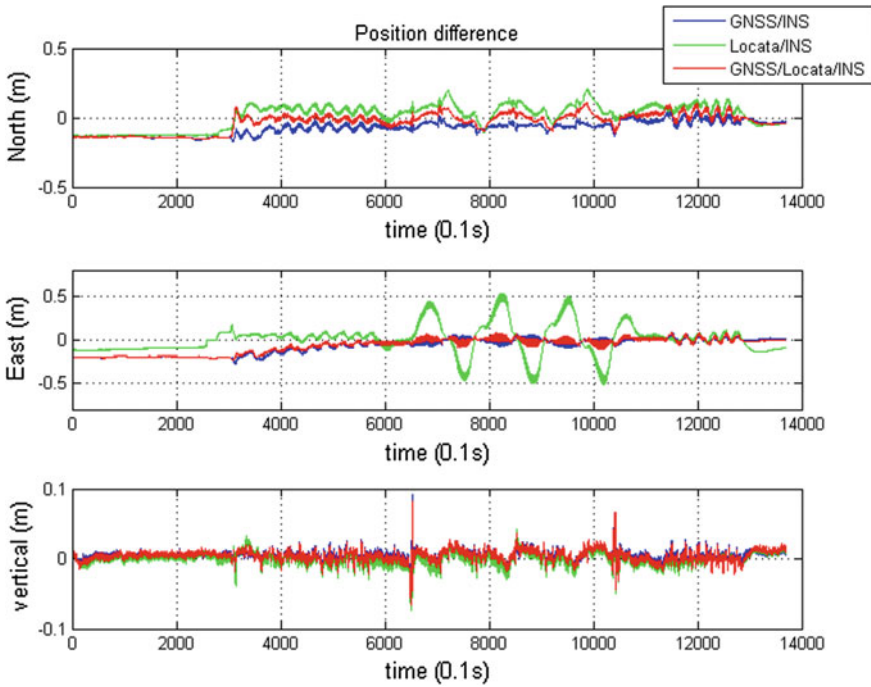


Fig. 43.5 Position difference comparison between GOF-based triple-integration system and local systems

which is lower than that of CKF and FKF’s 0.142 m. This is an improvement of 6.4 %.

The posteriori estimate error covariance $\hat{P}(k)$ was computed. Figure 43.4 shows the square root of the converged position and velocity error covariance $\sqrt{\hat{P}(k)}$ of the CKF and GOF during the last 100 s of the test. The blue colour and red colour denote the CKF and GOF solutions respectively. The left three plots compare the position error covariance in the three direction components, and the right three plots are the corresponding velocity error covariance. It can be seen that the GOF estimate position and velocity error covariance are improved for all three directions in comparison with the results of conventional centralised filtering.

Further comparisons of the solutions of the local systems PPP-GNSS/INS and *Locata*/INS with respect to the triple-integration system are made. Figure 43.5 shows the position comparison of the two local systems with the GOF solutions. It can be seen that the triple-integration approach provides the best positioning solutions for the horizontal (north and east) direction components. From Table 43.1 it can be seen

that the MRSEs of the PPP-GNSS/INS and *Locata*/INS solutions are 0.144 and 0.192 m respectively, while the MRSE of the GOF-based triple-integration solution is 0.132 m. This is lower than either the PPP-GNSS/INS or *Locata*/INS solutions.

As in the case of Fig. 43.4, the position and velocity precisions were investigated. The square root of the a posteriori estimate error covariance was also computed. Figure 43.6 shows the comparison of the square root of the estimated covariance for the GOF-based triple-integration system and the local systems during the last 100 s of the test. The triple-integration solution is plotted in red, and those of the local systems PPP-GNSS/INS and *Locata*/INS are plotted in blue and green respectively. The left three plots are the comparison of the positioning error covariance in three direction components, and the right three plots illustrate the velocity error covariance. It can be seen that the GOF-based triple-integration system has the smallest estimated position and velocity error covariance.

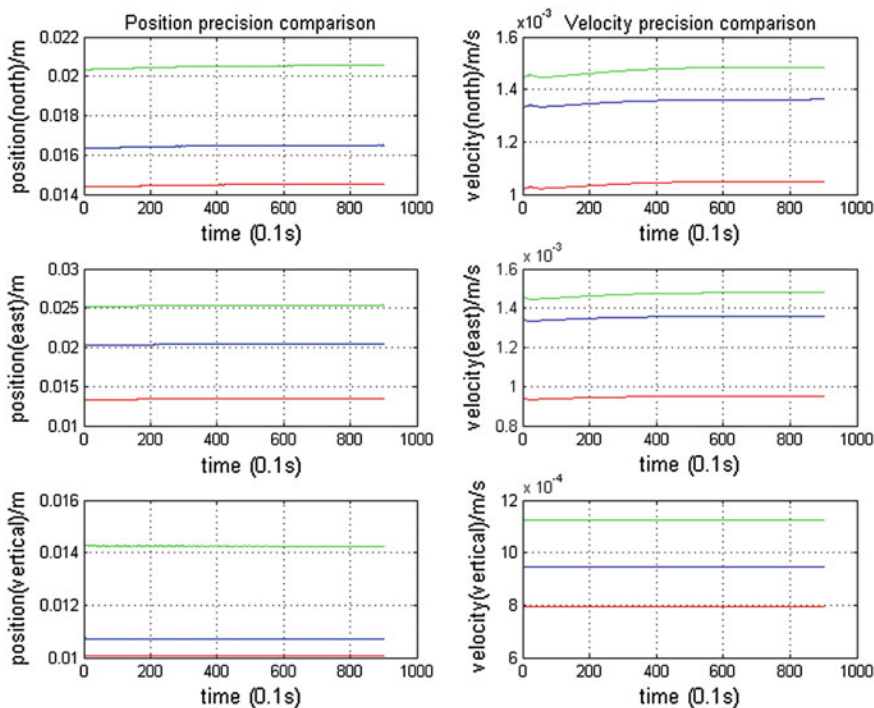


Fig. 43.6 Position and velocity precision comparison between GOF-based triple-integration solution local PPP-GNSS/INS system solution, and local *Locata*/INS system solution for the last 100 s of the test

43.5 Concluding Remarks

This paper describes a PPP-GNSS/*Locata*/INS triple-integration algorithm implemented using a loosely-coupled filtering approach. Conventional centralised filtering and decentralised filtering were discussed. In order to keep the reliability character of decentralised filtering, but to improve its accuracy, the GOF algorithm was developed, based on the recently developed information space concept. The GOF utilises all the information resources, including the raw measurements, the local predictions, and the global predictions. In order to evaluate the system performance, a field experiment was conducted. The comparison between GOF, CKF and FKF triple-integration approaches indicated that the GOF approach does indeed provide the most accurate positioning solution. The a posteriori error covariance for the GOF solution is also smaller than in the case of the other two algorithms. A comparison of the GOF-based triple-integration solution with the local PPP-GNSS/INS and *Locata*/INS solutions shows that the GOF-based approach is superior to the alternative approaches.

Acknowledgments The first author wishes to thank the Chinese Scholarship Council (CSC) for supporting her studies at the University of New South Wales.

References

1. Han S, Wang J (2011) Quantization and colored noises error modeling for inertial sensors for GPS/INS integration. *IEEE Sens J* 11(6):1493–1503
2. Jaradat M, Abdel-Hafez MF (2014) Enhanced, delay dependent, intelligent fusion for INS/GPS navigation system. *IEEE Sens J* 14(5):1545–1554
3. Caron F, Dufflos E, Pomorski D, Vanheeghe P (2006) GPS/IMU data fusion using multisensor Kalman filtering: introduction of contextual aspects. *Inf Fusion* 7:221–230
4. Seo J, Lee JG (2005) Application of nonlinear smoothing to integrated GPS/INS navigation system. *J Global Pos Syst* 4(2):88–94
5. Gao S, Zhong Y, Zhang X, Shirinzadeh B (2009) Multi-sensor optimal data fusion for INS/GPS/SAR integrated navigation system. *Aerosp Sci Technol* 13(45):232–237
6. Lo C, Lynch JP, Liu M (2013) Distributed reference-free fault detection method for autonomous wireless sensor networks. *IEEE Sens J* 13(5):2009–2019
7. Li X, Zhang W (2010) An adaptive fault-tolerant multisensory navigation strategy for automated vehicles. *IEEE Trans Veh Technol* 59(6):2815–2829
8. Li Y (2014) Optimal multisensor integrated navigation through information space approach. *Phys Commun SI Indoor Navig Tracking Part A* 13:44–53
9. Carlson NA, Berarducci MP (1994) Federated Kalman filter simulation results. *J Inst Navig* 41(3):297–321
10. Carlson NA (1996) Federated filter for computer-efficient, near-optimal GPS integration. In: *IEEE position location and navigation symposium (PLANS)*, Atlanta, Georgia, USA, 22–26 April 1996, pp 306–314

Chapter 44

Pulsar Navigation Profile Folding and Measurement Based on Loop Tracking

Xinyuan Zhang, Ping Shuai and Liangwei Huang

Abstract Pulsar navigation is a possible measurement of spacecraft navigation autonomous navigation, and profile folding and measure is one of the core technologies of it, whose performance determines the capability of the final navigation result. Current profile folding measures are commonly based on the presumption that the time of photon arrival can be transformed to the solar system barycentre and the presumption that folding profile is exactitude. However, on the real navigation scene, since the position and the velocity of the spacecraft are not exactly acquirable, folding profile would bring deformation and therefore cause the phase measure error. In order to solve this problem, this paper introduces the loop tracking to the pulsar navigation profile folding and measurement processing. By lock the difference between measure phase and predict phase, the period feedback can be realized to rise the accuracy of the pulsar phase measurement. The result shows that this method could improve the profile folding and phase measure result, and suits for dynamic pulsar navigation.

Keywords Pulsar navigation · Loop tracking · Profile folding · Phase measurement

44.1 Introduction

X-ray pulsar is a kind of nature beacon that has the ability to emit periodic and predictable X-ray pulse. With this nature, a novel autonomic celestial-based navigation method could be realized by measuring the pulse and achieving the high precise timing information. Pulsar navigation can support almost all the missions in solar system while does not suffer from low performance in situations where long

X. Zhang (✉) · P. Shuai · L. Huang
Qian Xuesen Laboratory of Space Technology, Youyi Road 104, Beijing 100191, China
e-mail: blitzer@buaa.edu.cn

range navigation is required. As a result, pulsar navigation is regarded as the best solution for future deep space interplanetary missions [1, 2].

During the pulsar navigation, phase estimation, whose role is calculation the phase difference between the measure pulse and predict pulse, is one of the pre-conditions to establish the navigation equation. Since the X-ray probe of pulsar navigation records the time of arrival (TOA) of received photons rather than the pulse profile, the profile rebuilding should be applied before phase measurement. Ideally, the TOA measured at spacecraft local time can be conversion to Barycentric Coordinate Time (TCB), the pulse profile can be folded under the inertial frame [3], and therefore a veracious profile can be generated. However, such operation relies on the precondition that the position of spacecraft is acknowledged, which is incompatible with the goal of navigation. As a result, in more practical scene, photon folding has to be actualized under the spacecraft local time frame. Because of the movement of the spacecraft, the observation pulse period will not equal to the predict period built at the inertial frame. This problem will bring a transform to the rebuilt profile, which will ulteriorly introduce a phase measurement system error, and finally deteriorate the pulsar navigation performance.

In order to solve this problem, multiple methods of epoch folding, phase estimation as well as tracking loop are combined. The tracking loop is designed to eliminate the dynamic caused by spacecraft movement and to improve the rebuilt profile precision. And ultimately reduce the pulsar navigation system error raised by epoch folding transformation.

44.2 Epoch Folding Transformation Due to Pulse Period Disaccord

The basic method of epoch folding is recording the TOA of photon in the given observation span T , slipping the each TOA result with the predict period of P_0 , and counting the number of photons that in bins which is uniformly divided form the pulsar cycle. This process can be illustrated as Fig. 44.1, and more details of epoch folding can be found in [4].

Fig. 44.1 Demonstration of profile folding

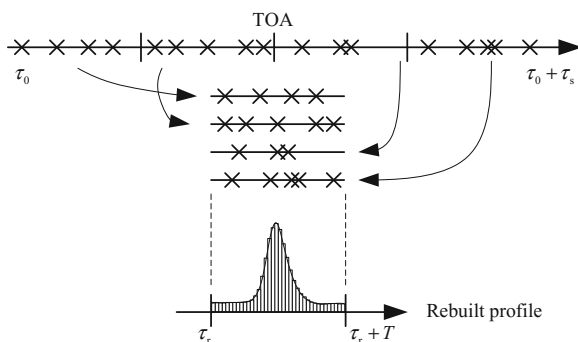
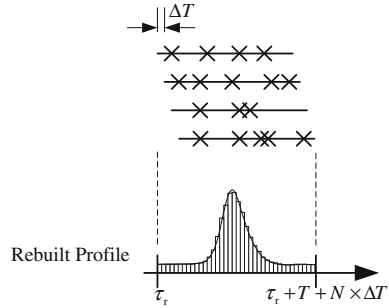


Fig. 44.2 Demonstration of profile folding slipping



Abstractly, as the period P used in epoch folding equals to the actual period which is obtained at the spacecraft, each TOA can be converted to the reference cycle. But when these two periods are not equivalent, the cycle slipping will introduce a cumulate error with the increasing distance between the TOA and the reference cycle. Consequentially, a rebuilt profile transformation will happen as (Fig. 44.2).

Based on the origin of profile transformation, an approximate describe of this transformation can be written as:

$$\tilde{h}(\phi) = \frac{1}{\Delta\phi} \times h(\phi) \otimes [u(\phi) - u(\phi - \Delta\phi)] + n(\phi) \tag{44.1}$$

where $\tilde{h}(\phi)$ and $h(\phi)$ represent rebuilt profile and standard profile respectively, $u(\phi)$ is the step function, $u(\phi) - u(\phi - \Delta\phi)$ makes a rectangle window whose width is $\Delta\phi = \frac{T \times (P - P_0)}{P_0^2}$, and $n(\phi)$ represents the residual measure noise. With the increasing of the observation span T and the growing of the difference of the predict period and the actual period $P - P_0$, the rectangle window will widen. Since the rectangle window can be regarded as a low-pass filter, the above-mentioned tendency will erase the pulse characteristic of the rebuilt profile. This transformation is shown as Fig. 44.3. The simulation B0523+21 photon observation is used to fold the profile with the $\Delta\phi$ equal to 0, 0.1 and 0.2 correspondingly. The practical and theoretical result is plotted as the real line and the dotted line in each subplot.

Based on the Eq. 44.1, when the rebuilt transformation exists, the Cramer-Rao bound of the phase estimation can be written as

$$CRB_{CC}(\phi) = T^{-1} \frac{\int_0^1 \left\{ \lambda'(\phi)^2 \times [\alpha \times \tilde{\lambda}(\phi) + \beta] \right\} d\phi}{\int_0^1 [\alpha \times \lambda'(\phi)^2 \times \tilde{\lambda}'(\phi)^2] d\phi} \tag{44.2}$$

where α and β are the known as source arrival rates and effective background. It is shown that the phase estimation is affected not only by the observation span, but also by the relative period error $(P - P_0)/P_0$. Adverse to tradition view that

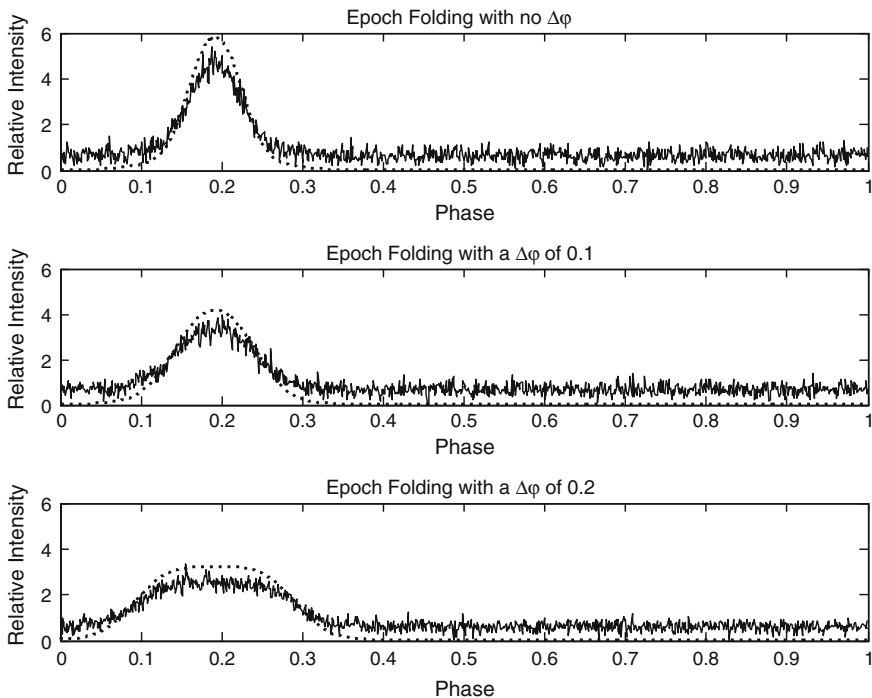
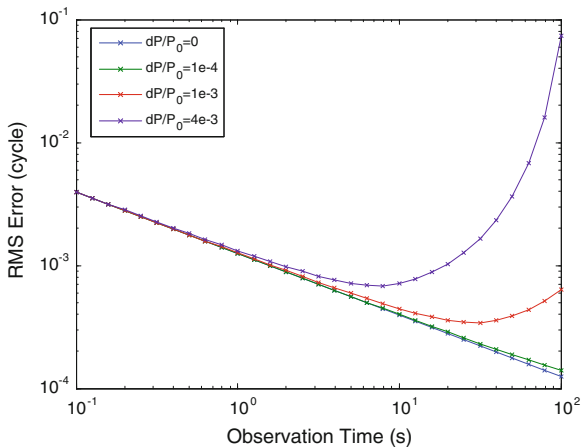


Fig. 44.3 Profile folding deformation due to period error

increasing the observation span will benefit the phase estimation, when relative period error is present, extending the observation span might worsen the phase estimation performance as Fig. 44.4 shows.

Fig. 44.4 Relationship of phase measurement accuracy and observation time as period error presence



44.3 Tracking Loop for the Pulsar Navigation Information Processing

One of the possible solutions to deal with the period error is using tracking loop to automatically estimate the receiving pulse period, correcting the reference period before epoch folding, and commuting the phase estimation error due to rebuilt profile transformation.

44.3.1 Overview of the Pulsar Navigation Tracking Loop

A typical tracking loop is composed by a phase or frequency discriminator, a low-pass loop filter and a voltage-controlled oscillator (VCO). By continually adjusting the output signal, the tracking loop makes the phase or frequency of output signal ails after the input signal. And the aims such as signal can be achieved (Fig. 44.5).

Since the pulsar navigation works on the band of X-ray rather than radio band frequently used by communication and navigation, when tracking loop is introduced to pulsar navigation, some modification should be applied to deal with the profile rebuilt. As a result, the tracking loop for pulsar navigation can be illustrated as Fig. 44.6.

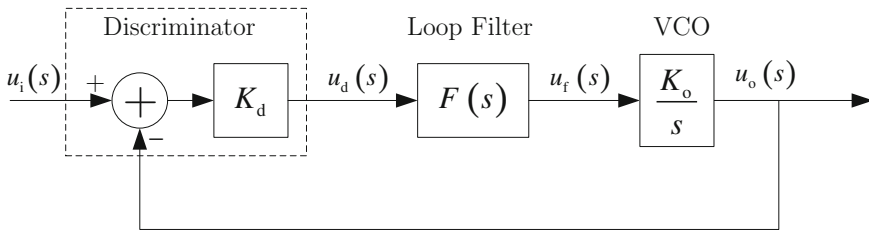


Fig. 44.5 Basic configuration of tracking loop

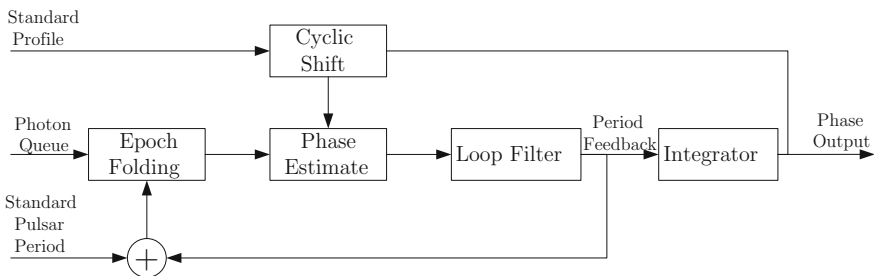


Fig. 44.6 Pulsar navigation signal processing tracking loop

44.3.2 Period Tracking Process

Because of the integral relationship between the pulsar period and the phase $f(t) = \phi'(t)$, when the pulsar phase is locked by the tracking loop, the phase change rate equals to the change value of the period [5] (Fig. 44.7).

Therefore, if the smoothed phase difference between the observation profile and the predict profile can be feedback, the pulsar period could be modified before epoch folding as

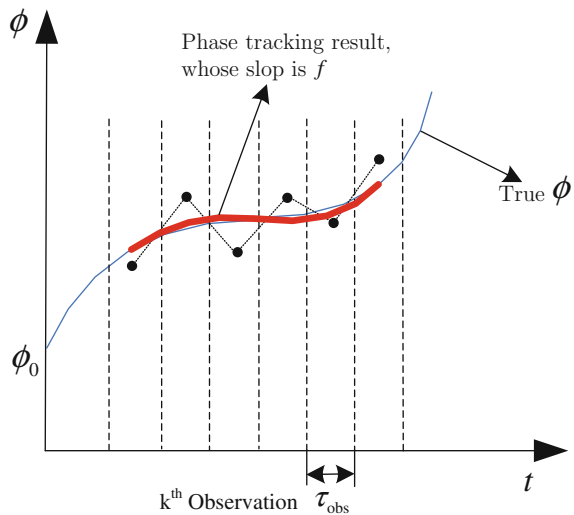
$$P = P_0 - (\phi'_{PLL}(t))^{-1} \tag{44.3}$$

This composes the second loop of the pulsar navigation information processing. From the Eq. (44.2), it can be found that the impact of period error on phase estimation is not simply linear. It should be admitted that, current tracking loop theory seems not be able to deal with such error in pulsar navigation tracking loop. But an efficient method will keep the expectation of the phase estimation equal to 0, which would make the pulsar navigation tracking loop feasible.

44.3.3 Pulsar Navigation Tracking Loop Transfer Function

Different from traditional tracking loop, the tracking loop designed for pulsar navigation contains two feedback terms of period feedback and phase feedback. The period feedback works as standard PLL, whose mission is smoothing the phase difference, and correcting the period employed for epoch folding. The effect of the phase feedback is to modify the phase of standard profile, so as to steady the

Fig. 44.7 Pulsar navigation phase tracking



tracking loop at the linear area. The pulsar navigation tracking loop is operated with the rhythm of the span of epoch folding. The phase estimation equals to the phase discriminator, whose transfer function is K_d . Because there is meaningless to use VCO here, an integrator with the transfer function of $\frac{T_s z^{-1}}{1-z^{-1}}$ is introduced to replace the VCO where T_s is the computing step of the tracking loop. Considering the period variability, a second-order loop filter is applied here, whose transfer function can be written as

$$F(z) = \frac{b_0 + b_1 z^{-1}}{K_d(1 - z^{-1})} \quad (44.4)$$

Therefore, the overall transfer function of the pulsar navigation tracking loop is

$$H_\phi(z) = \frac{T_s(b_0 z^{-1} + b_1 z^{-2})}{(1 - z^{-1})^2 + T_s(b_0 z^{-1} + b_1 z^{-2})} \quad (44.5)$$

At the view of frequency, Eq. (44.4) describes a low-pass filter, which can smooth the phase estimation result by restraining the high frequency component of it. Ignoring the quantization error and the error relative to oscillator of the PLL, only taking the phase estimation noise σ_{Phase} into account, the PLL will reduce the output noise as [6, 7]

$$\sigma_{\text{PLL}} \approx \sigma_{\text{Phase}} \sqrt{B_L T_s} \quad (44.6)$$

where the parameters of T_s and B_L present the sampling period and the loop noise band. The former equals to the span of epoch folding, and the latter is determined by the second-order loop filter, those exponential value is

$$B_L = 0.53 \sqrt{\frac{b_0 + b_1}{T_s}} \quad (44.7)$$

44.4 The Performance Simulation of the Pulsar Navigation Tracking Loop

In order to evaluate the pulsar navigation tracking loop described in this paper, a series of simulation is applied to test the signal processing of Crab (B0531+21). The detail parameter is set based on the reference of [4, 8]. Supposing the detector is installed on a LEO spacecraft with the orbit of 600 km. The totally continual observation is 7200 s. The photon TOA is generated based on the theoretic pulsar phase at the spacecraft local time that is simulated with the information of spacecraft position and the solar system ephemerides. Pulsar navigation tracking loop takes charge of tracking and processing the TOA record, obtaining the smooth

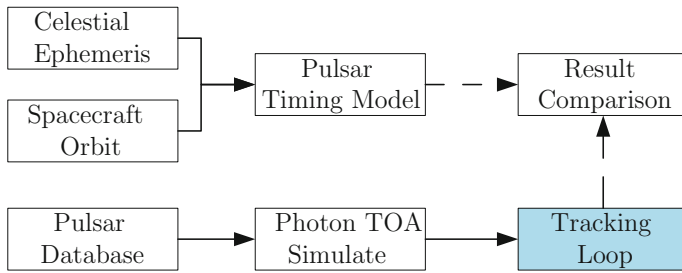


Fig. 44.8 Configuration of phase tracking simulation

phase output, which is compared to the theoretic pulsar phase so as to achieve the tracking error. The simulation logic diagram is illustrated as Fig. 44.8.

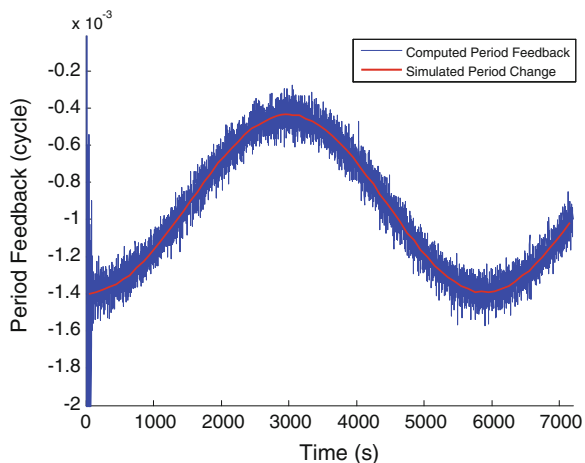
The processing step of the tracking loop is 1 s, which is equal to epoch folding rhythm. The band of the loop is 0.05 Hz to track the pulsar signal.

The period feedback result is shown in Fig. 44.9. The red line exemplifies the theoretic period transformation due to the movement of the spacecraft, and the blue line illustrates the period feedback obtained from the tracking loop. It can be found that when the tracking tends to steady, the period feedback could lock on the period change, which would be benefit to epoch folding result.

The phase estimation error is shown as Fig. 44.10. The blue line presents the direct phase estimation error without any correct method, whose standard variation is about 5.9×10^{-4} cycle. The red line shows the phase estimation error output get from the tracking loop, and the standard variation of this output is about 3.3×10^{-4} cycle. The phase estimation precision is improved by 44 %. If the movement of the spacecraft is more complex, the advance of the tracking loop will be more notable.

As a conclusion, the pulsar navigation tracking loop could deal with the phase estimation and the correction at the spacecraft dynamic condition. In the future X-ray pulsar navigation mission, this technique might be used to achieve the better navigation preferment.

Fig. 44.9 Period Feedback result



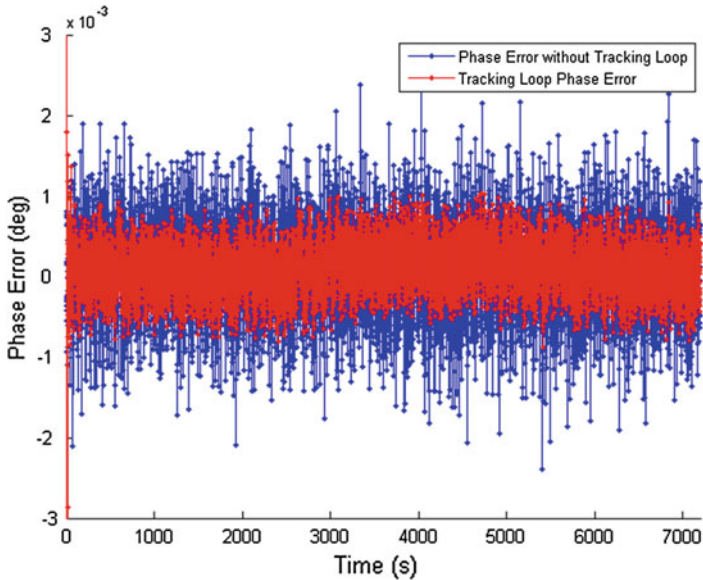


Fig. 44.10 Phase tracking result

Acknowledgments The authors hereby acknowledge with thanks to the financial supporting from National Natural Science Foundation of China (no. 61403391).

References

1. Sala J, Urruela A, Villares X et al (2004) Feasibility study for a spacecraft navigation system relying on pulsar timing information. ESA, Ariadna Study 03/4202
2. Sheikh S, Hanson J, Graven P, Pines D (2011) Spacecraft navigation and timing using X-ray pulsars. *Navigation* 58(2):165–186
3. Emadzadeh A, Speyer J (2011) *Navigation in space by X-ray pulsars*. Springer, New York
4. Huang LW (2013) *Theory and algorithm study in X-ray pulsar autonomous navigation based on pulsar timing model*. Tsinghua University, Beijing
5. Golshan R, Sheikh S (2007) On pulse phase estimation and tracking of variable celestial X-ray sources. In: ION 63rd annual meeting, pp 413–422
6. Aharoni R, Lee D (2001) On the achievability of the cramér-rao for poisson distribution. *IEEE Trans Inf Theory* 47(5):2096–2100
7. Xiang W, Xu B, Mou WH, Wang FX (2013) A clock error calibration algorithm based on phase lock loop in GNSS time synchronization receiver. *J National University Defense Technol* 35 (2):115–119
8. Lyne A, Graham-Smith F (2005) *Pulsar astronomy*, 3rd edn. Cambridge University Press, Cambridge

Chapter 45

Application of an Improved K Nearest Neighbor Algorithm in WiFi Indoor Positioning

Yeqing Fang, Zhongliang Deng, Chen Xue, Jichao Jiao, Hui Zeng, Ruoyu Zheng and Shunbao Lu

Abstract In this paper, K nearest neighbor algorithm is improved in fingerprint information matching in WiFi indoor positioning system. Because the original K nearest neighbor algorithm ignores the relationship between the neighboring points, it doesn't improve matching accuracy, and its positioning accuracy is not better. There is no effective treatment group sample points, matching time is greatly increased. In this paper, we study a modified K nearest neighbor algorithm in the application of WiFi indoor positioning. Because K neighbor points are given different weights according to certain rules, the matching accuracy is improved before matching method. And because of using grouping pretreatment of the sample space, the time of position matching accuracy is reduced. Therefore the positioning precision is improved.

Keywords WiFi · Indoor positioning · K nearest neighbor algorithm

45.1 Introduction

With the development of mobile Internet and mobile intelligent devices, Location Based Servers has become the frontier in the research of information technology. With the four major global satellite navigation system support, outdoor location service has been widely into people's life. People are in the indoor environment for more than 80 % of time, with increasing number of the large building, indoor location service have an even broader prospect in application of commercial application, public security and other aspects. In indoor environment, signal of satellite system can't be used because of building occlusion and multipath. At present, indoor positioning technology mainly contains WiFi technology, ultra wideband, pseudo

Y. Fang (✉) · Z. Deng · C. Xue · J. Jiao · H. Zeng · R. Zheng · S. Lu
School of Electronic Engineering, Beijing University of Posts and Telecommunications,
Beijing 100876, China
e-mail: 592202890@qq.com

satellite. WiFi technology has been widely used in the indoor localization for erection of low cost, convenient access, easy expansion, easy popularization, etc. And WiFi hot covers a wide range in the urban population concentration region, at the same time, more and more intelligent devices support WLAN network [1].

Indoor positioning method contains Arrival of Angle (AOA), Time of Arrive (AOA), Time Difference of Arrival (TDOA), hybrid positioning method of TDOA and AOA, Received Signal Strength Indication method. TOA and TDOA method needs high precision hardware synchronization. AOA method needs direction antenna. And in non-line-of-sight environment, they are seriously affected by multipath propagation. RSSI technology is widely used in WiFi positioning, and it works by fingerprint matching. Indoor environment is complex, but basic pattern remains the same. In a specific location, the characteristics of the wireless signal (signal number, phase, and intensity) presents the particularity highly. RSSI is the unique “fingerprint” to identify the location, and then calculate the position according to the fingerprint matching algorithm.

Fingerprint matching algorithm based on RSSI mainly contains nearest neighbor, K nearest neighbor algorithm. But they are only matched and get a single relation between training points in fingerprint database to be positioning point, and get nearest neighbor points or K nearest neighbors. They ignore the deeper relationship between neighbor points and other reference point. There is no further excavation and utilization of useful information on the RSSI fingerprint database and no effective grouping of fingerprint database information. Therefore it's difficult to further shorten matching time and improve the positioning accuracy.

45.1.1 WiFi Fingerprint Matching Positioning Method

The method is divided into two phases: offline training phase and online positioning phase [2]. On the offline training phase, the position information and AP signal strength information are associated together to form the RSSI signal vector into the database through the acquisition equipment acquiring training sequence. In accordance with certain dense degree, collecting all fingerprints in the region, then they completed the establishment of fingerprint database. As shown in Fig. 45.1.

Online positioning phase is to make similarity matching real-time receiving AP signal intensity with the fingerprint data in the database to get the best position after the fingerprint database is finished. As shown in Fig. 45.2.

45.2 K Nearest Neighbor Matching Algorithm

K nearest neighbor algorithm is matching RSSI vector, which is composed of the signal intensity of WiFi real time received by mobile terminal, with the fingerprint data in the fingerprint database in the ways of certain matching calculation. And

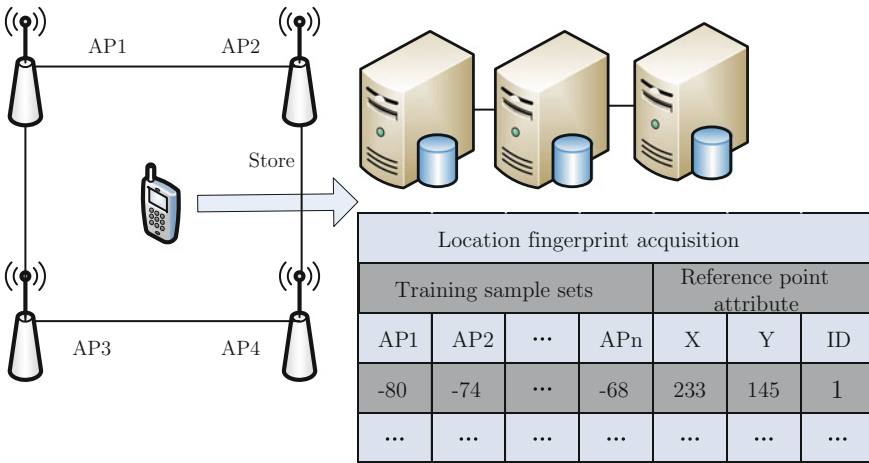


Fig. 45.1 Offline training phase

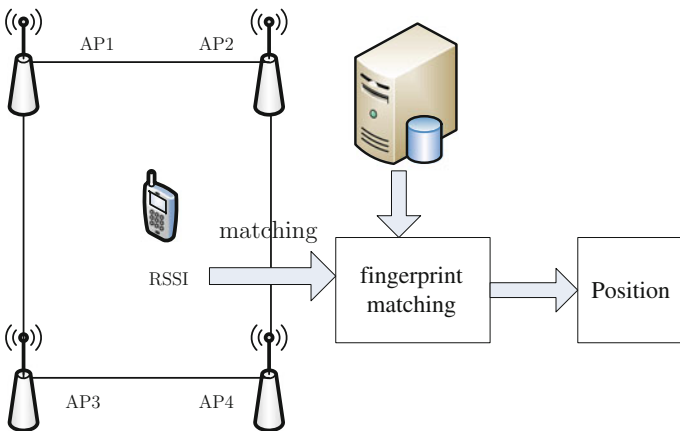


Fig. 45.2 Online positioning phase

selecting the K fingerprint data which has the minimum distance value and regarding mean position coordinates of K fingerprint information as the positioning result data.

Two example points in feature space distance is reflected in two instances of similar degree. The feature space of K nearest neighbor algorithm is generally an n-dimensional vector space R^n . Assuming that the mobile terminal real-time received n AP signal, the RSS vector is $[s_1, s_2, \dots, s_n]$. The mean RSS matrix in the fingerprint database is $[S_1, S_2, \dots, S_n]$. Design the fingerprint database has m reference points and n AP and m fingerprint data. And because of $S_i = [S_{i1}, S_{i2}, \dots, S_{in}]$, $1 \leq i \leq m$, measured distance formula between the vectors are shown as follows.

$$d_i = \left(\sum_{j=1}^n |s_i - S_{ij}|^q \right)^{\frac{1}{q}} \quad (45.1)$$

In the formula, when $q = 1$, it's known as the Manhattan distance and when $q = 2$, it's known as the Euclidean distance.

K nearest neighbor algorithm is selecting K reference points in choice from the d_i above in the way of the small to large. And then using the formula (45.2) computing the estimated location of the average coordinates as the test point:

$$(\hat{x}, \hat{y}) = \frac{1}{K} \sum_{i=1}^K (x_i, y_i) \quad (45.2)$$

Among them, (x_i, y_i) represent the i -th corresponding physical coordinates of K reference point. (\hat{x}, \hat{y}) represent the estimation of the test point coordinates. When $K = 1$, K nearest neighbor algorithm degenerate to the nearest neighbor algorithm.

The K nearest neighbor matching algorithm above has two serious problems. One is ignoring the relationship between the training points in the fingerprint database itself. The same weight will bring training points of large error. Therefore, noise points have the bad influence on the positioning results and the accurate rate is relatively low, and resulting in reduced accuracy positioning. Two is needing to match all the fingerprint data in all the fingerprint database. When the sample size is large in fingerprint database, searching is a large quantity, efficiency is very low.

45.3 Improved K Nearest Neighbor Matching Algorithm

45.3.1 Weighted for K Neighbor Points

K neighbor points are given different weights according to certain rules. The formula is as follows:

$$\begin{cases} (\hat{x}, \hat{y}) = \frac{1}{K} \sum_{i=1}^K w_i(x_i, y_i) \\ \sum_{i=1}^k w_i = 1 \end{cases} \quad (45.3)$$

In the formula, w_i is the weight of i -th nearest neighbor point. We need to give the K neighbor points different weight to improve the system positioning accuracy. According to signal strength based localization algorithm of active classical RFID system, the system selects some reference point and arrange reference tags in the location area in order to get several reference label which is nearest to locating tag,

by comparing the signal strength to be tracking tag and reference label values [3]. Then estimating the coordinates of positioning label based on these coordinates of weighted reference tags. The new design for weight is referring to this system' design idea above, the formula is:

$$w_i = \frac{\frac{1}{d_i^2}}{\sum_{j=1}^k \frac{1}{d_j^2}} \tag{45.4}$$

Square of Euclidean distance d reflects the changes of weights. The smaller the d , the greater the weight. And it's directly depended on the Euclidean distance of the current d_i . It enhance the weight ratio of Euclidean distance d_i .

45.3.2 Grouping Pre-Treatment

Space was divided by the actual geographic information, which each region of WiFi information will have obvious difference. Assume that the number of partition is s , we will clarify the information of fingerprint points into s groups. Each group form a virtual space, which is used to represent a corresponding physical space [4]. The specific group can be according to the signal amount above, every n element into a group. This will eventually form a group M signals, which are collected from m different physical space.

Getting a group of signal strength information after a scan and storing in an array A . Each of the scan results stored in the array list. When the packet is completed, we get the formation of M virtual space and each virtual space has N signal information. Now we construct an array $C_i = (1, 2, \dots, m)$ which can represent a virtual space, we define it as:

$$C_i[j] = \frac{\sum_{l=(i-1)n}^{in} list(l)[j]}{n} \tag{45.5}$$

Suppose that there are M access point to the location in the scene (AP) and N reference points (RP). The subscript $m = 1, 2, \dots, M$ and $n = 1, 2, \dots, N$ is used to present the m -th AP and the n -th RP. The q -th sampling RSS vector of the n -th AP is expressed as $S_n^q = (S_{n1}^q, S_{n2}^q, \dots, S_{nM}^q)$. Each RP can have a total of Q sampling, and all sampling RSS vector of a RP can be expressed as:

$$S_n = \begin{pmatrix} S_n^1 \\ S_n^2 \\ \vdots \\ S_n^Q \end{pmatrix} = \begin{pmatrix} S_{n1}^1 & S_{n2}^1 & \dots & S_{nM}^1 \\ S_{n1}^2 & S_{n2}^2 & \dots & S_{nM}^2 \\ \vdots & \vdots & \ddots & \vdots \\ S_{n1}^Q & S_{n2}^Q & \dots & S_{nM}^Q \end{pmatrix} \tag{45.6}$$

The next processing is to settle the Q sampling. Firstly, the sampling with the strongest RSS is divided into a group. For example, the strongest RSS vector in RSS is the i th AP, which is $S_{ni}^q \geq S_{nm}^q, i \neq m$. In this way, RSS vector is represented as $S_n^q(i)$, and $S_n(i)$ is used to represent a set of this class of RSS vector. In addition, we define N_{ni} scalar of a set size record, $N_{ni} = |S_n(i)|$ is corresponding to the number of RSS vector expressed in the group. Each RP sample number is Q , so it is $\sum_{i=1}^M N_{ni} = Q$.

Define a packet threshold $\alpha, 0 < \alpha < 1$. If the number of RSS vectors is $N_{ni} > \alpha Q$ in the group $S_n(i)$, then take the RSS vector mean within the group as a fingerprint of this RP, and the fingerprint representation for $\bar{S}_n(i) = (\bar{S}_{n1}(i), \bar{S}_{n2}(i), \dots, \bar{S}_{nM}(i))$. This reference point will be grouped into $C_i, i = 1, 2, \dots, M$, and $\bar{S}_n(i)$ is fingerprinting of the reference point in the group. It's different with existing fingerprint, which use all RSS vector averaging as its fingerprint.

As we can see, the algorithms in this paper, the total number of all reference points' fingerprint is greater than N . The surface seem to increase the amount of matching positioning stage, but due to the grouping algorithm, fingerprint quantity in fact is far small than N .

Using packet threshold is a method to remove some of the occasional poor sampling RSS training stage and also to reduce fingerprint mass, to save storage space and improve the response time. For those RSS vector whose number is less than the group, we directly ignore these vector. The grouping effect is shown in Fig. 45.3.

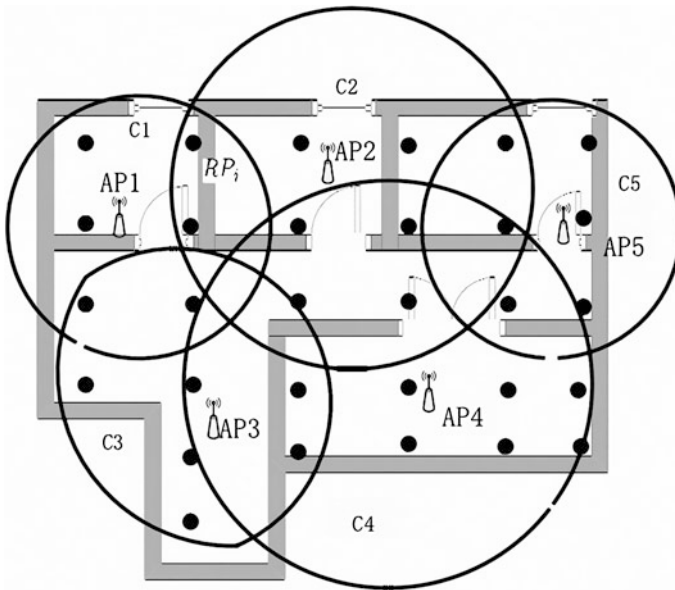


Fig. 45.3 Group effect

45.4 Application Examples and Results Analysis

The improved K nearest neighbor algorithm above is carried out in the actual environment that the smaller the distance between sample fingerprint, the more the number of the fingerprint database, the higher rate of accuracy in the different fingerprint spacing.

The improved K nearest neighbor algorithm has different reasonable weights, so matching accuracy is obviously higher than that of the ordinary K nearest neighbor algorithm. Performance in the fingerprint space which is larger than 10 m is obviously superior to the ordinary K nearest neighbor method, as shown in Fig. 45.4.

The improved K nearest neighbor algorithm is improved to join the training sample packet preconditioning and its matching time is obviously better than the ordinary matching time in response process in different spacing. It shown in Fig. 45.5.

Fig. 45.4 Matching accuracy

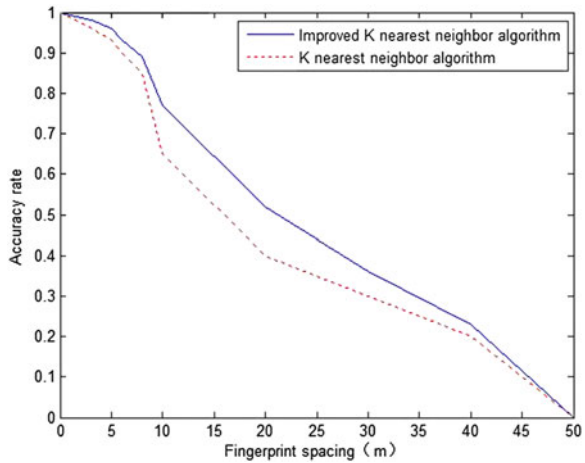


Fig. 45.5 Matching time

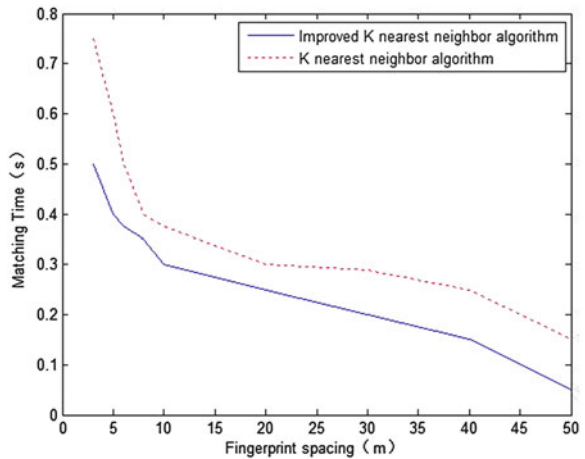
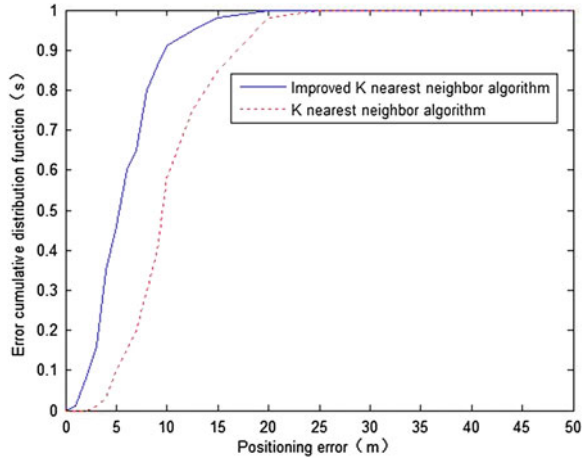


Fig. 45.6 Error cumulative distribution function



Finally, two kinds of matching method's error cumulative distribution function is obtained. The probability of ordinary K nearest neighbor matching method positioning accuracy, which is better than 10 m, is 60 %. The probability of improved K nearest neighbor matching method positioning accuracy, which is better than 6.4 m, is 60 %. We significantly improve the positioning accuracy. It shown in Fig. 45.6.

Acknowledgments Foundation project: The National High Technology Research and Development Program ("863"Program) of China (No. 2012AA120801) and The National Natural Science Foundation of China (No.61372110).

References

1. Xing P, Tian Z, Dongmei S (2008) Estimated distance and NNSS based indoor location algorithm. *Commun Technol* 41(4):152–153
2. Li S (2014) Application research on indoor location technology based on WiFi, vol 3. Nanjing Normal University, Nanjing
3. Feng Z (2009) Design and implementation of positioning system based on WiFi technology, vol 12. Beijing University of Post and Telecommunication, Beijing
4. Xiaojian W, Zheng X et al (2012) Design of infrastructure-free WiFi indoor localization. *J Commun* 33(z2):25–34

Chapter 46

A New Pulse Time-of-Arrival Estimation Method for X-Ray Pulsar Navigation

Qingqing Lin, Ping Shuai and Liangwei Huang

Abstract X-ray pulsar navigation (XPNAV) is an important technology for long time and high precious spacecraft autonomous navigation. XPNAV system can offer position, velocity and accurate time navigation information for spacecraft in near-earth orbit, geostationary orbit, elliptic orbit or interplanetary orbit. XPNAV has an extremely important military use and wide application prospect. In XPNAV system, pulse time-of-arrival (TOA) is a basic observation, and its estimation precious decides the precious of XPNAV's timing and position determination, thus high precious pulse TOA estimation is an important technology in XPNAV system. To improve the pulse TOA estimation precious of XPNAV system, we proposed a pulse TOA estimation method based on correlation information. First, based on the time observation data of photons radiating for X-ray pulsar to the spacecraft, the pulse folding profile is obtained by epoch folding. Then, we constructed the new observation equation by analyzing the correlation result between the folding profile and normal profile. At last, eigen-decomposed the new observation equation we can obtain two mutually orthogonal feature vectors, so the pulse TOA can be estimated by the orthogonal property of the two vectors. Simulation results show the validity of the proposed methods, and can improve the TOA estimation precious compared with the traditional correlation method.

Keywords X-ray pulsar navigation (XPNAV) · Time-of-arrival (TOA) · Correction method · Estimation precious

Q. Lin (✉)

Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology,
Beijing 100094, China
e-mail: linqinglqq@163.com

P. Shuai · L. Huang

China Academy of Space Technology, Beijing, China

© Springer-Verlag Berlin Heidelberg 2015

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015*

Proceedings: Volume III, Lecture Notes in Electrical Engineering 342,

DOI 10.1007/978-3-662-46632-2_46

46.1 Introduction

X-ray Pulsar-based Navigation (XPNAV) is a revolutionary technology, which is an effective way of autonomous navigation for deep space and interstellar probes as well as navigation satellite constellations and high Earth orbit satellites. In recent years, international spaceflight institutions had turned their attentions to XPNAV system for its important strategic significance for national defense and national economic construction [1–4].

In XPNAV system, pulse time-of-arrival (TOA) is a basic observation, the basic principle for XPNAV system is to obtain the location of the observation time using the TOA between spacecraft and the datum mark (such as solar system barycenter). The precious of TOA estimation decides XPNAV's timing and location accuracy, thus high precious pulse TOA estimation is an important technology in XPNAV system [5, 6].

Emadzadeh had proposed a series methods for X-ray pulse TOA estimation [7–10]. In [8, 9], the proposed methods uses epoch folding to obtain the integrated X-ray pulse profile, then [8] uses least square estimate method and [9] uses correlation technology to realize the comparison between the integrated X-ray pulse profile and the standard X-ray pulse profile. Based on the principle of maximum likelihood estimate technology, the pulse TOA had been estimated by maximum likelihood function utilizing the date of X-ray photons [10]. Then grid searching algorithm had been used to obtain the peak of the maximum like hood function [11]. Rinauro reconstruct the pulse TOA estimation problem as a circular shift parameter estimation problem, then solve the maximum likelihood function by discrete Fourier transformation [12]. By some new technologies, other X-ray pulse TOA estimation methods had been proposed recently. XIE restrained the Gauss noise by bispectrum processing, and proposed time offset measurement algorithm for pulsar integrated pulse profiles [13]. SU proposed a new time delay measurement algorithm to improve the precision utilizing paraboloidal interpolation method [14]. To improve the pulse TOA estimation precious of XPNAV system, we proposed a pulse TOA estimation method based on the Fourier transformation of correlation information. According to the high precision property of eigen-decomposed, the TOA estimation precision has been improved.

46.2 Basic Observation of XPNAV System

In XPNAV system, first we need to measure the time-of-arrival (TOA) of the pulse arriving to the spacecraft, then compared it with the TOA of the same pulse arriving to the solar system barycenter. The pulse TOA in the position of the solar system barycenter can be forecast by the pulsar timing model, but the calculation for the pulse TOA of the spacecraft is a complicated process. X-ray detector can obtained the photon TOAs when they hit the detecting material. But the photon TOA cannot

be used directly in XPNAV system because of the low X-ray photon flux of many pulsars, so we need to use epoch folding to obtain the integrated X-ray pulse profile [9]. The integrated X-ray pulse DOA is the basic observation of XPNAV system, then we can rewrite the pulse DOA as pulse phase angle ϕ based on the stably timing observation data of pulsar, the connection between DOA and ϕ is $\text{TOA} = \phi \cdot P$, where P is the pulse period. Then we can obtain the pulse TOA by comparing the observation epoch folding X-ray pulse profile with the standard pulse profile of the same pulsar.

46.3 Pulse DOA Estimation of X-Ray Pulsar

46.3.1 Observation Pulse Profile and Standard Pulse Profile

The standard pulse profile of pulsar is one of the most important data in XPNAV system, for the same pulsar, the relationship of its observation pulse profile and standard pulse profile at the same energy wave band is

$$r(t_i) = \alpha \cdot s(t_i - \tau_0) + \omega(t_i) \quad (46.1)$$

where, α is the range gene, τ_0 is the time delay between observation pulse profile and the standard pulse profile. In (1), we have the following assumptions:

- (a) It is assumed that the pulse period P of pulsar signal is known;
- (b) According to the epoch folding method of [10], divide the pulse period into N_b bin evenly while the length of each bin is Δ . Assumed the observation time contains N periods, then fold the data of the photon TOAs into N_b bin, so we obtained the observation pulse profile named $r(t_i)$, where $i = 1, 2, \dots, N_b$;
- (c) Sampled the standard pulse profile with uniformly-spaced of Δ , we can obtain the discrete standard pulse profile $s(t_i)$, where $i = 1, 2, \dots, N_b$, and $s(t_i) = s(t_i + P)$;
- (d) Only observe a single pulsar, then the other signals are equivalent to noise $\omega(t_i)$, while each time bin has more than 20 photons, the noise is Gaussian [15]. Then assumed the noise and pulse signals are uncorrelated.

46.3.2 Correlation Method for Pulse TOA Estimation

In XPNAV system, the key technology for pulse TOA estimation is high-precision time delay calculation between observation pulse profile and standard pulse profile, which is τ_0 in (1). The conventional method is correlation method, which estimates time delay of two signals by their correlation function. The correlation function is

$$\begin{aligned}
 R_{sr}(\tau) &= \sum_{i=1}^{N_b} [s(t_i - \tau)r(t_i)] \\
 &= \alpha \cdot \sum_{i=1}^{N_b} [s(t_i - \tau) \cdot s(t_i - \tau_0)] = \alpha \cdot R_{ss}(\tau - \tau_0)
 \end{aligned} \tag{46.2}$$

Due to the property of correlation function,

$$|R_{ss}(\tau - \tau_0)| \leq R_{ss}(0) \tag{46.3}$$

So, there is a peak of $R_{sr}(\tau)$ in the point of $\tau = \tau_0$, then the estimation of TOA is

$$\hat{\tau} = \arg \left\{ \max_{\tau} [R_{sr}(\tau)] \right\} \tag{46.4}$$

where, $\arg\{\cdot\}$ denotes the independent variable of the function, $\max[\cdot]$ denotes the maximum value.

The corresponding time delay value of the correlation functions' peak point is the estimation of pulse TOA.

46.3.3 Novel Method for Pulse TOA Estimation

The correlation method is simple, but its estimation accuracy is susceptible to the property of signal spectral and noise. To improve the accuracy of pulse TOA estimation, we proposed a novel method utilizing the reconstruct cross-correlation function based on subspace decomposition.

Assumed the discrete Fourier transform (DFT) of standard pulse profile $s(t_i)$ and observation pulse profile $s(t_i)$ are $\hat{S}(k)$ and $\hat{S}(k)$ respectively.

$$r(t_i) = \frac{1}{N_b} \sum_{i=1}^{N_b} \hat{R}(k) \cdot e^{j\frac{2\pi t_i k}{N_b}} \tag{46.5}$$

$$s(t_i) = \frac{1}{N_b} \sum_{i=1}^{N_b} \hat{S}(k) \cdot e^{j\frac{2\pi t_i k}{N_b}} \tag{46.6}$$

Reconstruct the cross-correction function

$$R(\tau) = \sum_{i=1}^{N_b} [s(t_i - \tau)r^*(t_i)] \tag{46.7}$$

From (46.6) and (46.7), we have

$$R(\tau) = \sum_{i=1}^{N_b} \left[\sum_{k=1}^{N_b} \hat{S}(k) \cdot e^{j\frac{2\pi(t_i-\tau)k}{N}} \cdot r^*(n) \right] = \sum_{k=1}^{N_b} \hat{S}(k) \cdot \left[\sum_{i=1}^{N_b} r^*(t_i) \cdot e^{j\frac{2\pi t_i k}{N_b}} \right] \cdot e^{-j\frac{2\pi \tau k}{N_b}} \quad (46.8)$$

Let

$$\begin{aligned} \Upsilon(k) &= \sum_{i=1}^{N_b} r^*(t_i) \cdot e^{j\frac{2\pi t_i k}{N}} = \sum_{i=1}^{N_b} [\alpha \cdot s(t_i - \tau_0) + \omega(t_i)]^* e^{j\frac{2\pi t_i k}{N_b}} \\ &= \left\{ \alpha \cdot \hat{S}(k) \cdot e^{-j\frac{2\pi k \tau_0}{N_b}} + \hat{W}(k) \right\}^* \end{aligned} \quad (46.9)$$

where,

$$\hat{W}(k) = \sum_{i=1}^{N_b} \omega(t_i) e^{-j\frac{2\pi t_i k}{N_b}} \quad (46.10)$$

From (46.8), (46.9) and (46.10), we have

$$\begin{aligned} R(\tau) &= \sum_{k=1}^{N_b} \hat{S}(k) \cdot \Upsilon(k) \cdot e^{-j\frac{2\pi \tau k}{N_b}} \\ &= \sum_{k=1}^{N_b} \left\{ \alpha^* \cdot |\hat{S}(k)|^2 \cdot e^{-j\frac{2\pi k \tau_0}{N_b}} + \hat{W}(k) \right\} \cdot e^{-j\frac{2\pi \tau k}{N_b}} \end{aligned} \quad (46.11)$$

Let

$$\mathfrak{R}(k) = \hat{S}(k) \cdot \Upsilon(k) = \alpha^* \cdot |\hat{S}(k)|^2 \cdot e^{-j\frac{2\pi k \tau_0}{N_b}} + \hat{W}(k) \quad (46.12)$$

According to (46.11) and (46.12) we know that $\mathfrak{R}(k)$ is the DFT of $R(\tau)$. So we can estimate the pulse TOA by $\mathfrak{R}(k)$, then rewrite $\mathfrak{R}(k)$ in matrix form

$$\mathbf{X} = \begin{bmatrix} \mathfrak{R}(1) \\ \mathfrak{R}(2) \\ \vdots \\ \mathfrak{R}(N_b) \end{bmatrix} = \mathbf{A}(\tau_0) \cdot \alpha^* + \mathbf{W} \quad (46.13)$$

where, \mathbf{X} denotes the $N_b \times 1$ array data vector, $\mathbf{W} = [\hat{W}(1) \hat{W}(2) \dots \hat{W}(N_b)]^T$ denotes the noise vector and it is constituted by the DFT of noise. $\mathbf{A}(\tau_0)$ denotes the steering vector of \mathbf{X} , and we have

$$\mathbf{A}(\tau_0) = \begin{bmatrix} e^{-j\frac{2\pi(\tau_0)}{N_b}} & & & & \\ & e^{-j\frac{2\pi(2\tau_0)}{N_b}} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & e^{-j\frac{2\pi(N_b\tau_0)}{N_b}} \end{bmatrix} \begin{bmatrix} |\hat{S}(1)|^2 \\ |\hat{S}(2)|^2 \\ \vdots \\ |\hat{S}(N_b)|^2 \end{bmatrix} \quad (46.14)$$

That $\mathbf{A}(\tau_0)$ is contains information of the pulse TOA.

Calculate the covariance matrix of \mathbf{X} , we have

$$\mathbf{R}_X = \mathbf{E}[\mathbf{X}\mathbf{X}^H] = \mathbf{A}(\tau_0)\mathbf{R}_S\mathbf{A}^H(\tau_0) + \mathbf{R}_W \quad (46.15)$$

where, $\mathbf{R}_S = \alpha^2$ and $\mathbf{R}_W = \mathbf{E}[\mathbf{W}\mathbf{W}^H]$ is the noise covariance matrix. According to (46.10), we have $\mathbf{R}_W = \sigma^2\mathbf{I}$, where $\sigma^2 = \sum_{i=1}^{N_b} |\omega(t_i)|^2$

Eigen-decompose \mathbf{R}_X , we have

$$\mathbf{R}_X = \sum_{i=1}^{N_b} \lambda_i \mathbf{u}_i \mathbf{u}_i^H \quad (46.16)$$

where, λ_i is the eigenvalue and \mathbf{u}_i is the corresponding eigenvector, where $i = 1, 2, \dots, N_b$. Then sort the eigenvalues in descending order, we have $\lambda_1 > \lambda_2 > \dots > \lambda_{N_b}$. We named the maximum eigenvalue λ_1 as the signal eigenvalue, and the subspace expanded \mathbf{u}_1 is the pulse subspace $\mathbf{U}_S = \mathbf{u}_1$. The remainder eigenvalues are corresponding to noise, where $\lambda_2 = \dots = \lambda_{N_b} = \sigma^2$, then the corresponding eigenvectors expanded the noise sunspace $\mathbf{U}_N = [\mathbf{u}_2, \dots, \mathbf{u}_{N_b}]$.

According the property between eigenvalues and eigenvectors, we have

$$\mathbf{R}_X \cdot \mathbf{u}_i = \sigma^2 \cdot \mathbf{u}_i, \quad i = 1, 2, \dots, N_b \quad (46.17)$$

Combining with (46.14), we have

$$\mathbf{A}^H(\tau_0) \cdot \mathbf{u}_i = \mathbf{0}, \quad i = 1, 2, \dots, N_b \quad (46.18)$$

According to (46.18), we know that the steering vector is orthogonal to the noise eigenvectors, while the steering vector contains the information of pulse TOA.

According to this property, the pulse TOA can be estimated by optimization search, so we can construct the searching vector as following

$$\mathbf{A}(t_i) = \begin{bmatrix} e^{-j\frac{2\pi(t_i)}{N_b}} & & & \\ & e^{-j\frac{2\pi(2t_i)}{N_b}} & & \\ & & \ddots & \\ & & & e^{-j\frac{2\pi(N_b t_i)}{N_b}} \end{bmatrix} \begin{bmatrix} |\hat{S}(1)|^2 \\ |\hat{S}(2)|^2 \\ \vdots \\ |\hat{S}(N_b)|^2 \end{bmatrix}, \quad i = 1, 2, \dots, N_b \quad (46.19)$$

According to (46.17), the searching function is denoted as

$$f(t_i) = \frac{1}{\mathbf{A}^H(t_i) \mathbf{U}_N \mathbf{U}_N^H \mathbf{A}(t_i)} \quad (46.20)$$

Then search for the maximum value of $\mathbf{f} = [f(t_1), f(t_2), \dots, f(t_{N_b})]$, and the corresponding time t_k is the estimation of pulse TOA, that is

$$\hat{\tau} = \underset{k}{\arg\{\max f(t_k)\}} \quad (46.21)$$

At last, summarized the computation process of the proposed method:

- Step 1. Calculate $R(\tau)$ according to (46.7);
- Step 2. Calculate the DFT of $R(\tau)$, then rewrite $\Re(k)$ into matrix form $\mathbf{X} = [\Re(1), \Re(2), \dots, \Re(N_b)]^T$;
- Step 3. Calculate the covariance matrix \mathbf{R}_X of \mathbf{X} ;
- Step 4. Eigen-decompose \mathbf{R}_X and construct the noise subspace \mathbf{U}_N ;
- Step 5. Construct the searching vector by (46.19);
- Step 6. Calculate the searching function by (46.15), then search for the maximum value of $\mathbf{f} = [f(t_1), f(t_2), \dots, f(t_{N_b})]$, and the corresponding time t_k is the estimation of pulse TOA.

46.4 Simulation Results

Simulation results are provided to verify the performance improvement of the proposed algorithm, which are compared with the correlation method. We choose PSR B1744-21A as the observation pulsar, the period of the pulsar is 0.0116 s, that is $P = 0.0116$ s. Its parameters of flux are: the flux of background $R_b = 5 \times 10^{-3}$ ph/s/cm²; the flux of pulsar $R_s = 1.09 \times 10^{-3}$ ph/s/cm². The two parameters are often denoted as the number of photons which the detector received in the 2 ~ 10 keV frequency range per unit area and per unit time. p_f denotes the pulse ratio, and $p_f = 0.6$.

Fig. 46.1 Standard profile and observation profile

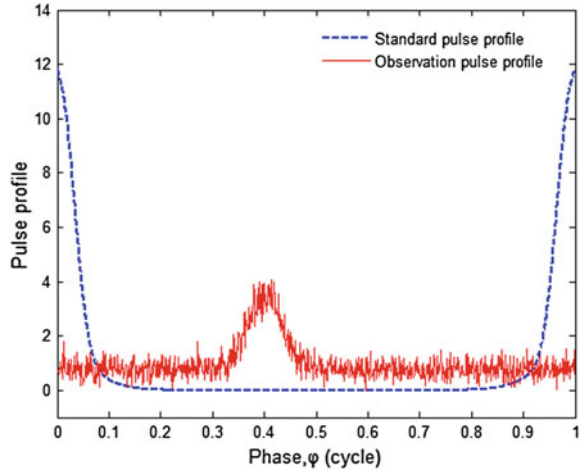


Figure 46.1 is the comparison diagram of Standard profile and observation profile by epoch folding. The imaginary line is the curve of standard profile, which is drawn by discrete data fitting. And the data is obtained from the Europe pulsar database of the Max Planck institution for Radio astronomy website. The full line is the curve of observation profile by epoch folding, and the length of the observation time is $t_{\text{obs}} = 200$ s, the area of detector is $A_d = 5$ m², the temporal resolution is $t_d = 100$ μ s, we choose the starting point of observation period $t_0 = 51545.0$ MJD, the phase position of pulse is set as $\phi_0 = 0.4$. The relationship of TOA and pulse position is equivalent, where $\tau_0 = \phi_0 \cdot P$, so we can express the pulse TOA by the pulse position.

Example 46.1 Choose PSR B1744-21A as the observation pulsar, the starting point of observation period $t_0 = 51545.0$ MJD, the length of the observation time is $t_{\text{obs}} = 2$ s, the area of detector is $A_d = 5$ m², the temporal resolution is $t_d = 100$ μ s, the phase position of pulse is set as $\phi_0 = 0.3$. Figure 46.2 is pulse TOA estimation by correction method and the proposed method respectively.

In order to compare the performance of the correction method and the propose method, normalization processing is carried out for the two curves in Fig. 46.2.

From Fig. 46.2 we know the spectrum of the proposed method is much sharp than the correction method, and the estimate value of the correction method is $\hat{\phi}_0 = 0.2974$ while the proposed method is $\hat{\phi}_0 = 0.2983$, so the estimation accuracy of the proposed method is improves compared with the correction method.

Example 46.2 In order to verify the performance of the method, conducted 300 times Monte-Carlo experiments under observation time length of $t_{\text{obs}} = 0.1$ s, 0.2 s, 0.5 s, 1 s, 2 s, 5 s, 10 s, 20 s, 50 s, respectively. And we obtained the RMSEs of the two methods in Fig. 46.3. The other parameters are setting as example 46.1.

Fig. 46.2 Normalization estimation spectrum

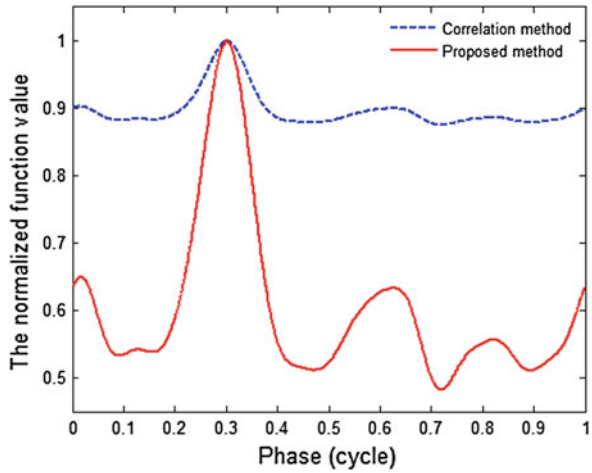
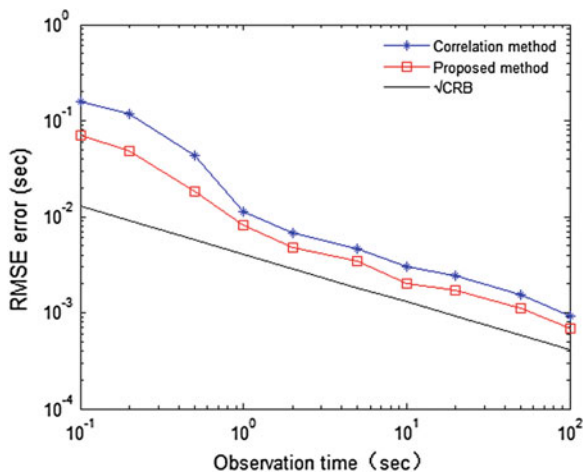


Fig. 46.3 RMSE in different observation time



From Fig. 46.3, we can know that the RMSEs of the two methods are decreased as the length of observation time increased, that is to say the estimation accuracy is increased. Also, at the same length of observation time, the accuracy of the proposed method is higher than the correction method.

46.5 Conclusion

In the XPNAV system, high precision solution of pulse TOA is of great significance. In order to improve the precision of pulse TOA estimation, a novel method is proposed. The cross-correction function of the standard pulse profile and the

observation pulse profile is analyzed in this paper, then based on the property of subspace decomposition, a high precision pulse TOA estimation method is proposed utilizing the DFT of the cross-correction function. Simulation results indicated that the spectrum of the proposed method is much sharper than the correction method, the estimation precision has been improved, and this can lay the foundation for the high precision autonomous navigation by XPNV system in the future.

References

1. Hanson JE (1996) Principles of X-ray navigation. Ph.D dissertation, Dept of Aeronautics and Astronautics, Stanford Univ, Stanford, CA
2. Emadzadeh AA, Speyer JL (2011) Navigation in space by X-ray pulsars. Springer, Berlin
3. Shuai P, Li M, Chen SL, et al (2009) Principle and method of X-ray pulsars navigation system. China Astronautic Publishing House, Beijing
4. Shuai P, Li M, Chen SL, et al (2007) Navigation principles using X-ray pulsars. *J Astronaut* 28 (6):1538–1543
5. Ashby N, Golshan AR (2008) Minimum uncertainties in position and velocity determination using X-ray photons from millisecond pulsars ION NTM 2008. San Diego, CA, pp 110–118
6. Mao Y, Song X, Chai F (2009) Analysis of pulsar time-of-arrival measurement error and dilution of precision. *J Geomatics Sci Technol* 26(2):140–143
7. Emadzadeh AA, Speyer JL (2009) Asymptotically efficient estimation of pulse time delay for X-ray pulsar based relative navigation. In: AIAA GN&C conference, Chicago, IL, pp 1–12
8. Emadzadeh AA, Speyer JL (2010) On modeling and pulse phase estimation of X-ray pulsars. *IEEE Trans Signal Process* 58(9):4484–4495
9. Emadzadeh AA, Speyer JL (2011) X-ray pulsar-based relative navigation using epoch folding. *IEEE Trans Aerosp Electron Syst* 47(4):2317–2328
10. Emadzadeh AA, Speyer JL (2009) Asymptotically efficient estimation of pulse time delay for X-ray pulsar based relative navigation. In: AIAA GN&C conference, Chicago, IL, pp 1–12
11. Golshan AR, Sheikh SI (2007) On pulse phase estimation and tracking of variable celestial X-ray sources. In: ION 63rd annual meeting, Cambridge, MA, pp 413–422
12. Rinauro S, Colonnese S, Scarano G (2013) Fast near-maximum likelihood phase estimation of X-ray pulsars. *Sig Process* 93(1):326–331
13. Zhenhua X, Luping X, Guangren N (2008) Time offset measurement algorithm based on bispectrum for pulsar integrated pulse profile. *J Phys* 57(10):6683–6688
14. Su Z, Xu L, Wang T (2011) A new time delay measurement algorithm for pulsar accumulated pulse profile. *J Astronaut* 32(6):1256–1261
15. Hanson JE, Sheikh SI, Graven P et al (2008) Noise analysis for X-ray navigation systems. In: 2008 IEEE/ION position location and navigation symposium, Monterey, CA, pp 101–110

Chapter 47

A Research of Code Tracking Loop for Navigation Signal Based on DS/FH Modulation

Zhuxi Yu, Jiaolong Wei, ZuPing Tang, Zhihui Zhou and Yuan Xue

Abstract Facing the increasingly complex electromagnetic environment, it is an attempt to introduce DS/FH hybrid spread spectrum technology to the field of TT&C. For DS/FH receiver, despreading and dehoppping proceed respectively. However if DS/FH navigation signal is dealt with in the same way, there is no difference between DS/FH navigation signal and the existing DS signal. The code tracking algorithm proposed in this paper makes full use of frequency hopping information so that the correlation peak becomes very sharp. Computer simulation shows that the code tracking accuracy of DS/FH navigation signal reaches at 0.0028 m (CNR = 40 dB Hz) and the maximum of the multipath error envelope is 0.062 m (MDR = -10 dB), far better than the performance of existing navigation signals.

Keywords Hybrid spread spectrum · Code tracking · Self-correlation · High precision · Anti-multipath performance

47.1 Introduction

DS/FH is widely used in the field of communication system for its outstanding anti-jamming performance, but its application in TT&C is rarely mentioned. The biggest difference of DS/FH's application in the field of communication and TT&C is the high-dynamic between the receiver and the transmitter and the extremely low SNR at the receiver.

Compared with the existing TT&C system used direct spread spectrum technology, the application of DS/FH signal will bring many new difficulties, such as: Doppler frequency hopping brought by carrier hopping, the discontinuity of carrier's initial phase and so on. These differences and difficulties limit the application

Z. Yu (✉) · J. Wei · Z. Tang · Z. Zhou · Y. Xue
School of Electronic Information and Communications, Huazhong University
of Science and Technology, Wuhan 430074, China
e-mail: 653579734@qq.com

of DS/FH in TT&C. Article [1] proposed using the assist of the hopping pattern to reduce the influence of frequency hopping on carrier tracking accuracy and present an approach to decrease the related accumulated energy loss caused by the discontinuity of carrier's initial phase. However, it is difficult to find code tracking algorithms of DS/FH from the existing publicly information. Based on this case, this paper focus on the discussion of the DS/FH code tracking algorithm, as the carrier tracking are involved in many documents [2–5], not tired in words here.

Under the premise of carrier synchronization, if the code tracking algorithm of DS/FH is the same with the traditional code tracking algorithm in TT&C, then the cross-correlation function of DS/FH system and traditional TT&C system will be the same which can not reflect the superiority of the DS/FH signal. In this paper, hopping information is used to assist code tracking. This code tracking algorithm can not only track DS/FH signal steady, but also can greatly improve the positioning accuracy of DS/FH navigation signal and anti-multipath performance.

47.2 Code Tracking Algorithm of DS/FH Signal

The code tracking loop in this article is still using delay locked loop (DLL), similar to the traditional DS system [6]. But in order to make better use of the frequency hopping information and to improve the tracking performance, when advanced code and delay code are generated in the locality, the carrier in the locality also generates a corresponding time delay, the code tracking loop is as shown in Fig. 47.1. Since the introduction of the local carrier delay, it brings the cross-correlation function with a factor which will make the correlation peak very sharp.

Assume that the received DS/FH signal is:

$$s(t) = \sum_k aD(t)c(t)e^{jw_k t} \quad (47.1)$$

Here, k represents the number of hopping points within a frequency hopping period, a represents the amplitude of signal, $D(t)$ represents the data bits, $c(t)$ represents the pseudo-random code, w_k represent the angular frequency of the k -hop.

Assume that the delay of pseudo-random code in the locality is τ at the hopping period of the k -hop, then the carrier in the locality is $e^{-jw_k(t-\tau)}$. The analysis of the signal's performance is as follows:

As shown in Fig. 47.2, for a DS/FH system with n hops, f_1, f_2, \dots, f_n are center frequencies after down converted. The frequency hopping period is T . In order to reach the best code tracking performance, the coherent integration time is set as $T_p = nT$. Based on the above assumptions, the output of the correlator is:

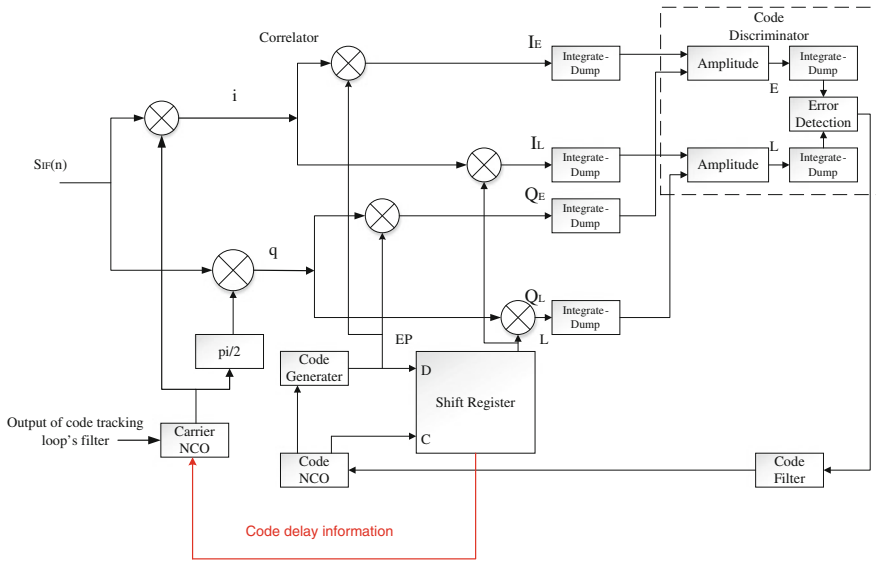


Fig. 47.1 Code tracking loop

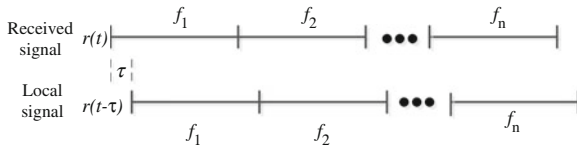


Fig. 47.2 Sketch map of cross-correlation

$$R_r(\tau) = E[r(t) \cdot r'(t - \tau)] = \frac{1}{n} \sum_{i=1}^n R_i(\tau) \tag{47.2}$$

Here, $R_i(\tau)$, $i = 1, 2, \dots, n$ represents the cross-correlation of the i -hop in Fig. 47.2.

For each piece of data:

$$\begin{aligned} R_i(\tau) &= E(r_i(t) \cdot r'_i(t - \tau)) \\ &= R(\tau) e^{j2\pi f_i t} \cdot e^{j2\pi f_i (\tau - t)} \\ &= R(\tau) e^{j2\pi f_i \tau} \quad i = 1, 2, \dots, n \end{aligned} \tag{47.3}$$

Here, $R(\tau)$ represents cross-correlation function.

Put Eq. (47.3) into (47.2):

$$R_r(\tau) = R(\tau) \sum_{i=1}^n e^{j2\pi f_i \tau} \quad (47.4)$$

Generally, frequency hopping system has equal interval between center frequencies, so there must be an arrangement (f_1, f_2, \dots, f_n) that makes it an arithmetic sequence. Without loss of generality, assume that $f_n > f_{n-1} > \dots > f_2 > f_1 = 0$ and f_1, f_2, \dots, f_n constitute an arithmetic sequence with a tolerance of Δf . Equation (47.4) can be converted into:

$$\begin{aligned} R_r(\tau) &= R(\tau) \sum_{i=1}^n e^{j2\pi f_i \tau} \\ &= R(\tau) \left(1 + e^{j2\pi \Delta f \tau} + e^{j2\pi (2\Delta f) \tau} + \dots + e^{j(2\pi(n-1)\Delta f) \tau} \right) \\ &= R(\tau) \left(\frac{1 - e^{j2\pi(n\Delta f)\tau}}{1 - e^{j2\pi\tau}} \right) \end{aligned} \quad (47.5)$$

According to Eq. (47.5), when $(n\Delta f)\tau$ is an integer, $R_r(\tau) = 0$, which means the cross-correlation function of DS/FH signal has several zero points. In particular, the closest zero point to the original point is $\tau_0 = \frac{1}{n\Delta f}$, $\tau_{0,c} = \frac{R_c}{n\Delta f}$ in chip unit (R_c is the rate of pseudo-random code). Therefore, when $n\Delta f \gg R_c$ (this condition is usually easily satisfied), $\tau_{0,c} \ll 1$, which means the main peak of $R_r(\tau)$ is more narrow than $R(\tau)$. In this case, the DS/FH signal has a higher tracking accuracy and a better anti-multipath performance than the DS signal.

Assume that $n = 31$, $\Delta f = 20.46$ MHz, $R_c = 10.23$ MHz, $R_r(\tau)$ is as shown in Fig. 47.3. As can be seen from the Fig. 47.3, $R_r(\tau)$ has a series of sub-peak similar to the BOC signal, but the amplitude of sub-peak is much smaller than BOC signal with the same order. Its first zero point is:

$$\tau_0 = \frac{R_c}{n\Delta f} = \frac{10.23 \times 10^6}{31 \times 20.46 \times 10^6} = \frac{1}{62} = 0.016 \text{ (chip)}$$

Discriminator curve of DS/FH signal is as shown in Fig. 47.4. It can be seen from the curve that the DS/FH signal has a high discriminator gain under this code tracking algorithm. However, the presence of multiple zero points also makes the tracking procedure fuzzy.

The code tracking accuracy and anti-multipath performance of navigation signal is related to its cross-correlation function. The narrower the main peak of the cross-correlation function is, the better code tracking accuracy and anti-multipath performance will be. Therefore, the code tracking algorithms proposed by this paper can offer better performance in code tracking accuracy and anti-multipath performance.

Fig. 47.3 Cross-correlation function of DS/FH signal

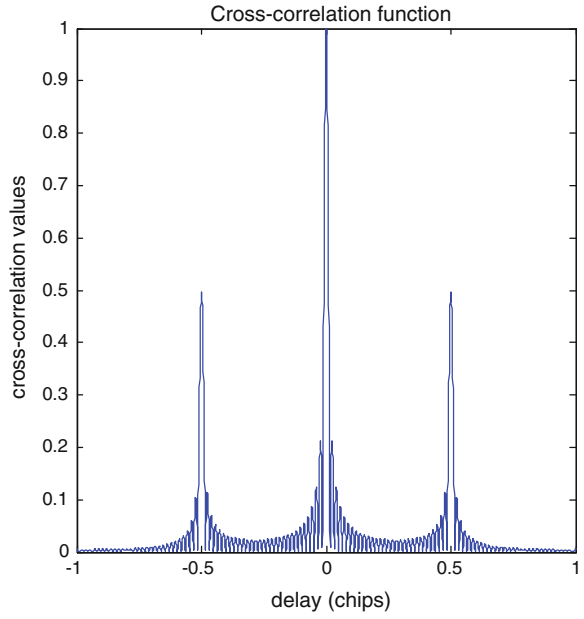


Fig. 47.4 Discriminator curve of DS/FH signal

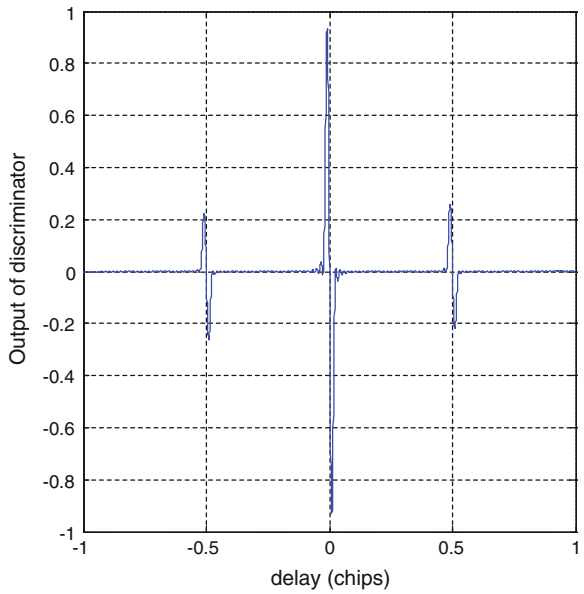


Table 47.1 Parameters of receiver (code tracking experiment)

Parameters	DS/FH	BPSK (10)
Sampling rate	50 MHz	
Frequency hopping period	31 Hops	–
Frequency hopping rate	1000 Hops/s	–
Bandwidth of frequency hopping channel	20.46 MHz	–
Simulation time	110 s	
CNR	40 dB Hz	
Bandwidth of code tracking loop	1 Hz	
Interval of correlator (single side)	0.01 T _c (1/10.23e6 s)	
Coherent integration time	31 ms	

47.3 Simulation

In order to verify the performance of the code tracking algorithm, this paper makes several experiments about DS/FH signal and compares it with the traditional DS system. For better comparison, DS signal uses BPSK (10).

47.3.1 Code Tracking Experiment

When conducting code tracking experiment, the parameters of receiver is as shown in Table 47.1.

According to the parameters of Table 47.1, this paper makes code tracking experiment about DS/FH signal and traditional BPSK (10) signal. The output of code tracking loop's filter is as shown in Fig. 47.5, it can be seen from the figure that the code tracking algorithm proposed in this paper can track DS/FH signal steady.

Under these experimental conditions, you can get the code tracking accuracy of DS/FH signal is 0.0028 m, code tracking accuracy of BPSK (10) signal is 0.1012 m. Obviously, the code tracking algorithm of this paper enhances code tracking accuracy by two orders of magnitude.

47.3.2 Anti-multipath Experiment

When conducting anti-multipath experiment, the parameters of receiver is as shown in Table 47.2

According to the parameters of Table 47.2, this paper makes anti-multipath experiment about DS/FH signal and traditional BPSK (10) signal. The multipath

Fig. 47.5 Output of code tracking loop's filter (DS/FH signal)

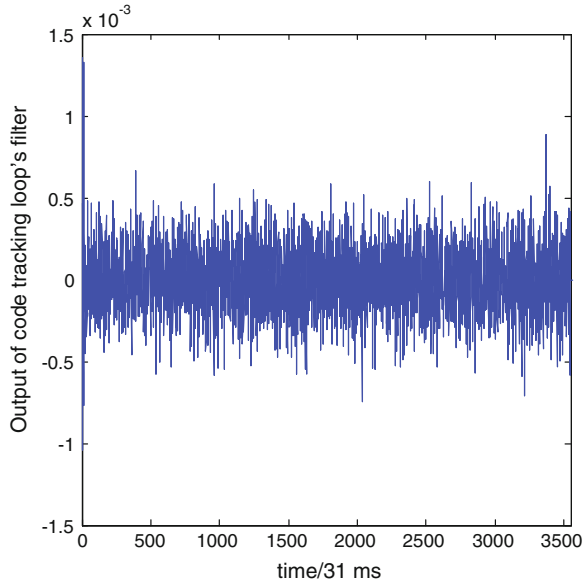


Table 47.2 Parameters of receiver (Anti-multipath experiment)

Parameters	DS/FH	BPSK (10)
Sampling rate	50 MHz	
Frequency hopping period	31 Hops	–
Frequency hopping rate	1000 Hops/s	–
Bandwidth of frequency hopping channel	20.46 MHz	–
Simulation time	110 s	
CNR	40 dB Hz	
Bandwidth of code tracking loop	1 Hz	
Interval of correlator (single side)	0.025 T _c (1/10.23e6 s)	
Coherent integration time	31 ms	
MDR	-10 dB	

error envelope of DS/FH signal and traditional BPSK (10) signal is as shown in Figs. 47.6 and 47.7.

It can be seen from Figs. 47.6 and 47.7 that the code tracking algorithm of this paper has a very advantageous anti-multipath performance. When MDR is -10 dB, the maximum value of multipath error envelope is 0.062 m, far better than the performance of existing navigation signals.

Fig. 47.6 Multipath error envelope of DS/FH signal

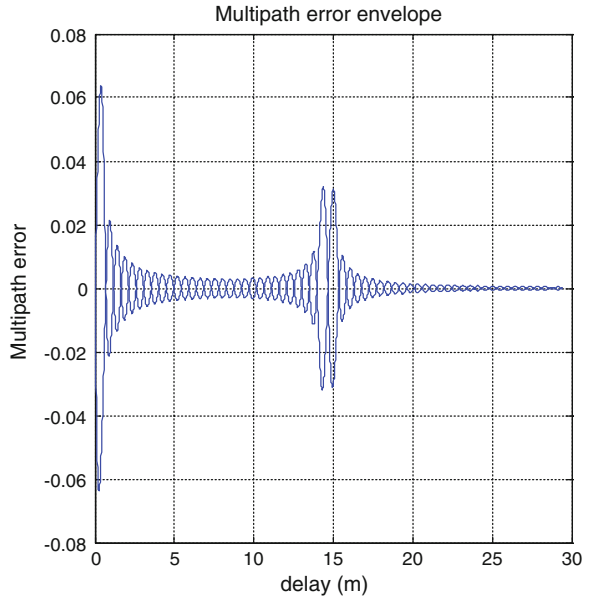
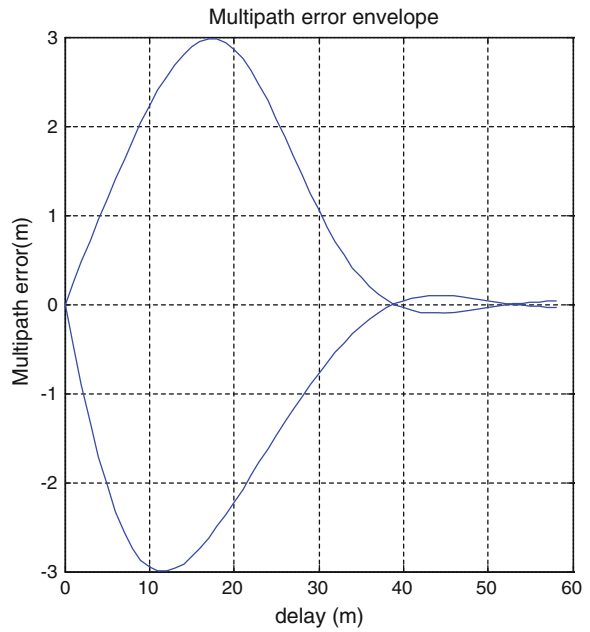


Fig. 47.7 Multipath error envelope of BPSK (10)



47.4 Conclusion

In order to improve code tracking accuracy of DS/FH signal, this paper proposes to use frequency hopping information to assist code tracking. Through simulation, we compare the performance of DS/FH signal with traditional DS signal. It can be verified that the code tracking algorithm proposed in this paper can not only maintain the stability of code tracking loop of DS/FH signal, but also shows great advantages on code tracking accuracy and anti-multipath performance. The code tracking algorithm proposed in this paper can provide some references and guidances in the field of TT&C.

References

1. Wei Z, Xiaolin Z (2012) Impact analysis on frequency hopping to carrier tracking precision in DS/FH hybrid spread spectrum receiver. *J Beijing Unive Aeronaut Astronaut* 38(10):1–5
2. Chen T, Wang J, Fan X (2013) A research of carrier tracking loop for high hopping rate DS/FH receiver. In: *Signal processing, communication and computing*, 5–8 Aug 2013
3. Chen Z, Wang Y, Yang W (2010) Research on carrier tracking in Hybrid DS/FH spread spectrum TT&C system. In: *Future computer and communication*
4. Meng S, Yang W, Chen Z (2010) Performance simulation analysis on Doppler frequency tracking of DS/FH signal for TT&C. In: *China satellite navigation conference*, 19–23 May 2010
5. Guangxuan Y, Xiantai G (2010) Study on rapid synchronization techniques of DS/FH Hybrid spread spectrum system. *Commun Technol* 43(6):26–28
6. Elliott DK, Christopher JH (2012) *Understanding GPS: principles and applications*, vol 8, Publishing House Of Electronic Industry, Beijing

Chapter 48

Method to Construct Database of X-ray Pulsar-Based Navigation Using Space-Based Observation Data

Xiao-long Hao, Qiang-wen Yang and Liang-wei Huang

Abstract In consideration of advancing the technology of X-ray pulsar-based navigation, a theoretical method for space-based building of navigation database is raised. This method uses locus integration of the Earth orbit to provide the analytical solution to measured parameters of navigation sources. Numerical analyses reveals that to meet the demands of navigation with hundreds of meters' accuracy, the testing database containing four pulsars can be built within two years with a 0.25 m^2 detector. For an extended database, it is recommended that X-ray binaries are included. The space-based observation data of X-ray binaries shall be collected before judging whether they will serve as navigation sources

Keywords X-ray pulsar-based navigation · Testing database · Space-based observation · Timing residual · X-ray binary

48.1 Introduction

The X-ray pulsar navigation technology is the most promising breakthrough technology which can achieve long-precision autonomous navigation, and has important application value and strategic significance, which has attracted international space agencies concerned [1]. In 2004, the US Department of Defense proposed X-ray pulsar navigation Research Program (XNAV), has completed the feasibility demonstration of key technologies and ground verification, the space flight experiments on the International Space Station and high-orbit satellites is upcoming. In addition, the European Space Agency (ESA), Russia, Germany, Japan, the United Kingdom, India, Australia and other countries or organizations have also launched the study on X-ray pulsar navigation theoretical methods and experimental validation.

X. Hao (✉) · Q. Yang · L. Huang
Beijing Tracking and Communication Technology Research Institute, Beijing 100094, China
e-mail: haoxl@beidou.gov.cn

Build the X-ray pulsar database is the fundamental to realize the pulsar navigation, which including X-ray pulsar preferred, X-ray pulsar catalog and navigation database design [1]. Considering the X-ray pulsar signals can't pass through the dense atmosphere, a massive pulsar data were already get using the ground large diameter radio telescope observations, so the navigation database Construction is main based on the ground data, and the space data is auxiliary [2]: Firstly, the precise angular position of the pulsar (ascension and declination) and timing model parameters was obtained through the ground long-term observational data; then the X-ray profile of the pulsar was obtained by space-based observations; and the suitable autonomous navigation database was constructed in the end.

The space-based X-ray pulsar observation has the unique advantage, which the ground radio pulsar observations cannot be compared. For example, Vela pulsar in the radio flux at 1.4 GHz for 1100 mJY, corresponding space-based X-ray detector with effective area of 2500 cm² (diameter of 28.2 cm), which is equivalent to diameter of 3.7 km of ground radio telescopes. So, in order to accelerate the development of X-ray pulsar navigation technology, this paper proposed to using the space-based observation data to construct the X-ray navigation pulsar database, and verifying the feasibility of the method through numerical calculation.

48.2 Space Observation and Pulsar Parameters Determination

48.2.1 Problems Description

In the process of constructing a database, there are two kinds of parameters need to observed: one is the angular position parameters, namely pulsar ascension and declination; the other is timing model parameters, namely pulsar rotation frequency and its first and second derivative [1]. Similar to the single ground station observations, this paper carried out only for one space-based observation satellite; for multiple satellites constitute VLBI network can be precisely measured the angular position of the pulsar parameters [3], will discuss in the future. The pulsar parameters observations utilizing a single satellite are obtained by accumulating the pulse arriving time using a least squares fit of the estimated. Using through binary pulsar orbit integration method can analyze the measurement accuracy of binary parameters [4], but has not yet considered the periodic motion of the detector;

Using the numerical integration along the satellite orbit can analyze the track accuracy [5], but has not yet estimate pulsar parameters accuracy. Therefore, this article using the integral method along orbits to derive the mathematical expression of pulsar parameters; then calculating the time requirement to build the pulsar navigation database at a given timing observation accuracy.

In general, the timing residuals is defined as the pulse arriving time difference between the forecasting by the fitted timing model and the actual measurement by

the detector. The timing residuals can characterize the influence of all parameter errors to the time delay calculation, which can also describe the accuracy of timing model parameter. The main determinant of the timing residual is the angular position parameter of the pulsar, which means the angular position is the most considerable measure parameter. The mathematical model of pulsar measured parameters is the expression of angular position estimation error and timing residual, with respect to observation time, the mathematical analytical expressions can be obtain through the integration in earth orbits around the sun, then the pulsar observation accuracy can be predicted according to the design life of the satellite. That is, if the requirement of the orbit accuracy is hundred meters, the corresponding timing residual of pulsar is in the order of microseconds. In this paper, the timing residuals is set as $1 \mu\text{s}$, with the given effective area of the detector and original parameter of pulsar, the pulsar observation time in orbit can be calculated.

48.2.2 Space Observation Mode Analysis

If the detector static or linear motion, the observation equation is linear correlated, even if cumulating long observational data, it does not have the observability. The observability depends on the periodic motion of the detector, as the detector follow the earth around the sun, to achieve the best observing effect, the observing data needs evenly distributed in different phase in the orbit.

This paper presents two modes of space observation: the continuous observation mode and the gapped observation mode as shown in Fig. 48.1. The continuous

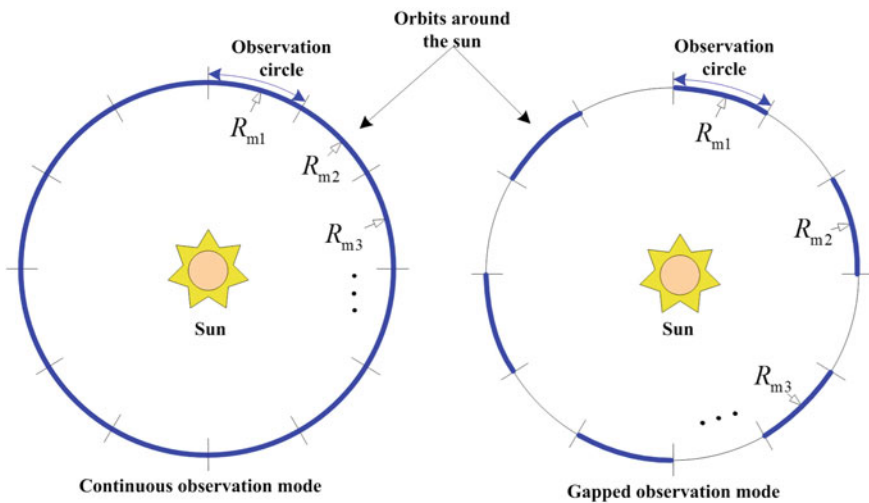


Fig. 48.1 Schematic diagram of the continuous observation mode and the gapped observation mode

observation mode refers to a continuous observation period followed by another period of observation; the gapped observation mode refers to a gap between the observation period with a certain intervals, each interval of equal length, and the observation data distributed throughout the orbit to achieve optimal measurement results. In the pulsar parameter determination algorithm, for both the observed modes, can use the approximate summation along the orbits. In this paper, the cumulative observation time is the actual observation time minus the interval, as shown in Fig. 48.1 in bold, denote t_Y .

48.2.3 Pulsar Angular Position and Timing Model Parameters Determination Algorithm

Set the time of arrival (TOA) of the pulse measured by the detector is t_{aOBS} , arriving at the Solar System Barycenter (SSB) is t_{aSSB} , the timing observation parameter is defined as $R_m = -c(t_{\text{aOBS}} - t_{\text{aSSB}})$, where c is the speed of light in vacuum. Regarded the high-order time delay as the noise, all the parameters are known except the angular position, the pulsar timing observation model is:

$$R_m = x \cos \delta \cos \alpha + y \cos \delta \sin \alpha + z \sin \delta \quad (48.1)$$

where x , y and z are the position coordinate components of the space-based detector in Barycentric Celestial Reference System (BCRS); α and δ denote the ascension and declination of pulsar in BCRS respectively. Given the original parameter of the pulsar angular position, using Gauss–Newton nonlinear least squares method, iterative estimating ascension and declination deviation (denoted as $\Delta\delta$, $\Delta\alpha$) to obtain refined angular position parameters [6]. The time residuals is $\Delta t = \Delta R_m/c$, then

$$\Delta R_m = \frac{\partial R}{\partial \delta} \Delta\delta + \frac{\partial R}{\partial \alpha} \Delta\alpha \quad (48.2)$$

where

$$\begin{aligned} \frac{\partial R_m}{\partial \delta} &= -x \cos \alpha \sin \delta - y \sin \alpha \sin \delta + z \cos \delta \\ \frac{\partial R_m}{\partial \alpha} &= -x \cos \delta \sin \alpha + y \cos \delta \cos \alpha \end{aligned} \quad (48.3)$$

Set the total observed data points is N , there

$$\Delta \mathbf{R}_m = \mathbf{A}[\Delta\delta, \Delta\alpha]^T \quad (48.4)$$

where, $\Delta \mathbf{R}_m = [\Delta R_{m1}, \Delta R_{m2}, \dots, \Delta R_{mN}]^T$

$$\mathbf{A} = \begin{bmatrix} (\partial R_m / \partial \delta)_1, & (\partial R_m / \partial \alpha)_1 \\ (\partial R_m / \partial \delta)_2, & (\partial R_m / \partial \alpha)_2 \\ \vdots & \\ (\partial R_m / \partial \delta)_N, & (\partial R_m / \partial \alpha)_N \end{bmatrix} \quad (48.5)$$

Solving variance of Eq. (48.4), assuming all observation points equal weights, as have the same variance $\varepsilon^2 = \text{var}(R_{mi})$, the estimated angular position covariance matrix is:

$$\text{var}\left(\left[\hat{\delta}, \hat{\alpha}\right]^T\right) = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T (\varepsilon^2 \mathbf{I}) \left[(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T\right]^T = (\mathbf{A}^T \mathbf{A})^{-1} \varepsilon^2 \quad (48.6)$$

The formula (48.6) can be rewritten as:

$$\text{var}\left(\left[\hat{\delta}, \hat{\alpha}\right]^T\right) \varepsilon^2 \left(\sum_{i=1}^N \mathbf{M}_i\right)^{-1} \quad (48.7)$$

where

$$\mathbf{M}_i = \begin{bmatrix} \left(\frac{\partial R_m}{\partial \delta}\right)^2 & \frac{\partial R_m}{\partial \delta} \frac{\partial R_m}{\partial \alpha} \\ \frac{\partial R_m}{\partial \delta} \frac{\partial R_m}{\partial \alpha} & \left(\frac{\partial R_m}{\partial \alpha}\right)^2 \end{bmatrix}_i \quad (48.8)$$

As the accumulated time of the pulsar t_Y more than 1 year, the continuous observation mode can be used; if t_Y is less than 1 year, only the gapped observation mode can be used. $\sum \mathbf{M}_i$ is the summation of formula (48.7) obtained through integration along the earth orbits around the sun. As the distance between detector and the earth can be ignored, $[x, y, z]^T = [x_E, y_E, z_E]^T$, $\partial R_m / \partial \delta$ and $\partial R_m / \partial \alpha$ can be expressed by the Earth's orbital parameters, Arguments is the true anomaly earth revolution orbit f_E . Approximated by summing integration along the track:

$$\sum_{i=1}^N \mathbf{M}_i = \frac{N}{2\pi} \int_0^{2\pi} \mathbf{M}(f_E) df_E \quad (48.9)$$

The single point observation accuracy ε^2 in formula (48.6) can be calculated by the TOA error equation given by Sheikh [7]:

$$\sigma_{\text{TOA}}^{\text{Sheikh}} = \frac{W \sqrt{[R_b + R_s(1 - p_f)](A_d \tau_{\text{obs}} W f_0) + R_s A_d p_f t_{\text{obs}}}}{2R_s A_d p_f t_{\text{obs}}} \quad (48.10)$$

where, W is the full width at half maximum (HWHM) of the pulse, R_b and R_s are the source of background traffic and traffic from the pulsar respectively, a general statistics is the number of photons per unit area per unit time within 2–10 keV band received; A_d is the effective detection area; p_f is the pulse proportion; t_{obs} is the observation period. The observation points in Eq. (48.7) is $N = t_Y/\tau_{\text{obs}}$, considering the high-order time delay error, set $\varepsilon^2 = 100c^2\sigma_{\text{TOA}}^{\text{Sheikh}}$, Eq. (48.7) can be rewritten as:

$$\text{var}([\hat{\delta}, \hat{\alpha}]^T) = \frac{50\pi c^2 W^2 ([R_b + R_s(1 - p_f)](Wf_0) + R_s p_f)}{(R_s p_f)^2 A_d t_Y} \left(\int_0^{2\pi} \mathbf{M}(f_E) df_E \right)^{-1} \quad (48.11)$$

It shows that there is no direct relation between the estimated variance of the angular position and the observation points and the observation period τ_{obs} , which depends on the cumulative time of observation and the Earth motion around the sun. Equation (48.11) can be rewritten as:

$$\begin{aligned} \text{var}(\Delta t) = & c^{-2} \left[(\partial R_m / \partial \delta)^2 \text{var}(\hat{\delta}) + (m / \partial \alpha)^2 \text{var}(\hat{\alpha}) \right. \\ & \left. + 2(\partial R_m / \partial \delta)(\partial R_m / \partial \alpha) \text{cov}(\hat{\delta}, \hat{\alpha}) \right] \end{aligned} \quad (48.12)$$

Equations (48.11) and (48.12) is the pulsar parameters accuracy analysis model. Given the cumulative observation time, the angular position of the pulsar and the accuracy of timing residuals can be calculated; on the contrary, set the timing residuals, the cumulative time required for observation can be resolved.

48.3 Preferred Pulsar and Cataloging Method in Construct Database

48.3.1 The Preferred Method of Pulsar Sources

The pulsar emits radio, infrared, visible light, ultraviolet and X-rays signals; the dynamic model, pulse profile, pulse period and variance ratio, and the radiation flow intensity are different. It's necessary to preferred pulsar and construct the X-ray pulsar navigation database which suitable for actual implementation. Navigation pulsar preferred criteria are divided into qualitative and quantitative areas [1]:

- (1) Qualitative criteria: include precise angular position of pulsars, high signal to noise ratio of the pulse profile, high-precision pulse timing model, a higher flow rate of X-ray radiation, short pulse period, a sharp pulse shape, and long-term stability of the pulse cycle. From the pulsar formation and radiation

mechanism, the isolated rotation-powered pulsars has relatively stable pulse period and radiation flow, can meet the spacecraft autonomous navigation applications, so the preferred pulsar should be the isolated rotation-powered pulsars.

- (2) Quantitative criteria: the quality factor can be used to evaluate and calibrate the X-ray pulse signal quality. As given the effective area of the detector, the signal integration time and the X-ray background radiation flow rate, the measurement accuracy of the pulse arrival time depends on the pulse signal quality and pulse contour shape. According to the measured variance analysis of TOA based on SNR, the pulsar quality factor is expressed as:

$$Q_x = \frac{R_s p_f^2}{W^2 [p_f + W f_0 (1 - p_f)]} \quad (48.13)$$

where the symbols definition is the same with Eq. (48.10).

48.3.2 Cataloging Candidate Pulsar Sources

The X-ray pulsar cataloging is to analyze the characteristic parameters of X-ray pulsars which have been detected already, then sorting, classification, and building a layered structure of the pulsar data table for updating data, access, query and calls. Considering the pulsar catalog structure, there can be three basic directories: statistical directory, detailed catalog and candidate list [1].

Currently, there are more than 2000 pulsars have been found and cataloged, which constitute the statistical directory of pulsar. According to the preferred criteria of both qualitative and quantitative methods, 55 pulsars was selected to constitute detailed catalog. Among them, there are 24 isolated rotation-powered pulsars (IRPSR), 6 binary rotation-powered pulsars (BRPSR) and 25 X-ray binaries (XB). Tables 48.1, 48.2 and 48.3 show the detail parameter of the pulsar [7–10], where the value of p_f and W is taken from the predicted value of reference [7] or Australia Telescope National Center (ATNF) pulse width data. The angular position of BRPSR can reference the ATNF catalog and the angular position of XB can derived from the PSR J which given in reference [7].

48.4 Results Analysis

48.4.1 Numerical Results

Set the background noise value R_b is 5×10^{-3} ph/cm²/s, the effective area of the space detector A_d is 0.1, 0.25, 0.5 and 1 m² respectively, as the timing residuals is 1 μ s, the total observation time t_Y can be calculated. Table 48.4 lists the top 30 pulsars with t_Y values in ascending order. Note the effect of the Earth around the

Table 48.1 Source parameters of isolated rotation-powered pulsars

No.	PSR	f_0^{-1}/s	p_f	$W(s)$	$R_s/(ph/s/cm^2)$
1	B0531+21	0.033085	0.70	0.001670	1.54×10^0
2	B1937+21	0.001558	0.86	0.000021	4.99×10^{-5}
3	B1821-24	0.003054	0.98	0.000055	1.93×10^{-4}
4	B0540-69	0.050499	0.67	0.002500	5.15×10^{-3}
5	B1823-13	0.101487	0.60	0.005800	2.63×10^{-3}
6	B1509-58	0.150658	0.65	0.016000	1.62×10^{-2}
7	J1124-5916	0.135310	0.10	0.002500	1.70×10^{-3}
8	J1846-0258	0.325684	0.10	0.005900	6.03×10^{-3}
9	J0205+6449	0.065686	0.10	0.003300	2.32×10^{-3}
10	J1811-1925	0.064670	0.10	0.003200	1.90×10^{-3}
11	J1617-5055	0.069340	0.10	0.003500	1.37×10^{-3}
12	B0833-45	0.089290	0.10	0.004500	1.59×10^{-3}
13	B1951+32	0.039500	0.10	0.002000	3.15×10^{-4}
14	J0030+0451	0.004870	0.10	0.000240	1.96×10^{-5}
15	J1024-0719	0.005160	0.10	0.000590	1.37×10^{-6}
16	B0355+54	0.156382	0.10	0.003900	1.79×10^{-5}
17	B1920+10	0.226518	0.10	0.007400	4.30×10^{-5}
18	J2124-3358	0.004931	0.10	0.000510	1.28×10^{-5}
19	B0656+14	0.384891	0.10	0.018400	3.17×10^{-5}
20	J2229+6114	0.051624	0.10	0.004000	2.01×10^{-4}
21	J0537-6910	0.01611	0.10	0.000810	7.93×10^{-5}
22	J1420-6048	0.06818	0.10	0.003400	7.26×10^{-4}
23	B1706-44	0.10245	0.10	0.001900	1.59×10^{-4}
24	J1930+1852	0.13686	0.27	0.002500	2.16×10^{-4}

Table 48.2 Source parameters of binary rotation-powered pulsars

No.	PSR	f_0^{-1}/s	p_f	W/s	$R_s/(ph/s/cm^2)$
1	B1957+20	0.001610	0.60	0.000035	8.31×10^{-5}
2	J0218+4232	0.002323	0.73	0.000350	6.65×10^{-5}
3	J0437-4715	0.057570	0.28	0.000969	6.65×10^{-5}
4	J0751+1807	0.003480	0.70	0.000700	6.63×10^{-6}
5	J1012+5307	0.005260	0.75	0.000690	1.93×10^{-6}
6	B1259-63	0.047760	0.10	0.002400	5.10×10^{-4}

sun, for a total observation time t_Y which less than 1 year, the space-based observation period is consider as 1 year. See from Table 48.4:

- (1) If $A_d = 0.5 m^2$, there are 19 pulsars can achieve the timing residuals of $1 \mu s$ within 1 year observation, the PSR B1821-24 and PSR B1937+21 have better timing observation accuracy in both radio and X-ray frequency band, there

Table 48.3 Source parameters of X-ray binaries

No.	Name	f_0^{-1}/s	p_f	W/s	$R_s/(\text{ph/s/cm}^2)$
1	XTE 1751-305	0.0023	0.06	0.00046	1.81×10^{-1}
2	SAX J1808.4-3658	0.0025	0.04	0.00050	3.29×10^{-1}
3	B1728-337	0.0028	0.10	0.00055	4.49×10^{-1}
4	B1758-250	0.0030	0.10	0.00061	3.74×10^0
5	B0614+091	0.0031	0.10	0.00061	1.50×10^{-1}
6	XTE J1814-338	0.0032	0.12	0.00064	3.88×10^{-2}
7	B1617-155	0.0032	0.10	0.00065	4.19×10^1
8	B1813-140	0.0033	0.10	0.00065	2.10×10^0
9	B1636-536	0.0034	0.10	0.00069	6.58×10^{-1}
10	B1820-303	0.0036	0.10	0.00073	7.48×10^{-1}
11	B1908+005	0.0036	0.10	0.00073	2.99×10^{-4}
12	B1731-260	0.0038	0.10	0.00076	2.99×10^{-2}
13	XTE J1807-294	0.0053	0.08	0.00150	1.18×10^{-1}
14	XTE J0929-314	0.0054	0.05	0.00110	1.05×10^{-2}
15	J0635+0533	0.0338	0.10	0.00680	1.70×10^{-3}
16	1E 1024.0-5732	0.0610	0.10	0.01200	1.70×10^{-3}
17	AO 0538-66	0.0692	0.10	0.01400	4.27×10^{-1}
18	GRO J1744-28	0.4670	0.10	0.00850	3.80×10^1
19	B0115-737	0.7160	0.10	0.01300	1.50×10^{-3}
20	B1656+354	1.2400	0.100	0.02300	4.49×10^{-2}
21	GRO J1750-27	4.4500	0.10	0.08100	8.08×10^{-2}
22	B1119-603	4.8179	0.10	0.08800	2.99×10^{-2}
23	B1627-673	7.7000	0.10	0.14000	7.48×10^{-2}
24	B1744-24A	0.0116	0.60	0.00093	1.10×10^{-3}
25	GRO J1948+32	18.700	0.10	0.34000	7.31×10^{-1}

timing residuals are 0.24 and 0.02 μs in radio frequency respectively [11], which can be used as the preferred pulsar sources to build a database.

- (2) Some pulsars have better timing observation accuracy in radio frequency band, but not in X-ray frequency, such as PSR J0437-4715, its timing residuals are 0.03 μs [11] in radio but achieve 1 μs timing residuals in X-ray frequency band needs 90 years observation with $A_d = 0.5 \text{ m}^2$, which also difficulty to measure its pulse profile, original phase in X-ray frequency band, so this type can't consider as the preferred pulsars.
- (3) There are only 4 isolated rotation-powered pulsars in the top 20 pulsars: PSR B0531+21, PSR B1821-24, PSR B1937+21 and PSR B0540-69; only one binary rotation-powered pulsars PSR B1957+20, the others are X-ray binaries.

From the number of candidates, it's apparently that only choose the isolated rotation-powered pulsars is not enough to build the database, X-ray binaries can be considered as an effective supplementation. Considering the lack of radio

Table 48.4 Required accumulated observation time t_Y for the timing residual of 1 μs (in years)

NO.	Name	Type	$A_d = 0.1 \text{ m}^2$	$A_d = 0.25 \text{ m}^2$	$A_d = 0.5 \text{ m}^2$	$A_d = 1 \text{ m}^2$
1	B1617-155	XB	0.00043	0.00017	0.00009	0.00004
2	<i>B0531+21</i>	<i>IRPSR</i>	<i>0.00405</i>	<i>0.00162</i>	<i>0.00081</i>	<i>0.00041</i>
3	B1758-250	XB	0.00429	0.00172	0.00086	0.00043
4	B1813-140	XB	0.00862	0.00345	0.00172	0.00086
5	B1728-337	XB	0.02908	0.01163	0.00582	0.00291
6	B1820-303	XB	0.03083	0.01233	0.00617	0.00308
7	B1636-536	XB	0.03122	0.01249	0.00624	0.00312
8	GRO J1744-28	XB	0.03392	0.01357	0.00678	0.00339
9	<i>B1821-24</i>	<i>IRPSR</i>	<i>0.03620</i>	<i>0.01448</i>	<i>0.00724</i>	<i>0.00362</i>
10	<i>B1937+21</i>	<i>IRPSR</i>	<i>0.04053</i>	<i>0.01621</i>	<i>0.00811</i>	<i>0.00405</i>
11	B0614+091	XB	0.10901	0.0436	0.02180	0.01090
12	B1957+20	BRPSR	0.12031	0.04812	0.02406	0.01203
13	XTE 1751-305	XB	0.14782	0.05913	0.02956	0.01478
14	SAX J1808.4-3658	XB	0.16396	0.06558	0.03279	0.01640
15	XTE J1814-338	XB	0.36302	0.14521	0.0726	0.03630
16	B1731-260	XB	0.92815	0.37126	0.18563	0.09282
17	XTE J1807-294	XB	1.82588	0.73035	0.36518	0.18259
18	<i>B0540-69</i>	<i>IRPSR</i>	<i>3.04340</i>	<i>1.21736</i>	<i>0.60868</i>	<i>0.30434</i>
19	B1744-24A	XB	3.33933	1.33573	0.66787	0.33393
20	AO 0538-66	XB	20.01224	8.00490	4.00245	2.00122
21	XTE J0929-314	XB	24.06032	9.62413	4.81206	2.40603
22	<i>B1823-13</i>	<i>IRPSR</i>	<i>39.84015</i>	<i>15.93606</i>	<i>7.96803</i>	<i>3.98402</i>
23	<i>B1509-58</i>	<i>IRPSR</i>	<i>41.27643</i>	<i>16.51057</i>	<i>8.25529</i>	<i>4.12764</i>
24	J0218+4232	BRPSR	64.10714	25.64286	12.82143	6.41071
25	<i>J1124-5916</i>	<i>IRPSR</i>	<i>96.35034</i>	<i>38.54014</i>	<i>19.27007</i>	<i>9.63503</i>
26	<i>J1846-0258</i>	<i>IRPSR</i>	<i>116.2085</i>	<i>46.48342</i>	<i>23.24171</i>	<i>11.62085</i>
27	<i>J0205+6449</i>	<i>IRPSR</i>	<i>182.3901</i>	<i>72.95603</i>	<i>36.47801</i>	<i>18.23901</i>
28	B1656+354	XB	214.4759	85.79037	42.89518	21.44759
29	<i>J1811-1925</i>	<i>IRPSR</i>	<i>226.9788</i>	<i>90.79153</i>	<i>45.39576</i>	<i>22.69788</i>
30	<i>J1930+1852</i>	<i>IRPSR</i>	<i>429.6507</i>	<i>171.8603</i>	<i>85.93015</i>	<i>42.96507</i>

observation of the most X-ray binaries, its period stability, temporary characteristic and eclipse depends on the space observation, so we suggest evaluating its feasibility after accumulating a certain space-based data.

48.4.2 Database Construction Strategy

In order to accelerate the implementation of X-ray pulsar navigation technology, here we proposed to develop and launch a low orbit pulsar observation experimental satellite to construct the navigation database, which can utilize the existing small satellite platform. As the auto-navigation method using X-ray binaries need

future deep study [5, 12], this paper only considers the isolated rotation-powered pulsars to construct trail database.

Suppose the lifetime of the experimental satellite is 3 years, the timing residuals requirement is 1 μs , according to the numerical results in Table 48.4, the first 4 pulsars to construct the database can choose PSR B0531+21, PSR B1821-24, PSR B1937+21 and PSR B0540-69. If the effective area of the space detector A_d is 0.1, 0.25, 0.5 and 1 m^2 , the corresponding construct database observation period is 3 years, 1.2 years, 1 year and 1 year; if add the 5th pulsar PSR B1823-13 in the observation task, the corresponding construct database observation period is 39.8 years, 15.9 years, 8 years and 4 years. Actually, the real observation time is usually larger than the above time when considering the shadow of the earth and the observation efficiency of the space detector.

48.5 Conclusions

This paper presents a direct use of space-based observation data to construct navigation pulsar database. The numerical analysis results shown that by launching a experimental pulsar navigation satellite, with an effective area of 0.25 m^2 of the space detector, the experimental navigation database construction with four pulsars within 2 years: PSR B0531+21, PSR B1821-24, PSR B1937+21 and PSR B0540-69, which can get 100 m orbit precision. The database construction chose X-ray binaries and auto-navigation using the X-ray binaries will be studied in the future.

References

1. Ping S, Ming L, Shao-long C (2009) The principle and method of X-ray pulsar navigation system. Chinese Aerospace Press, Beijing
2. Lorimer DR, Kramer M (2005) Handbook of pulsar astronomy. Cambridge University Press, Cambridge
3. Rodin A, Sekido M (2002) Pulsar VLBI observations. In: Proceedings of the 6th European VLBI network symposium, Bonn, Germany, 25–28 June 2002
4. Blandford R, Teukolsky SA (1976) Arrival-time analysis for a pulsar in a binary system. *Astrophys J* 205:580–591
5. Huang LW, Liang B, Zhang T et al (2012) Navigation using binary pulsars. *Sci China Ser G* 55(3):527–539
6. Damour T, Deruelle N (1986) General relativistic celestial mechanics II: the post-newtonian timing formula. *Ann Inst H Poincaré (Physique théorique)* 44:263–292
7. Sheikh SI (2005) The use of variable celestial X-ray sources for spacecraft navigation. Maryland University, College Park
8. Cheng-shi Z, Ding C, Hong-bing C et al (2011) Research on X-ray pulsar navigation sources. *Prog Astron* 29(3):335–342
9. Manchester RN, Hobbs GB, Teoh A et al (2005) The Australia telescope national facility pulsar catalogue. *Astron J* 129:1993–2006

10. The ATNF Pulsar Catalogue [EB/OL] 2005[2013]. <http://www.atnf.csiro.au/research/pulsar/psrcat>
11. Hobbs G, Archibad A, Arzoumanian Z et al (2010) The international pulsar timing array project: using pulsars as a gravitational wave detector. *Class Quantum Gravity* 27:1–10
12. Huang LW, Liang B, Zhang T (2013) Pulse phase and doppler frequency estimation of X-ray pulsars under conditions of spacecraft and binary Motion and its application in navigation. *Sci China, Ser G* 56(4):848–858

Chapter 49

Physical Analysis on Pulsar-Based Navigation System: Preliminary Designs of Timing Model and a New Prototype of X-Ray Detector

Yuanjie Du

Abstract Six X-ray pulsars (PSR J0218+4232, PSR J0437-4715, PSR B0531-21, PSR J1939+2134, PSR J1024-0719 and PSR J0751+1807) are selected for the mission of deep-space autonomous navigation, and the relevant parameters (including position, proper motion, period and period derivative, etc.) are also listed. Based on the XMM-Newton X-ray timing data and the radio ephemeris for the young pulsar PSR B1509-58, we adopted both absolute timing and relative timing methods to fold its pulse profile in the 0.2–12 keV band. In addition, for the sake of both pulsar-based deep-space autonomous navigation and scientific research, we take the Crab pulsar as an example to fit its X-ray (0.2–30 keV band) spectrum, and obtain its power-law fitted spectrum. Then we estimate the navigation accuracy of the detectors with different effective areas. Finally, we discuss a preliminary probe design principle of X-ray detectors for the pulsar navigation, and give the rough estimations on the effective area and sensitivity.

Keywords Pulsars autonomous navigation X-ray detector crab pulsar · Timing model

49.1 Introduction

In the vast universe, there is a class of very special celestial bodies, namely pulsars, which have a great potential value of application. They are a class of compact objects that radiate electromagnetic pulse, have extremely stable rotation and contain the four interactions (the strong interaction, the weak interaction, electromagnetic force and gravity). Therefore, pulsars are a natural “extreme physical

Y. Du (✉)

Qian Xuesen Laboratory of Space Technology, No. 104 Youyi Road,
Haidian District, Beijing, China
e-mail: dyj@nao.cas.cn

laboratory”. As one of the four discoveries in astronomy in the 1960s, scientists have obtained twice “the Nobel Prize in Physics” because of pulsars. Professor Hewish and his student Bell found the first pulsar in 1967, Hewish therefore won the Nobel Prize in 1974; US Hulse and Taylor discovered a binary pulsar, thus indirectly confirmed the existence of gravitational waves predicted by Einstein, and won the Nobel Prize in 1993. Pulsars are one class of the most important research celestial objects in astronomy community.

Electromagnetic radiation from a pulsar is mainly concentrated in a bundle-like lighthouse. For each rotation, its radiation beam sweeps the Earth, we therefore receive a pulse. It is generally believed that the pulse period of a pulsar is its rotation period. Since now, we have found more than 2500 pulsars, their rotation periods are in a distribution from 1.4 ms to 12 s.

It is generally believed that, pulsars are rotating neutron stars with strong magnetic field (10^8 – 10^{12} G), their mass is about 1.4 solar masses, while their radius is only 10 km. Because the moment of inertia of a pulsar is very large (about 10^{45} g cm²), the change rate of its rotation period is very small, only a nanosecond or less per year, or even only 0.1 picoseconds for millisecond pulsars. Due to its magnetic dipole radiation and loss of relativistic particles, this leads to reduction in the rotation energy, so its rotation gradually slows down. Pulsars can be divided into normal pulsars and millisecond pulsars, and most pulsars were discovered in the radio band. Normal pulsars have slightly large rotation periods and strong magnetic fields (10^{11} – 10^{12} G). Millisecond pulsars have rotation periods less than 10 ms, and their period derivative are 10^{-18} – 10^{-21} s s⁻¹ (i.e., their characterized magnetic field strength are 10^8 – 10^9 G).

Some young pulsars and millisecond pulsars have a full-band pulsed radiation from radio (10^{-6} eV) to the gamma-ray band (tens or even hundreds of GeV), the X-rays and gamma rays dominated the pulsar’s radiation luminosity. If the radiation beam of a pulsar sweeps the earth, we can receive a corresponding pulse signal, but the received radio or X-ray pulse signals for a rotation period are often weak (much smaller than the system noise of radio receiver or X-ray detector), so the actually observed pulse profiles are accumulated within a few hundred to a few thousand cycles, with considerations of the various effects that affect the obtained signal. The typical value of duty cycle (duty cycle or pulse width) of a radio pulse is only 5–10 %, while the duty cycle of millisecond pulsars is usually larger than those of young pulsars. Radio pulses usually have a single peak, double peaks or multimodal peaks, some radio pulses have a very sharp spike. If the observed SNR were excellent, the radio pulses from a pulsar are hardly indistinguishable.

Pulsar timing is to study the time of arrival (TOA) pulse, due to the influence of the Earth’s rotation and revolution movement, so the measured TOA is relative to the Solar System Barycenter (SSB). Pulsar’s period and position changes (pulsar has a high space velocity), the interstellar medium, the effect of the electromagnetic wave dispersion (DM), the Earth’s rotation and revolution movement, relativistic effects, the position change of the reference coordinate system, the intrinsic changes of pulsar’s radiation and other effects will affect the TOA to some extent. With technology improving and experience accumulation of pulsar observations, the

SNR of a pulse profile we observed is growing significantly. In addition to the intrinsic changes of pulsar radiation, the other major influence on TOA has been largely studied and become clear, but there would be some improvements for a few issues.

In sum, the pulsar research is in a transition phase from basic research to applied research, the observation data in each band is increasingly extensively. Some novel and detailed observations have been found, this will be a strong guidance to physics of pulsars, the prospects will be brighter.

In the following, we will briefly introduce the existing methods for deep space navigation:

- (1) Global Positioning System (GPS) can only provide the navigation services to the low-orbit satellite and objects on the Earth's surface. It becomes quite powerless for a spacecraft of deep space exploration or interplanetary flight.
- (2) The Deep Space Network (DSN), formed by three 70-m radio antenna in California, Madrid and Canberra, can provide communication and navigation services for a spacecraft within Mars orbit. The service requires helps from ground stations, and the determined position accuracy is about a few kilometers to tens of kilometers. In order to meet the communication and navigation needs for the Mars mission in future, the DSN has been expanded and a satellite network for Mars has been proposed to build. The network consists of small satellites and Mars relay satellites system which can provide independent and efficient navigation and communication services for a spacecraft to Mars.
- (3) Star tracker or sun sensor (sun sensor) can also be used as a deep-space autonomous navigation service. But limited by its angular resolution, the spacecraft's position accuracy determined by this method changes significantly over time and have accumulative effect on the errors up to thousands of kilometers.
- (4) The idea of pulsar navigation had been proposed long ago. In the 1970s, shortly after the discovery of pulsars, it is considered to be used as a navigation radio source. Because radio flux of a pulsar is extremely low, it is needed to use a large-diameter antenna to obtain a pulse profile with a long integration time. Combined with too much equipment payload and much high positioning errors may be caused, it is not suitable for space navigation in the radio band. X-ray emission from many pulsars (including millisecond pulsars) has a slightly higher count rate, only a short integration time is needed to get a pulse profile with a high signal-to-noise ratio, the navigation candidates of X-ray pulsars are also very rich. In sum, the X-ray band overcomes the shortcomings of radio band, the advantages of easy miniaturization of equipment and data processing makes it suitable for deep space spacecraft autonomous navigation.

The basic principle of X-ray pulsar autonomous navigation is to get the TOA of an X-ray pulse measured by X-ray detector mounted on the spacecraft, then we compare the measured TOA with the one forecasted by the pulsar timing model, the

projected distance relative to line-of-sight direction in the solar system is finally estimated. Using four pulsars and some methods such as Kalman filtering, we can ultimately determine the spacecraft's position and time error of onboard clock. It can also be combined with the method of inertial navigation, only one pulsar may meet the needs of deep space navigation. Therefore, selection of navigation X-ray pulsars and efficient X-ray detector are the key techniques which directly determine the accuracy of positioning.

NASA approved a space satellite mission, namely "The Neutron star Interior Composition ExploreR (NICER)", with an energy band of 0.2–12 keV, is expected to be installed on the International Space Station (ISS) in 2016. The main instrument of NICER is X-ray Timing Instrument (XTI). XTI has 56 concentrators (nested grazing incidence focusing optics) and X-ray detectors, which are combined into an effective area of about 2000 cm² (at 1.5 keV). For each focused optical mirror system with 24 nesting cones, the effective receiving area is about 50 cm²; the viewing angle is 15 arcmin; a silicon drift detector (SDD) is used. GPS timing and location is also adopted, which provides the absolute time stamp for each X-ray photon better than 0.3 μs [1]. The main scientific concept of NICER is to study the internal composition of neutron stars, kinetics and mechanism of particle acceleration in pulsars' magnetosphere. The practical value is to validate the technique of X-ray pulsar navigation and X-ray communications.

Many teams had carried out the related works on pulsar navigation. For example: a team set up a semi-physical simulation system to verify the key technology of X-ray pulsar navigation [2]; Deng et al. [3] used 4 ms pulsars to simulate the pulsar navigation results for a spacecraft on a way to Mars, the simulated error location is better than ~20 km.

49.2 Database of Navigation Pulsars

49.2.1 Preliminary Selection of Navigation Pulsars

More than 2500 pulsars have been found, more than 150 have X-ray pulses. However, only 52 ms pulsars have X-ray pulsed radiation, half of which is located in globular clusters that are not suitable for navigation. Therefore, the number of alternative navigation pulsars is about 30, including young pulsars and millisecond pulsars.

Selection of the suitable pulsars is the premise for pulsar navigation. The main selection criteria depend on the pulsar's own physical properties (e.g., position, X-ray flux, pulse width and timing model, etc.), and also rely on the performance properties of the X-ray telescope (such as effective collection area, sensitivity, onboard clock and system errors).

To meet the future navigation algorithms and positioning accuracy, we qualitatively determine that navigation pulsars are a portfolio of young and millisecond pulsars. The main selection criteria are listed in the following:

- (1) precise observed position coordinates, small space velocity (the best values of these parameters can be derived by pulsar VLBI observations);
- (2) good astronomical environment (far away from the Galactic plane), low background noise, high SNR X-ray flux with smaller variability, simple pulse profile with narrow and sharp peaks;
- (3) few glitch (non-regularity) phenomena, have long-term stable pulse profile;
- (4) combination of pulsars in a reasonable spatial distribution (when the spacecraft is flying in some certain orbits, the Earth or the Moon and other celestial bodies may sometimes obscure certain X-ray pulsars, which affects the whole-track navigation);
- (5) excellent timing model (with considerations of various timing noise).

We preliminarily determine the navigation sources: PSRs J0437-4715, J1024-0719, J1939+2134, J0534+2200, J0218+4232 and J0751+1807, their corresponding parameters are listed in Table 49.1, where P is a rotation period of a pulsar; \dot{P} is the period derivative; RA and DEC are right ascension and declination; D is the distance; PM_RA and PM_DEC are proper motion in RA and DEC directions; F_x is the flux in the X-ray band. Figure 49.1 shows the locations of six pulsars in the galactic coordinates.

We applied the European VLBI Network (EVN) observations of PSR J0218+4232 which had five epochs lasting for 3 years, each epoch had 6 h continuous observation, and then we processed the data and obtained the position, distance and proper motion of PSR J0218+4232 that are listed in Table 49.1. In addition to PSR J0218+4232, the parameters of other pulsars are taken from ANTF pulsar catalogue (<http://www.atnf.csiro.au/research/pulsar/psrcat/>) [10]. We also convert the proper motion of PSR J0751+1807 in the ecliptic coordinate system presented in the literature [11] to the ones in the equatorial coordinate system.

49.2.2 Timing Model of Pulsars

In order to achieve the goal of pulsar navigation, we need a precise timing model to predict the TOA of an X-ray pulse from a selected pulsar. The TOA of a pulse is obtained by a correlation calculation between the observed pulse and the standard pulse template. Standard pulse template is a pulse profile with high SNR obtained by a long observation time, then it is fitted by the Gaussian or circular Gaussian (Mises) functions. Timing model is a text file that contains a lot of pulsar parameters (for example: position, rotation frequency and its first, second, multi-order (etc.) derivatives, proper motion, dispersion measure (DM), binary model and its parameters, the solar system ephemeris, timing noise model parameters, etc.), it is generally obtained in the radio band.

Table 49.1 The selected X-ray navigation pulsars and their parameters

PSR	P (ms)	PI (s/s)	RA (hms)	DEC (dms)	D (kpc)	PM_RA (mas/yr)	PM_DEC (mas/yr)	Fx (ph/s/cm ²)
J0437-4715	5.76	5.73e-20	04:37:15.883	-47:15:09.0318	0.16	121.68 [4]	-71.82 [4]	1.3×10^{-4}
J1024-0719	5.16	1.85e-20	10:24:38.688	-07:19:19.1702	0.53	-35.3 [5]	-48.2 [5]	6.7×10^{-6}
J1939+2134	1.56	1.05e-19	19:39:38.559	+21:34:59.13	8.33	-0.46 [6]	-0.66 [6]	1.3×10^{-5}
J0534+2200	33.08	4.23e-13	05:34:31.973	+22:00:52.06	2.0	-14.7 [7]	2.0 [7]	1.5
J0218+4232	2.32	7.74e-20	02:18:06.354	+42:32:17.419	6.3	5.35 [8]	-3.74 [8]	3.7×10^{-5}
J0751+1807	3.481	7.70e-21	07:51:09.157	+18:07:38.603	0.62	-1.53 [9]	-5.83 [9]	1.1×10^{-5}

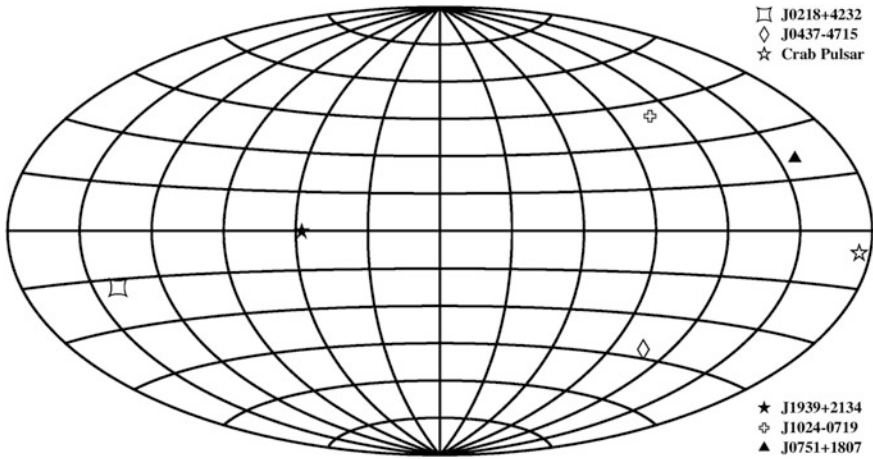


Fig. 49.1 Positions of navigation pulsars in the galactic coordinate

The most widely used software for pulsar timing is “Tempo2” [11], which is developed from the old timing software “Tempo”. Tempo2 uses high-precision solar planetary ephemeris (DE200, DE405, DE414 and DE421) derived from NASA’s Jet Propulsion Laboratory (JPL), and the ephemeris of DE200 and DE405 are commonly used (Note: The latest version of JPL ephemeris is DE431).

For different pulsars, the TOA and corresponding timing model can be obtained in the radio observations, and Tempo2 can get accurate timing model parameters with high precisions. This timing model could be applied to the high-energy band, Tempo2 can be used to obtain X-ray and gamma-ray pulse profiles. However, this approach is called “relative timing” in terms of X-ray band. The radio timing model takes into account the effect of interstellar medium dispersion for radio electromagnetic waves. Dispersion measure (DM) can be used to characterize the amount of this effect. DM is not constant, and has small changes each year. Coupled with the effects of the intrinsic timing red noise, the results are sometimes unsatisfactory with respect to high-precision timing models. DM has very small impact on high energy (X-ray and gamma-ray) emission, which can be negligible. Therefore, we should refer to the radio method (autocorrelation of observation and template pulse profiles, then we can directly get TOA of an X-ray profile, we finally fit the timing model (without DM parameter) and get the optimized parameters, this method is called as “absolute timing” in the X-ray band. The absolute timing model for a radio-quiet pulsar (e.g. the Geminga pulsar) is obtained directly in the X-ray band.

In order to compare the absolute timing and relative timing methods, we first obtained the X-ray data of PSR B1509-58 observed by EPIC-PN onboard XMM-Newton satellite (obsid: 0312590101; observing modes: timing; observation time: 2006 February 20; Length: 31,946 s; time resolution: 30 μ s); then we used the absolute and relative timing methods to obtain its pulse profile of 64 bins, as shown in Fig. 49.2. Using the relative timing ephemeris derived from the Parkes telescope,

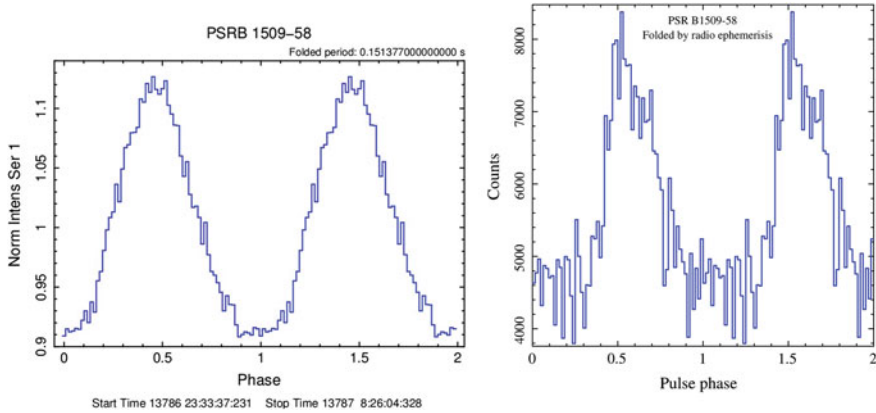


Fig. 49.2 X-ray pulse profiles (XMM-Newton) of PSR B1509-58 (64 bins); (*Left panel* absolute timing; *Right panel* relative timing)

we can get the X-ray profile of PSR B1509-58 via tempo2 software with the “photons” plugin. Xronos package in HEASFOT is used as absolute timing software. As shown in Fig. 49.2, there is big difference between the two pulse profiles obtained by the two methods, and the corresponding error of fold period (imputed to the X-ray observation date) is $1.6\text{E-}6$. Many effects (e.g., different systematic errors observed by different instruments and DM errors in radio timing model) can account for this difference, better data is required to identify the advantages and disadvantages of the two methods.

49.3 Physical Design of X-ray Detector

Conducting wide energy-band, high time and energy resolution X-ray observations on pulsar-like compact objects can reveal fundamental problems of magnetospheric physics and equation of state of neutron stars, as well as applications of pulsar navigation services. High-performance X-ray detector is the key factor to achieve these scientific goals and pulsar navigation applications. We have sustained demand on the X-ray detector onboard astronomical satellites.

This chapter addresses these needs, taking the Crab pulsar as an example, by fitting its X-ray pulsed spectrum, we explore the selection of energy range for X-ray detector (Sect. 49.3.1); estimate the navigation accuracy for the X-ray detectors with different effective area (Sect. 49.3.2); propose a X-ray detector with a new prototype of X-ray focusing mirrors (Sect. 49.3.3), and estimate its sensitivity (Sect. 49.3.4).

49.3.1 Energy Range of X-ray Detector

The Crab pulsar is a young “neutron star” with a period of 33 ms, its wind nebulae was first discovered by the Chinese people in 1054. It is the strongest pulsar with steady X-ray pulsed flux. Because the spectrum of its wind nebulae is very stable, it is the calibration source for X-ray satellites.

There are two main components for X-ray pulsed emission from young pulsars, the first one is non-thermal synchrotron radiation from the extremely relativistic charged particles in the magnetosphere; the other is thermal X-ray emission produced by the hot spots with a very high temperature ($10^7\text{--}10^8$ K) in pulsar’s polar cap regions. While the X-ray unpulsed radiation is produced by synchrotron radiation from charged particles in the wind nebulae and other noise sources. Through precise timing and spectral observations, we can separate non-thermal and thermal radiation from pulsed emission. The useful X-ray signal for pulsar navigation is pulsed emission, the unpulsed X-ray emission is a noise source.

In order to fit the phase-averaged spectrum of Crab pulsar in the X-ray band (0.1–30 keV), we obtained the spectral data from the literature [12], and found that the power-law spectrum can be a good fit to these data (see Fig. 49.3). The obtained spectral density equation is as follows:

$$dN/dE = 0.453E^{-1.724}$$

Integrate the equation above in the soft X-ray band, we can further get the integral spectrum for the Crab pulsar, the results is shown in Fig. 49.4. It is shown that the different values of the integral upper limit E_{max} (10, 15, 30 keV)

Fig. 49.3 X-ray pulsed spectra fitting for the crab pulsar

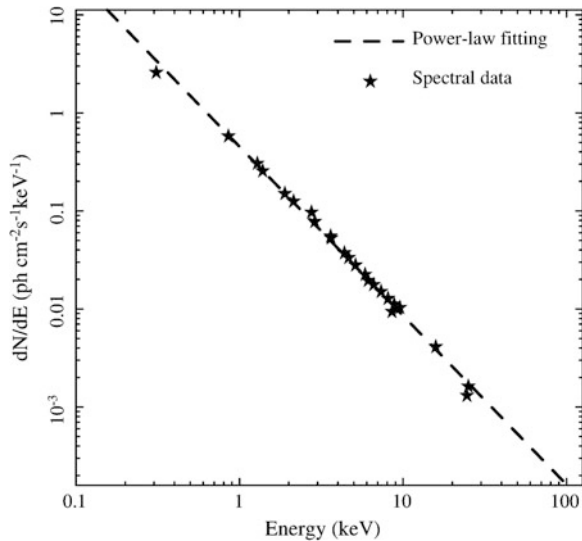
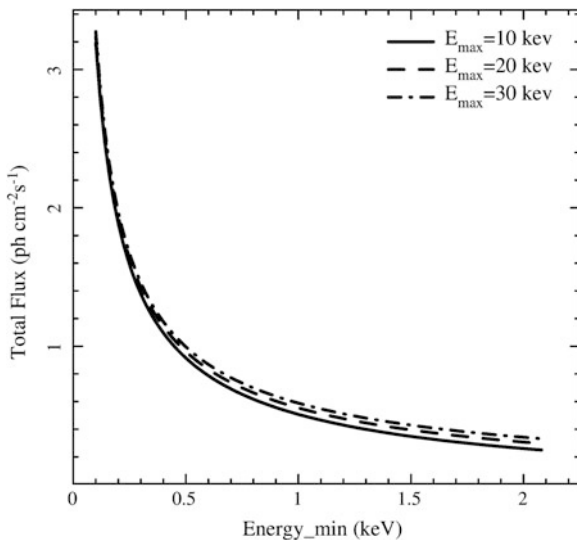


Fig. 49.4 The influence of the lower limit of photon energy (E_{\min}) on the X-ray integrated spectrum of the crab pulsar



contribute little to the integral spectrum, but the integral lower limit E_{\min} contribute much to the integral spectrum. When E_{\min} is smaller, the total photon number (integral spectrum) is larger. For example: when $E_{\min} = 0.5$ keV, the obtained integral photon flux is $\text{Flux} = 0.98 \text{ ph cm}^{-2} \text{ s}^{-1}$; when $E_{\min} = 1$ keV, the resulting $\text{Flux} = 0.5 \text{ ph cm}^{-2} \text{ s}^{-1}$. This is very meaningful to the selected energy range of X-ray detector, it is better to select the smaller E_{\min} within the realm of technical possibility. However, when E_{\min} is too small (e.g. $E_{\min} \leq 0.1$ keV), the detector is more susceptible to the diffuse X-ray noisy photons, which results in a low SNR for an observation. Additionally, in terms of technique, the quantum efficiency at low energy end of X-ray silicon drift detector (SDD) is very small, the effective area decreases significantly. Therefore, the energy range of X-ray detector should be chosen at $\sim 0.2\text{--}15$ keV.

49.3.2 Estimation of TOA Precision for Pulsar Navigation

In order to estimate the influence of X-ray detector's different effective area on navigation TOA accuracy, we still took the Crab pulsar as an example, and used its pulse width and observed SNR to obtain the rough TOA accuracy.

Assuming that the number of observed X-ray noise photons follows a Poisson distribution, the noise photon σ_{noise} is square root of the total number, so the SNR of an observed pulse profile can be written as:

$$\begin{aligned}
 SNR &= \frac{N_{pulse}}{\sigma_{noise}} = \frac{N_{pulse}}{\sqrt{(N_B + N_{unpulse})d_c + N_{pulse}}} \\
 &= \frac{N_{pulse}}{\sqrt{[B_X + F_X(1 - P_f)]A_e t_{obs}d_c + F_X A_e P_f t_{obs}}} \tag{49.1}
 \end{aligned}$$

where P_f is the pulse fraction that is a percentage of the pulsed flux relative to the total flux; d_c is the duty cycle, i.e. the proportion of the entire pulse cycle; F_X is the total X-ray flux; B_X is X-ray background noise flux, which has a relationship with the energy band of the detector; A_e is the effective area; t_{obs} is the observation time.

Thus, the TOA accuracy can be determined by the half width of the pulse profile and SNR, namely:

$$\sigma_{TOA} = 0.5W/SNR \tag{49.2}$$

where W is the pulse width, σ_{TOA} is the accuracy of TOA.

We used the existing X-ray pulse profile data of the Crab pulsar to give a rough estimation of the TOA accuracy with different detector efficient area (1200, 2500, 5000 and 10000 cm^2), energy range (1–15 and 0.2–15 keV) and integration time (100, 500, 1800 and 3600 s), the results are shown in Fig. 49.5.

It is found that: when an X-ray detector has an energy range 0.2–15 keV, an effective area of 2500 cm^2 and a long observation (integration) time of 1800 s, this leads to an accuracy of 4 μs , and the corresponding TOA position error (spherical radius) is about 1.2 km, which was able to meet the needs of deep space navigation.

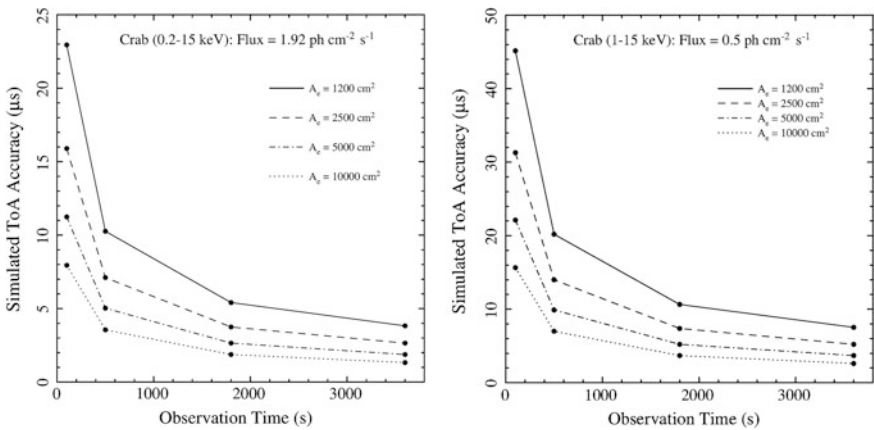


Fig. 49.5 TOA precisions of X-ray pulsar navigation

49.3.3 Structure Design of X-ray Detector

Focused detectors can not only meet the detection need of a large collection area, but also can greatly reduce the interference of background X-ray photons or cosmic rays. Because of its focus capability, the SDD with a very small area is required, which leads to a small parasitic capacitance and easy detection of low energy photons.

We intend to use the revolution paraboloid focusing system to focus X-rays. Using confocal multilayer formed by a paraboloid of revolution nested focusing system, coupled with the backend of SDD and the associated electronic system, we thus constitute an X-ray detector cell, then assemble a group of detector cells to meet the needs of arbitrarily large collection area. The transmittance and quantum efficiency of a SDD vary greatly with photon energy, so the effective area is a function of X-ray photon energy, and generally refers to the one at a specific energy (e.g. 1 keV). Through optimized design, an X-ray detector cell, using 11 nesting mirrors, has an effective area of 125 cm^2 (at 1 keV). 20 Such cells can form a large detector with an effective area of 2500 cm^2 (at 1 keV), the relation between effective area and energy are shown in Fig. 49.6. The effective area of the X-ray detector is at peak segment in the 0.8–2 keV band, and the effective area decreases rapidly in the <0.2 and >5 keV bands.

This type of X-ray detector has high energy resolution, fast response time, and small weight. It is easy to realize the needs of large area through an array of combinations.

49.3.4 Estimation on Sensitivity of X-ray Detector

The sensitivity of an X-ray detector is the minimum measured flux for some objects (e.g., pulsars), its value depends on the location of this pulsar, background noise flux, performance of the detector (e.g., effective area, quantum efficiency, etc.) and

Fig. 49.6 The relation between the effective area and energy for the X-ray detector

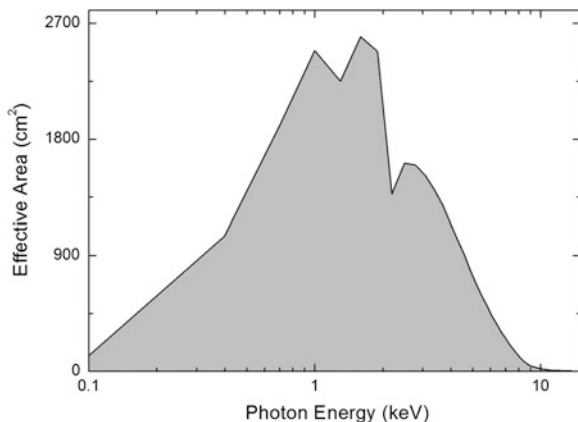


Fig. 49.7 The relation between the sensitivity and signal-to-noise for the X-ray detector

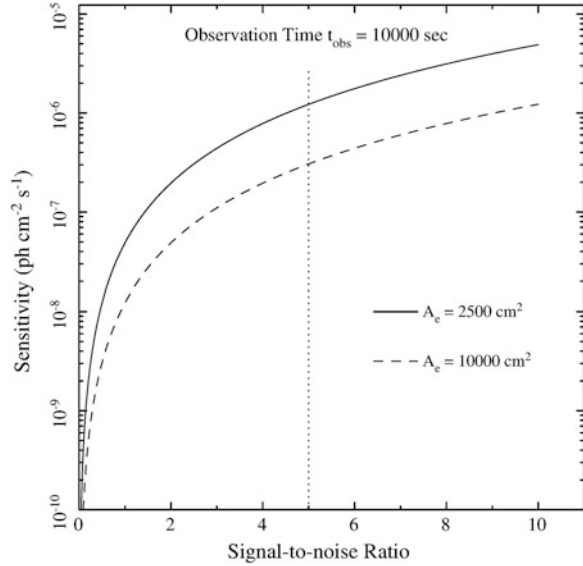


Table 49.2 The sensitivities of the X-ray detector with various effective areas

A_e (cm ²)	S (SNR = 5) (erg/cm ² /s)	S (SNR = 3) (erg/cm ² /s)
2000	3.47E-014	1.25E-014
2500	2.77E-014	9.99E-015
10000	6.85E-015	2.46E-015

observation time (t). Sensitivity is the most important factor for the performance of the X-ray detector. Background noise is mainly from two sources: (1) the diffuse X-ray background radiation in Milky Way; (2) the unpulsed radiation from pulsar nebulae. Therefore, the sensitivity of X-ray detector depends on the background noise near the observed source. In high-energy astrophysics, it is generally believed that the signal is valid when SNR is greater than or equal to 5 (or 3).

Taking the Crab pulsar as an example, we can derive the relationship between the sensitivity and SNR using the Eq. (49.1). The result is presented in Fig. 49.7. When taking SNR = 5, we calculated the corresponding sensitivity of the X-ray detector with different effective area (2000, 2500 and 10000 cm²), as shown in Table 49.2. The newly designed detector with an effective area of 2000 cm² is comparable to NICER.

49.4 Conclusion

Selection of proper navigation pulsars from the 30 qualified pulsars is the premise and foundation to achieve the great goal of pulsar navigation. The selection criteria is mainly dependent on pulsar's physical properties (e.g., position, X-ray flux, pulse width and pulsar timing model). It also has a strong correlation on the X-ray telescope (e.g., effective collection area, sensitivity, time resolution, stability of onboard clock and other systematic errors). Therefore, we need to launch a scientific satellite, equipped with the same type of X-ray telescope, to continually observe a group of pulsars for a long time. The absolute and relative timing models for selected pulsars are expected to obtain. When comparing with these two methods, we can further filter out the available navigation pulsars for deep space autonomous navigation services in future.

Acknowledgments I am grateful to Dr. Baoquan Li in National Space Science Center (CAS) for useful discussions. This work is supported by the National Natural Science Foundation of China (grant No. 11303069 and 11373011).

References

1. Gendreau KC, Arzoumanian Z, Okajima T (2012) The neutron star interior composition explorer (NICER): an explorer mission of opportunity for soft x-ray timing spectroscopy. *Proc SPIE* 8443:13
2. Liu L, Zheng W, Tang G, Sun S (2012) The hardware-in-loop simulation system of X-ray-pulsars-based navigation. *J Nat Univ Defense Technol* 34(5):10–14
3. Deng XP, Hobbs G, You XP et al (2013) Interplanetary spacecraft navigation using pulsars. *Adv Space Res* 52(9):1602–1621
4. Deller AT, Tingay SJ, Bailes M, Reynolds JE (2009) Precision southern hemisphere VLBI pulsar astrometry II: measurement of seven parallaxes. *Astrophys J* 701(1243):J0437–4715
5. Verbiest JPW, Bailes M, Coles WA et al (2009) Timing stability of millisecond pulsars and prospects for gravitational-wave detection. *Mon Not R Astron Soc* 400(951):J1024–0719
6. Cognard I, Bourgois G, Lestrade JF et al (1995) High-precision timing observations of the millisecond pulsar PSR 1937+21 at Nancay. *A&A* 296(169):J1939+2134
7. Ng CY, Romani RW (2006) Proper motion of the crab pulsar revisited. *Astrophys J* 644(445) (Crab)
8. Du Y, Yang J, Campbell RM, Janssen G, Stappers B, Ding C (2014) VLBI-measured proper motion and parallax of the -ray millisecond pulsar PSR J0218+4232. *ApJL*, 782(L38):J0218+4232
9. Nice DJ, Splaver EM, Stairs IH et al (2005) A 2.1 msolar pulsar measured by relativistic orbital decay. *Astrophys J* 634(1242):J0751+1807
10. Manchester RN, Hobbs GB, Teoh A, Hobbs M (1993) The Australia telescope national facility pulsar catalogue. *Astron J* 2005:129
11. Hobbs GB, Edwards RT, Manchester RN (2006) TEMPO2, a new pulsar-timing package—I. An overview. *Mon Not R Astron Soc* 369:655
12. Kuiper L, Hermsen W, Cusumano G et al (2001) The crab pulsar in the 0.75–30 MeV range as seen by CGRO COMPTEL. A coherent high-energy picture from soft X-rays up to high-energy gamma-rays. *A&A* 378:9

Chapter 50

GNSS P2P Cooperative Positioning System for Multiple Search-and-Rescue Robots

Yulong Song and Baowang Lian

Abstract Multi-robots are often better choices to complete search-and-rescue tasks in some large-scale disasters, such as earthquakes, mine accidents, forest fires, etc. However, effects such as shadowing and fading for GNSS signals limit the positioning ability which is most important for search-and-rescue robots. To improve positioning availability and reliability, the article proposes a GNSS Peer-to-Peer cooperation positioning system for multi-robots search-and-rescue. The peers share GNSS-only data among neighbors as aiding information under light block scenario, and share both GNSS and terrestrial ranging data under deep indoor scenario. A particle filtering algorithm, the Monte Carlo numerical approximation of Bayesian filtering, is proposed to estimate position of peers utilizing both the prior information coming from robot motion model and posterior information provided by pseudo-range and terrestrial range observations, and the algorithm flow chart is presented. As a result, the acquisition time could be reduced and sensitivity could be improved for peers under light block scenario, and position could be solved under deep indoor scenario with fewer than 4 visible satellites. Simulation results show that the positioning error of particle filtering is less than that of Non-Bayesian filtering, and the error is about 5 m for low-cost receivers.

Keywords Multi-robots · Search-and-Rescue · Global navigation satellite system · Peer-to-Peer cooperative positioning · Particle filter

50.1 Introduction

Robots can perform better in search-and-rescue tasks in the disaster scene, and satellite navigation is a key factor. The performance of GNSS receiver is widely known to be very good in open sky conditions but may decline strongly in difficult

Y. Song (✉) · B. Lian

School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China
e-mail: songyulong8982@126.com

© Springer-Verlag Berlin Heidelberg 2015

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015*

Proceedings: Volume III, Lecture Notes in Electrical Engineering 342,

DOI 10.1007/978-3-662-46632-2_50

situations like earthquakes, mine accidents, forest fires, etc. When the number of visible satellites is less than 4, a receiver can not compute its position. The augmentation systems such as AGNSS and DGNSS are centralized network with fixed infrastructure. They cover limited places, and may not be available in certain disaster site. At this point, if multiple search-and-rescue robots share location-related information, each node as an aid to other ones, making up a GNSS P2P (Peer-to-Peer) cooperative positioning system, then they can greatly improve positioning availability and robustness, reduce positioning time-consuming, and own flexibility to adapt to specific environments, even to deep indoor scenario.

The concept of cooperative positioning comes from mobile robot navigation [1], and it has been introduced in satellite navigation field in recent years, showing a strong application value. GNSS P2P cooperative positioning system is the extension of the augmentation like AGNSS and DGNSS, and it can not only improve positioning availability, reliability and accuracy, but also provide new solution ideas in anti-multipath [2], anti-spoofing interference [3], high-precision positioning [4, 5] and other fields.

GNSS cooperative positioning system can be divided into two categories: GNSS-only and hybrid GNSS cooperative positioning system. In GNSS-only system, each user has a GNSS receiver for positioning and outputting observations and a communication system for exchanging data with other users, while in hybrid system user also has a terrestrial ranging system. Roberto Garello et al. introduced the structure and implementation of the two systems in [6, 7]. In this paper, a more practical cooperative positioning mechanism is proposed for search-and-rescue multi-robots combining both GNSS-only and hybrid GNSS cooperative positioning system. Algorithm implementation and simulation results are provided.

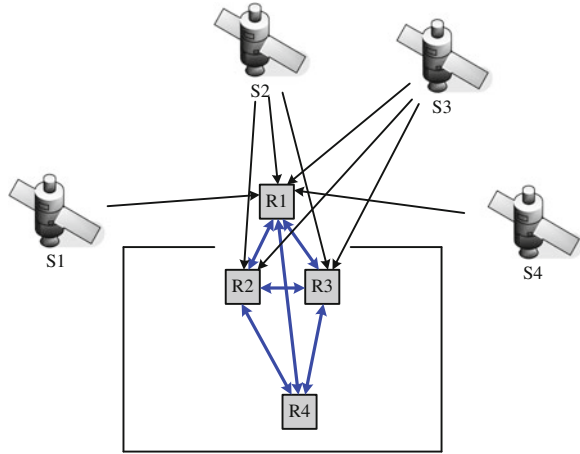
50.2 GNSS P2P Cooperative Positioning for Search-and-Rescue

50.2.1 Scenario Analysis

Disaster scenes are much complex, and can be divided into two categories: light block scenario and deep indoor scenario. Under light block scenario each mobile agent has a sufficient number of visible satellites with low C/N_0 , and could reduce average acquisition time and improve sensitivity by use of GNSS-only data shared by neighboring agents.

Under deep indoor scenario some of the agents have insufficient number of visible satellites, so they can not compute position independently. As shown in Fig. 50.1, double arrow lines represent communication links between every two nodes, and single arrow lines represent signals transmitting from a satellite to a node. Robot node R2, R3 and R4 can not compute position independently, but can make it when their terrestrial ranges are employed.

Fig. 50.1 Hybrid cooperative positioning in deep indoor scenario



Pseudo-range measurement of agent m respect to satellites s in P2P network is show as Eq. (50.1):

$$\rho_{sm} = \sqrt{(x_s - x_m)^2 + (y_s - y_m)^2 + (z_s - z_m)^2} + c\delta_m + v_{sm} \quad (50.1)$$

where (x_s, y_s, z_s) is ECEF coordinate of satellite s ; (x_m, y_m, z_m) is ECEF coordinate of agent m ; δ_m is clock offset of agent m ; v_{sm} is pseudo-range error of agent m respect to satellite s .

Terrestrial range measurement between agent m and n is show as Eq. (50.2):

$$r_{nm} = \sqrt{(x_n - x_m)^2 + (y_n - y_m)^2 + (z_n - z_m)^2} + \mu_{nm} \quad (50.2)$$

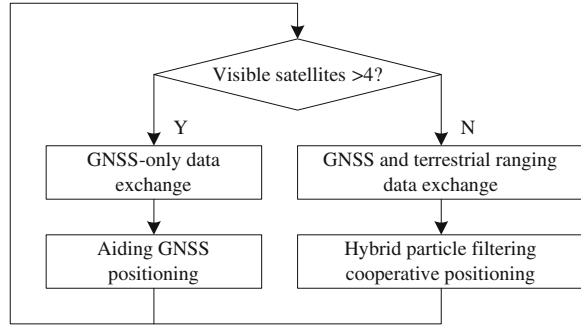
where (x_n, y_n, z_n) is ECEF coordinate of agent n ; μ_{nm} is the range error between agent m and n .

For the scenario shown in Fig. 50.1, each node has four state variables, i.e. XYZ coordinates and the clock offset, with 16 unknowns in total. The number of pseudo-range measurements shown as formula (50.1) is 8, and that of terrestrial range measurements shown as formula (50.2) is 6. What's more, there are 2 more equations for equal heights of R1–R4, and also 16 equations in total. As a result, every node's position can be solved.

50.2.2 Positioning Mechanism

Mobile agents can compute position using only GNSS and can improve positioning performance aided by neighbouring agents under light block scenario, while under deep indoor scenario they may be not able to compute position independently. To solve this problem, a new positioning algorithm merging pseudo-range and

Fig. 50.2 Cooperative positioning for search and rescue



terrestrial range measurement should be proposed. In this article, a hybrid cooperative positioning particle filtering algorithm is presented, though a little increase in the amount of calculation. A search-and-rescue cooperative positioning mechanism shown in Fig. 50.2 is proposed. Use GNSS-only data shared by neighbouring agents as aiding information to compute position when visible satellites are enough, while use particle filter to merge both GNSS pseudo-range and terrestrial range measurements and compute position when there are not enough visible satellites.

50.3 GNSS P2P Cooperative Positioning Algorithm

50.3.1 GNSS-Only Data Exchange

Agents share GNSS-only data under light block scenario, including visible satellites number, position, clock offset, C/N_0 , Doppler frequency, code phase and secondary code phase. Information coming from augmentation system like AGNSS or DGNSS is also a part, while synchronization of peers is the key problem. There are a lot of research literatures about synchronization of peers in P2P network. In addition, the reliability of different nodes' aiding information varies as different range and position accuracy. Weight sum of aiding information is employed.

Low-cost receiver can reduce mean acquisition time and improve receiver sensitivity by exchanging GNSS-only data, achieving similar performance to AGNSS system. When a high-end receiver is employed in P2P network, the whole performance will be greatly improved [8].

50.3.2 Terrestrial Ranging Data Exchange

Agents share not only GNSS data but also terrestrial ranging under deep indoor scenario. RSS (Received Signal Strength) and TOA (Time of Arrival) are two candidate ranging methods.

RSS method measure signal strength of signal from neighbouring nodes, and compute the range by use of propagation attenuation model. Many commercial Wifi and ZigBee modules provide RSS ranging, so it is easy to realize. However, this method is very susceptible to interference, and accuracy is very poor with error on the order of tens of meters, which is not suitable for search-and-rescue applications.

TOA method gets range by measuring the signal trip time between two nodes. To avoid the stringent synchronization requirements between nodes, two-way TOA ranging method is utilized. The two-way trip time is show as formula (50.3):

$$t = 2\Delta + \varepsilon \quad (50.3)$$

where Δ is the time of one way trip; ε is the processing time of neighboring node.

$$d = c\Delta = c(t - \varepsilon)/2 \quad (50.4)$$

where c is the speed of light.

Ultra Wide Band signal can provide decimeter-level TOA ranging accuracy because of its nanosecond-level width pulse. As a result, UWB is used for terrestrial ranging in multi-robots search-and-rescue cooperative positioning system.

50.3.3 Hybrid Cooperative Positioning Algorithm

50.3.3.1 Bayesian Filtering

Optimal filter can estimate the state of time-varying system in real time utilizing noisy measurement. The first proposed optimal filter is LMS linear optimal filter, and linearization is the first step. Bayesian filter is a recursive optimal filter based on Bayesian inference. The recursive process is divided into two stages: prediction and update. Bayesian filter takes full advantage of the priori and posterior probability distribution of the state to compute posterior system state. Kalman Filter (KF, EKF, UKF), particle filter, grid filter are applications of Bayesian filtering [9].

State x_k of search-and-rescue robots in hybrid cooperative positioning consists of ECEF coordinates and clock offset, and observation z_k consists of pseudo-range and terrestrial range measurements. Measurements are only concerned with the current state. States are hidden, and only concerned with the state of the previous time. All those fit Markov model shown in Fig. 50.3, so Bayesian filtering should be taken into consideration.

System model of agent m in hybrid cooperative positioning system is shown as formula (50.5) and (50.6):

$$\mathbf{x}_m^{(k)} = f(\mathbf{x}_m^{(k-1)}, \boldsymbol{\omega}_m^{(k)}) \quad (50.5)$$

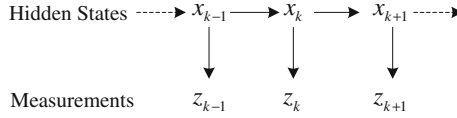


Fig. 50.3 Hidden Markov model

$$\omega_m^{(k)} \sim N\left(0, \mathbf{Q}_m^{(k)}\right) \tag{50.6}$$

where $\mathbf{x}_m^{(k)}, \mathbf{x}_m^{(k-1)}$ is the state of node m at time slot k and $k - 1$ respectively and defined as formula (50.7) for search-and-rescue robots. Assuming each robot is equipped with speed and geomagnetism sensor, and real-time speed and heading angle can be obtained, then velocity in ECEF $(\dot{x}, \dot{y}, \dot{z})$ is input, error brought by clock offset is $b = c\delta$, and \dot{b} can be considered as 0; $f(\cdot)$ is state transition function as formula (50.8); $\omega_m^{(k)}$ obey normal distribution with expectation of 0 and variance of $\mathbf{Q}_m^{(k)}$.

$$\mathbf{x}_m^{(k)} = (x, y, z, b, \dot{x}, \dot{y}, \dot{z}, \dot{b}) \tag{50.7}$$

$$f\left(\mathbf{x}_m^{(k-1)}, \omega_m^{(k)}\right) = \begin{bmatrix} \mathbf{I} & \Delta t_k \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{x}_m^{(k-1)} + \begin{bmatrix} \Delta t_k \mathbf{I} \\ \frac{\Delta t_k^2}{2} \mathbf{I} \end{bmatrix} \omega_m^{(k)} \tag{50.8}$$

where \mathbf{I} is a 4×4 identity matrix; $\mathbf{0}$ is a 4×4 zero matrix; elapse time between two state is $\Delta t_k = t_k - t_{k-1}$.

Measurement model of agent m in hybrid cooperative positioning system is shown as formulas (50.9) and (50.10):

$$\mathbf{z}_m^{(k)} = h\left(\mathbf{x}_m^{(k)}, \mathbf{P}_m^{(k)}, \mathbf{S}_m^{(k)}, \mathbf{v}_m^{(k)}\right) \tag{50.9}$$

$$\mathbf{v}_m^{(k)} \sim N\left(0, \mathbf{R}_m^{(k)}\right) \tag{50.10}$$

where $\mathbf{z}_m^{(k)}$ is the measurement at time slot k ; $\mathbf{P}_m^{(k)}$ is the position set of neighboring nodes of node m and $\mathbf{P}_m^{(k)} = \left\{ \mathbf{p}_n^{(k)} \mid n \in M_m^{(k)} \right\}$; $\mathbf{S}_m^{(k)}$ is the visible satellites set of node m and $\mathbf{S}_m^{(k)} = \left\{ \mathbf{p}_s^{(k)} \mid s \in S_m^{(k)} \right\}$; $\mathbf{v}_m^{(k)}$ is the measurement noise vector, and obey normal distribution with expectation of 0 and variance of $\mathbf{R}_m^{(k)}$; $h(\cdot)$ is the measurement function.

(1) Prediction Stage

Transition of state is Markov stochastic process, so priori *p.d.f.* of state can be obtained using Chapman-Kolmogorov equation as shown in formula (50.11):

$$p(\mathbf{x}^{(k)} | \mathbf{z}^{(1:k-1)}) = \int p(\mathbf{x}^{(k)} | \mathbf{x}^{(k-1)}) p(\mathbf{x}^{(k-1)} | \mathbf{z}^{(1:k-1)}) d\mathbf{x}^{(k-1)} \quad (50.11)$$

where $p(\mathbf{x}^{(k)} | \mathbf{x}^{(k-1)})$ is the system state transition probability at time slot k , and can be obtained via formula (50.8); $p(\mathbf{x}^{(k-1)} | \mathbf{z}^{(1:k-1)})$ is posterior *p.d.f.* at time slot $k - 1$.

(2) Update Stage

Posterior *p.d.f.* of state can be obtained using Bayesian rule when current measurement $\mathbf{z}_m^{(k)}$ is available.

$$p(\mathbf{x}^{(k)} | \mathbf{z}^{(1:k)}) = \frac{p(\mathbf{z}^{(k)} | \mathbf{x}^{(k)}) p(\mathbf{x}^{(k)} | \mathbf{z}^{(1:k-1)})}{p(\mathbf{z}^{(k)} | \mathbf{z}^{(1:k-1)})} \propto p(\mathbf{z}^{(k)} | \mathbf{x}^{(k)}) p(\mathbf{x}^{(k)} | \mathbf{z}^{(1:k-1)}) \quad (50.12)$$

where $p(\mathbf{z}^{(k)} | \mathbf{x}^{(k)})$ can be obtained using formula (50.9); $p(\mathbf{z}^{(k)} | \mathbf{z}^{(1:k-1)})$ is constant relative to state x , and can be ignored in practice.

50.3.3.2 Particle Filtering

Particle filter (PF) is the Monte Carlo numerical approximation of Bayesian filtering. PF makes use of a set of particles to estimate posterior probability of system state indirectly, and it can be generally applied to nonlinear systems and non-Gaussian noise. Also, it is easy to program.

Particle set is made up by a number of particles with states and weight coefficients as shown below:

$$\{x_{m,i}^{(k)}, w_{m,i}^{(k)}\}_{i=1}^N \quad (50.13)$$

where N is the number of particles; $x_{m,i}^{(k)}$ is the state of particle i for node m at time slot k ; $w_{m,i}^{(k)}$ is the weight coefficient of particle i for node m at time slot k .

Posterior probability estimated by particles is shown as formula (50.14):

$$p(x_m^{(k)} | z_m^{(1:k)}) \approx \sum_{i=1}^N w_{m,i}^{(k)} \delta(x_m^{(k)} - x_{m,i}^{(k)}) \quad (50.14)$$

where $\delta(x)$ is Dirac delta function, and the weight coefficient can be presented by formula (50.15):

$$w_{m,i}^{(k)} = p\left(\mathbf{z}_m^{(k)} \mid \mathbf{x}_{m,i}^{(k)}\right) \quad (50.15)$$

where $p\left(\mathbf{z}_m^{(k)} \mid \mathbf{x}_{m,i}^{(k)}\right)$ is likelihood function of measurement as shown below:

$$p\left(\mathbf{z}_m^{(k)} \mid \mathbf{x}_{m,i}^{(k)}\right) = \prod_{n \in M_m^{(k)}} p_{nm}\left(r_{nm}^{(k)} - \left\|\hat{\mathbf{p}}_n^{(k)} - \mathbf{p}_{m,i}^{(k)}\right\|\right) \cdot \prod_{s \in S_m^{(k)}} p_{sm}\left(\rho_{sm}^{(k)} - \left\|\mathbf{p}_s^{(k)} - \mathbf{p}_{m,i}^{(k)}\right\| - b_{m,i}^{(k)}\right) \quad (50.16)$$

where p_{nm} is p.d.f of ranging error between node m and n ; $r_{nm}^{(k)}$ is the range measurement between node m and n ; $\hat{\mathbf{p}}_n^{(k)}$ is the estimated position of node n ; $\mathbf{p}_{m,i}^{(k)}$ is the position of i -th particle for node m ; p_{sm} is p.d.f of pseudo-range error of node m respect to satellite s ; $\rho_{sm}^{(k)}$ is pseudo-range measurement of node m respect to satellite s ; $\mathbf{p}_s^{(k)}$ is the position of satellite s ; $b_{m,i}^{(k)}$ is the error brought by clock offset of node m ; $\|\cdot\|$ represents Euclidean distance.

The Fig. 50.4 shows the GNSS P2P hybrid cooperative positioning particle filter algorithm flowchart. The bigger N is the performance is better, and the algorithm is very suitable for computer processing.

50.4 Simulation Results and Analysis

Figure 50.5 shows the actual trajectory and estimated trajectories for a rescue robot agent under light block scenario. It can be seen that the trajectory estimated by non-Bayesian filtering LMS algorithm fluctuates much heavier with the positioning error of 15 m. The particle filtering algorithm takes full advantage of priori information coming from system model and posteriori information coming from measurement model, leading to a smaller position error, where the error of velocity measurement sensor is several centimeters per second.

Figure 50.6 shows the horizontal position error for 4 mobile agents in GNSS P2P cooperative positioning network under deep indoor scenario. Assuming each agent is equipped with a low-cost receiver and the position accuracy is about 15 m, and the ranging error utilizing UWB TOA method is at decimeter level, simulation results show that the position accuracy could be about 5 m for those agents without enough visible satellites with the help of hybrid cooperative positioning particle filtering algorithm merging both pseudo-range and terrestrial range measurement.

Fig. 50.4 Particle filtering positioning flow for agent m

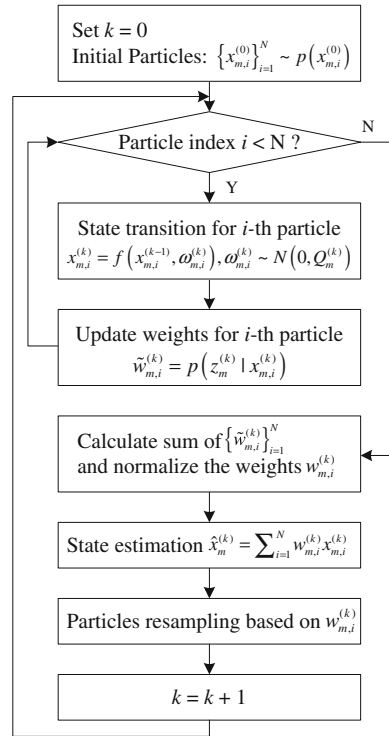


Fig. 50.5 Trajectory for search and rescue robot

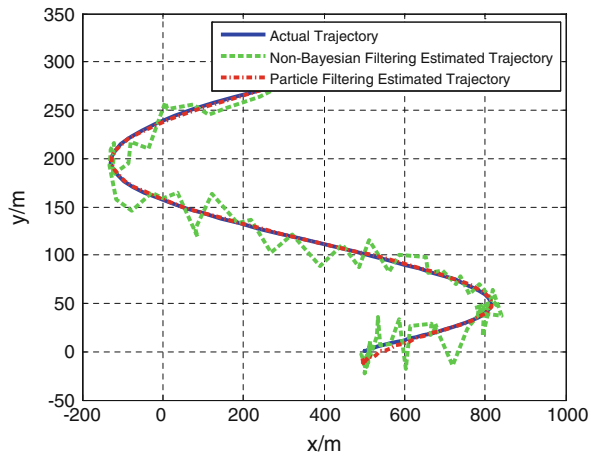
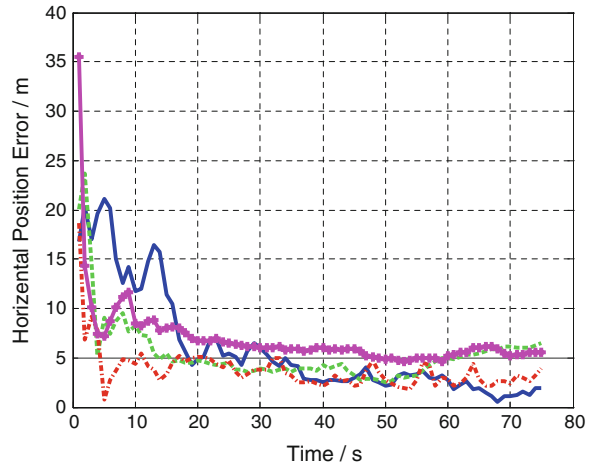


Fig. 50.6 Horizontal RMSE for particle filtering in hybrid cooperative positioning system



50.5 Conclusions

GNSS P2P cooperative positioning system is constructed by multiple mobile agents equipped with GNSS receivers, communication systems and other sensors, and could be considered as the extension of AGNSS and DGNSS augmentation. The cooperative positioning system does not require fixed infrastructure, and could suit challenge environment by use of GNSS positioning information shared by neighboring peers. What's more, the system can also solve position under deep indoor scenario with the help of terrestrial ranging system, which is very suitable for multi-robots search-and-rescue tasks at complicated disaster sites.

This paper analyzes the cooperative positioning mechanisms for both light block scenario and deep indoor scenario. Under light block scenario there are enough visible satellites with low C/N_0 for every peer. They share GNSS-only data to aid each other, and the performance is similar to AGNSS. While under deep indoor scenario there are often not enough visible satellites for some peers, so they can not compute their positions using only GNSS. To solve this problem, terrestrial ranging data is also shared among peers to complete hybrid cooperative positioning. There are too many unknown state variables for system model and measurement model of hybrid cooperative positioning, so the linearization of equations set are not easy. A particle filtering algorithm, the Monte Carlo numerical approximation of Bayesian filtering, is proposed to solve the nonlinear problem, and it is very suitable for computer processing. This paper presents the simulation of particle filtering cooperative positioning algorithm under deep indoor scenario, which verifies the feasibility and positioning performance. Horizontal positioning error for low-cost receiver is about 5 m.

References

1. Kurazume R, Hirose S, Nagata S et al (1996) Study on cooperative positioning system. In: IEEE international conference on robotics and automation, Minneapolis, USA, pp 1421–1426, 22–28 April 1996
2. Jau P-H, Chang F-R, Du DHC et al (2014) Mitigating the multi-path effect of GNSS in the urban area by using a cooperative wilcoxon method through VANET, 2014 IEEE/ION, Monterey, CA, pp 1088–1095, 5–8 May 2014
3. Heng L, Work DB, Gao GX (2013) Cooperative GNSS authentication: reliability from unreliable peers. *Inside GNSS* 8(5):70–75
4. Kiam JJ, Cárdenas JM, Henkel P (2014) Cost-effective cooperative RTK positioning for rowing boats. In: ION ITM 2014, San Diego, Calif, pp 541–550 27–29 Jan 2014
5. S Banville, P Collins, P Tétreault et al (2014) Precise cooperative positioning: a case study in canada. In: ION GNSS 2014, Tampa, FL, 8–12 Sept 2014
6. Garelo R, Presti LL, Corazza GE et al (2012) Peer-to-peer cooperative positioning. Part I GNSS Aided Acquis *Inside GNSS* 7(2):55–63
7. Garelo R, Samson J, Spirito MA et al (2012) Peer-to-Peer cooperative positioning. Part II: Hybrid Devices GNSS Terr Ranging Capab *Inside GNSS* 7(4):56–64
8. Lina D, Claudio P, Francesco B et al (2010) Impact of high-end receivers in a Peer-to-peer cooperative localization system. In: UPINLBS 2010, Helsinki, Finland, 14–15 Oct 2010
9. Sarkka S (2013) Bayesian filtering and smoothing, Cambridge University Press, Cambridge

Chapter 51

A Position Approach Based on the Special Sub-frame in TDD LTE-A System

Dongyan Wei, Zhili He, Xuping Gong, Ying Xu and Hong Yuan

Abstract With the urgent need for indoor and outdoor seamless navigation service, it is imperative to develop high precise positing technique suitable for various scenarios. In this paper, considering the widely deployment and well coverage of the wireless cellular communication system, a position approach is proposed based on time duplex division long term evolution advanced (TDD LTE-A) network. A novel complex frame structure for positing is proposed based on the current TDD LTE-A protocol. In GP (Guard Period) of the special sub-frame, i.e., the downlink-uplink switch period of TDD system, the navigation (NV) signal is added for positing. NV signals are transmitted simultaneously from different base station (BS) with orthogonal PN (Pseudo Noise) code. By measuring the distance to different BSs which are time synchronized in TDD system, user's position can be fixed as in the GNSS system. In this paper, the theoretical performance of the proposed position scheme will be analyzed, and the impact to the current system will also be discussed. The results show that, by properly allocating the time length of NV signal based on the coverage of the cell, signal conflicts with the current LTE-A signal can be effectively avoided.

Keywords TDD LTE-A · Special sub-frame · Positing

51.1 Introduction

Recently, location based service (LBS) has become a very important field for both researchers and engineers. The applications of LBS arise from both commercial navigation services and emergency rescue. On the one hand, this is due to commercial services such as vehicle navigation, location based search service, and

D. Wei (✉) · Z. He · X. Gong · Y. Xu · H. Yuan
Academy of Opto-Electronics, Chinese Academy of Science, Beijing 100094, China
e-mail: weidongyan@aoe.ac.cn

machine-to-machine (M2M) network. On the other hand, in the United States, all wireless providers have been mandated by Federal Communications Commission (FCC) to report the user's location with specified accuracy.

Several positing methods have been proposed to solve the location problem [1]. Global navigation satellite system (GNSS) is a promising positing technology that can provide high precision location service globally and have been widely used. However, in dense urban and indoor scenarios, the performance of GNSS will be severely degraded, since the number of satellites in view is often insufficient to fix the position. Therefore, non-satellite navigation techniques have been well studied recently, such as cellular system based position, Wireless local area network (WLAN) based navigation, vision assisted navigation and INS assisted navigation. With globally indoor and outdoor seamless coverage, navigation technique based on cellular communication system is a promising technique especially for indoor and urban canyon scenarios. In [2], Round Trip delay Time (RTT) based position method is proposed in WCDMA network. In [3], fingerprint assisted cell-ID position approach is proposed. In [4], TDOA and AOA based location scheme is proposed for wideband CDMA cellular systems. All the above mentioned techniques can fix a position with the precision of about 100 m, which can't meet the FCC E911 obligation. In the latest LTE and LTE-A protocol, dedicated positioning reference signal (PRS) is reserved. However, the pseudo-range measurement precision is very low since only 180 kHz (12 carriers) bandwidth is allocated to PRS. In addition, PRS occupies many frequency-time resources, which impacts the spectrum efficiency of system.

In this paper, considering the special sub-frame, i.e., the downlink and uplink switch sub-frame in time division duplex (TDD) LTE-A system, a novel location scheme is proposed by adding navigation (NV) signal in guard period (GP), in which both the base station (BS) and user equipment (UE) keep silence to avoid uplink and downlink signal collision. By properly designing the duration of the NV signal with the restriction of current LTE-A protocol, we will see that the proposed scheme will compact with the current LTE-A signal frame. In other words, the proposed navigation signal will cause no influence to the current LTE-A signal. In addition, since all the bandwidth and transmission power in current LTE-A network can be used for NV signal, the pseudo-range measurement precision can be greatly improved in the proposed scheme.

The rest of this paper is organized as follows. Section 51.2 refers to the frame structure in the current TDD LTE-A system. In Sect. 51.3, the proposed navigation scheme is introduced. Furthermore, interference with the current system is analyzed in Sect. 51.4. In Sect. 51.5, examples and simulation results of the proposed scheme are shown. Finally, conclusions are in Sect. 51.6.

51.2 Frame Structure in TDD LTE-A System

In LTE-A, both TDD and FDD are used [5]. The duration of one LTE-A radio frame is 10 ms, which contains 2 half frame with 5 sub-frames each. In TDD LTE-A, the same frequency resource is used for both uplink (UL) and downlink (DL). UL and DL data are carried in different sub-frames in TDD model. According to the data traffic, the distribution of subframes for DL and UL can be adapted. As shown in Fig. 51.1, the transition sub frame between DL and UL is called special sub frame, which consists of downlink pilot time slot (DwPTS), guard period (GP) and uplink pilot time slot (UpPTS).

To avoid the signal conflicts between DL and UL, the time duration of GP should be adapted according to the cell size. Nine different special subframe configurations are supported in TDD LTE-A as shown in Table 51.1, where T_s is a basic time unit that satisfies $307200 \cdot T_s = 10 \text{ ms}$ (the length of 1 radio frame).

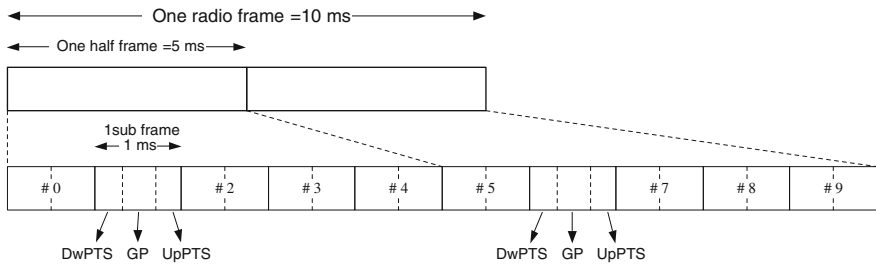


Fig. 51.1 Frame structure of TDD LTE-A system

Table 51.1 Special sub-frame configuration in TDD LTE-A system

Configuration	DwPTS	GP	UpPTS
0	$6592 \cdot T_s$	$21936 \cdot T_s$	$2192 \cdot T_s$
1	$19760 \cdot T_s$	$8768 \cdot T_s$	
2	$21952 \cdot T_s$	$6576 \cdot T_s$	
3	$24144 \cdot T_s$	$4384 \cdot T_s$	
4	$26336 \cdot T_s$	$2192 \cdot T_s$	
5	$6592 \cdot T_s$	$19744 \cdot T_s$	$4384 \cdot T_s$
6	$19760 \cdot T_s$	$6576 \cdot T_s$	
7	$21952 \cdot T_s$	$4384 \cdot T_s$	
8	$24144 \cdot T_s$	$2192 \cdot T_s$	
9	$13168 \cdot T_s$	$13168 \cdot T_s$	

In addition, since UEs locate at different place of the cell, the UL transmission delays to BS are different among the UEs. In TDD LTE-A protocol, timing advance (TA) process is employed to make sure signals from different UEs will reach the BS at the same time. Every UE is informed by the BS when to start transmitting, depending on the distance from UE to BS. The greater the distance between the BS and the UE, the earlier the UE starts to transmit.

In TDD cellular system, time synchronization is required among the BSs to prevent inter-cell interference. In the current commercial networks, the synchronization precision is about $\pm 1.5 \mu\text{s}$, which is very small for data communication. However, for positioning, effective range deviation due to the clock error is measured by multiplying the time error by velocity of light c ($3 \times 10^8 \text{ m/s}$), which is about 450 m. In fact, with the development of low cost and high performance clock, the synchronization precision of the BS can be highly improved. Alternatively, by distributing dedicated monitoring stations, signals from different BSs can be tracked, and the clock bias can be tracked and broadcast to the user through the data communication link. Thus, to simplify the analysis, in this paper, we assume the clock biases among the BSs are known to the user.

51.3 Proposed Navigation Scheme

As shown in Fig. 51.2a, GP is a transition period in TDD system, in which both BS and user keep silence. In this paper, based on GP, navigation (NV) signal is added in the special subframe as shown in Fig. 51.2b. In addition, GP1 is reserved to avoid the signal collision between NV and DwPTS signal of the neighbor cells. And GP2 is reserved to avoid the signal collision between NV and UpPTS signal of the neighbor users. It will be shown that under the proper design of the length of GP1 and GP2, NV signal will cause no influence to the current LTE-A signals.

In the NV period, navigation signal modulated by the Cell-ID of each cell is broadcasted by every BS simultaneously. In this paper, DSSS (Direct Sequence Spread Spectrum) signal is employed for the ranging code as in GNSS. As shown in

Fig. 51.2 The proposed frame structure based on TDD LTE-A

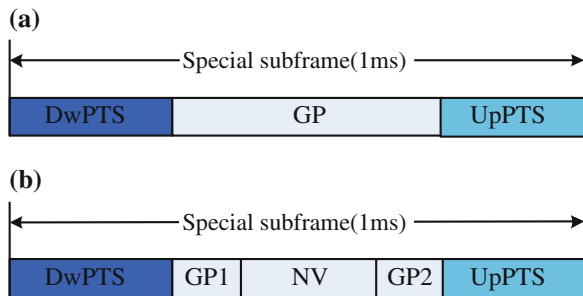


Fig. 51.3 Cell structure in cellular communication system

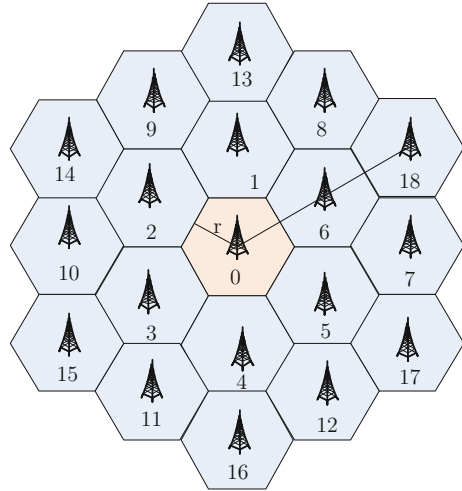


Fig. 51.3, benefits from the cell structure, signals from several BSs can be received by UE. In realistic scenario, signals from the BSs of 2 to 4 tiers can be seen by the user. In this paper, to simplify the analysis, only 19 cells (2 tiers) are considered, and we will take cell 0 in the center as an example.

By demodulating the spectrum spread signal from different BS, the persuade range between UE and every BS can be obtained using the code or carrier phase of the signal. By delivering the positions of the BSs to UE via the data communication link of LTE-A, UE’s position can be fixed just as in GPS.

51.4 Interference Analysis

1. Interference between DwPTS and NV

For each UE, the transmission delays of the signals from various BSs are different, since the distance to every BS is different. Therefore, the NV signal from the near BS may be interfered with the DwPTS from the farer BS as shown in Fig. 51.4. Obviously, to prevent the confliction between DwPTS and NV of different BS, the duration of GP1 should be long enough so that the NV signal of the BS with farrest distance will arrive at the UE without interference from DwPTS signal of neighbor cell. In Fig. 51.4, we can conclude that $GP1$ should be larger than $delay_{max}^i - delay_{min}^i$, where $delay_{max}^i$ and $delay_{min}^i$ is respectively the maximum and minimum transmission delay from the BSs to the i -th user. Considering the distribution of the users in the cell, $GP1$ should be subject to

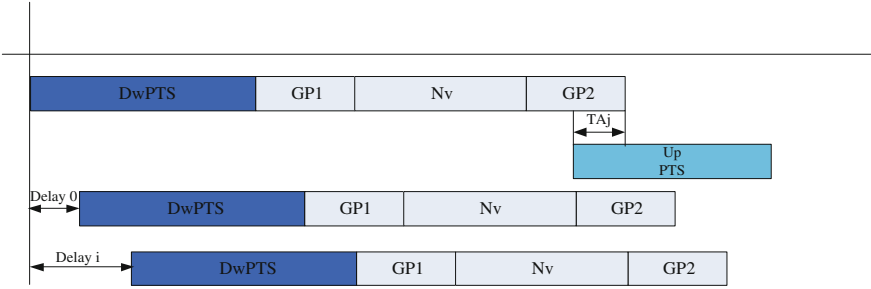


Fig. 51.4 Interference among DwPTS, NV and UpPTS

$$\begin{aligned}
 GP_1 &= \max_{all\ the\ user} (delay_{max}^i - delay_{i,min}) \\
 &\leq \max_{all\ the\ user} (delay_{max}^i) - \min_{all\ the\ user} (delay_{i,min})
 \end{aligned}
 \tag{51.1}$$

Obviously, both parts in the above inequality get the extreme value when UE locates at the center of the cell. Then $\max_{all\ the\ user} (delay_{max}^i) = 4r/c$, which is the transmission latency from the BS of outer tier to BS of cell 0 (e.g., BS18 to BS0). And $\min_{all\ the\ user} (delay_{i,min}) = 0$, since the transmission latency is 0 from BS0 to UE, under the assumption that UE locates at the center of the cell. Then, GP_1 can be obtained as

$$GP_1 = 4r/c - 0 = 4r/c
 \tag{51.2}$$

where r is the radius of the cell.

2. Interference between UpPTS and NV

As mentioned in Sect. 51.3, under the TA process of TDD LTE-A, UE starts the transmission a little earlier to keep uplink time alignment at BS. Therefore, NV signal from all the BSs should be ended before the UE start to transmit UpPTS. Additionally, considering the transmission latency of the BS' signal to UE, GP_2 should be larger than $delay_{max}^i + TA_i$, where TA_i is the time length of TA process of UE*i*. Considering the distribution of the users in the cell, GP_2 should be subject to

$$GP_2 = \max_{all\ the\ user} (delay_{max}^i + TA_i)
 \tag{51.3}$$

To simplify the analysis, we assume UE locates on the line between BS0 and BS18. Then, BS15 is the farrest BS to UE, and $delay_{max}^i = d_1/c$ where d_1 is the distance from BS15 to UE. TA length TA_i is d_2/c , where d_2 is the distance from BS0 to UE. Obviously, the right part of Eq. (51.3) gets the extreme value when UE locates at edge of the cell, and

$$GP_2 = 5r/c + r/c = 6r/c \quad (51.4)$$

3. Interference among NV signals

According to the above analysis, by reserving GP1 and GP2 at the two sides of NV signal, the collision with DwPTS and UpPTS signal can be effectively prevented. Thus, for the NV signal, interference only comes from the NV signals of other cells. Different from the communication system, in the position system, the NV signal of every cell has to be demodulated. So the expression of SINR (Signal to Interference plus Noise Ratio) for the NV signal of every cell will be discussed. In fact, with the reservation of GP1 in the proposed signal structure, the interference among the NV signals only happens at some of the time. In other words, SINR will vary with time t . For the NV signal of cell j , $SINR$ can be obtained as

$$SINR^j(t) = \frac{P_j |h_j|^2 / B}{n_0 + \sum_{i \in \mathbb{Z}, i \neq j} \kappa_{ij}(t) P_i |h_i|^2 / B} \quad (51.5)$$

where P_i is the transmission power of the i -th BS, h_i is the channel coefficient between the i -th BS and the user, B is the system bandwidth, n_0 is the additive Gaussian white noise. The interference flag $\kappa_{ij}(t)$ in Eq. (51.5) means whether the NV signals from BS i and BS j collides at time t . $\kappa_{ij}(t) = 1$ means the NV signal of BS i and BS j arrive at the user at the same time t , and $\kappa_{ij}(t) = 0$ means there is no signal of cell i arrive at the user at time t .

In cellular communication system, there exists heavy near-far effect for users in the center of the cell, since the receiving signal power of BS0 is much higher than that of other BS's. Thus, for the NV signal from BS0, it can be easily decoded with a large SINR. For the NV signals from other BSs, several near-far effect mitigation techniques can be applied, such as successive interference cancellation (SIC) [6] and double transmission technique (DTT) [7]. In our analysis, we assume that the NV signal from BS0 has been subtracted when decoding the signals from other BS. Thus, the SINR expression for other BS1-BS18 can be rewritten as

$$SINR^j(t) = \frac{P_j |h_j|^2 / B}{n_0 + \sum_{i \in \mathbb{Z}, i \neq j, i \neq 0} \kappa_{ij}(t) P_i |h_i|^2 / B} \quad (51.6)$$

51.5 Examples and Simulation Results

In this section, an example of the proposed position scheme is given based on the current LTE-A protocol. Configurations with the longest (configuration 0) and shortest (configuration 4) GP length are chosen. And the parameters in Table 51.2

Table 51.2 Simulation parameters

Parameters	Value
Carrier frequency	2.0 GHz
Number of cells	19 cells (2 tier)
Cell radius	1 km
Bandwidth	10 MHz
Transmission power	3 dBm
Noise PSD	-174 dBm/Hz
BS antenna gain	14 dBi
Path loss	$128.1+37.6l \text{ g}(d)$, d in km
Parameter of GP (configuration 0)	GP:0.7141 ms GP1:0.0133 ms, NV:0.6807 ms, GP2:0.02 ms
Parameter of GP (configuration 4)	GP:0.0714 ms GP1:0.0133 ms, NV:0.0380 ms, GP2:0.02 ms

are used for simulation. To prevent the interference to cells out of the 2 tiers in Fig. 51.3, the transmission power of the NV signal is set to 3dBm, which is about 20 dB lower than the power of signals in LTE-A.

1. DOP value

The parameter of geometric Dilution Precision (DOP) plays an important role in positioning. Since the distribution of base station is usually in horizontal plane, only Horizontal DOP (HDOP) is considered as shown in Fig. 51.5. Benefits from the distribution of the cell structure, it can be seen that the HDOP in cell 0 is about 0.5 and 1.1 respectively, when 2 tiers and 1 tires are considered.

2. SINR

SINR of the NV signal from each BS is shown in Fig. 51.6 for cell edge UE and cell center UE respectively. Obviously, SINR of the signals from inner tier is higher than that of outer tier. In addition, NV signals from the BSs of inner tier arrive at the UE earlier than the NV signals from the BSs of outer tier; therefore, the SINR is higher at the beginning.

3. Theoretical pseudo-range measurement precision

Based on SINR of the NV signals from each BS, theoretical pseudo-range measurement precision (TPRMP) can be obtained as

$$\text{TPRMP} \geq \frac{\sqrt{2T_s B_r}}{\sqrt{T_w \cdot \text{SINR} \cdot (2\pi B_{rms})^2}} \quad (51.7)$$

where T_s is the sampling rate, B_r is the bandwidth, B_{rms} is the mean square root bandwidth, and T_w is the duration of the NV signal. As shown in Table 51.2,

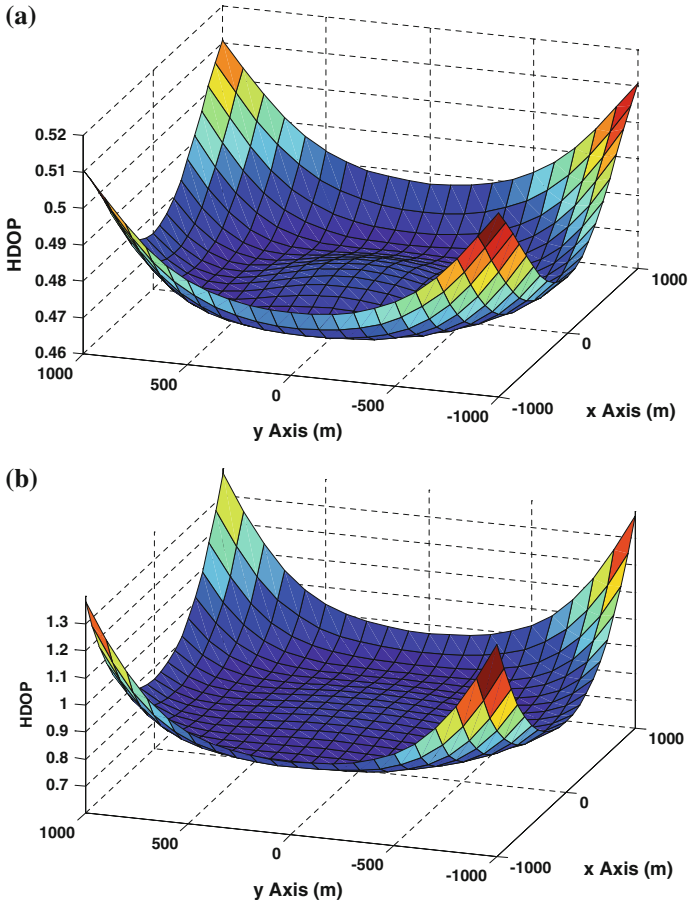


Fig. 51.5 HDOP of the center cell **a** 2 tiers of BSs are considered **b** 1 tier of BSs are considered

$T_w = 0.6807$ ms under subframe configuration 0, and $T_w = 0.0380$ ms under subframe configuration 4.

The TPRMP of each BS are shown in Fig. 51.7 for the special sub-frame configuration 0 and 4. UEs locate at both the center and the edge of the cell are considered. For sub-frame configuration 0, the TPRMP for center UE is about 0.4 and 1.2 m for the BSs of inner tier and outer tier. The TPRMP for edge UE is about 1.5 and 3 m for the BS of inner tier and outer tier. And for sub-frame configuration 0, the TPRMP for center UE is about 1 and 5 m for the BSs of inner tier and outer tier. The TPRMP for edge UE is about 4 and 10 m for the BS of inner tier and outer tier.

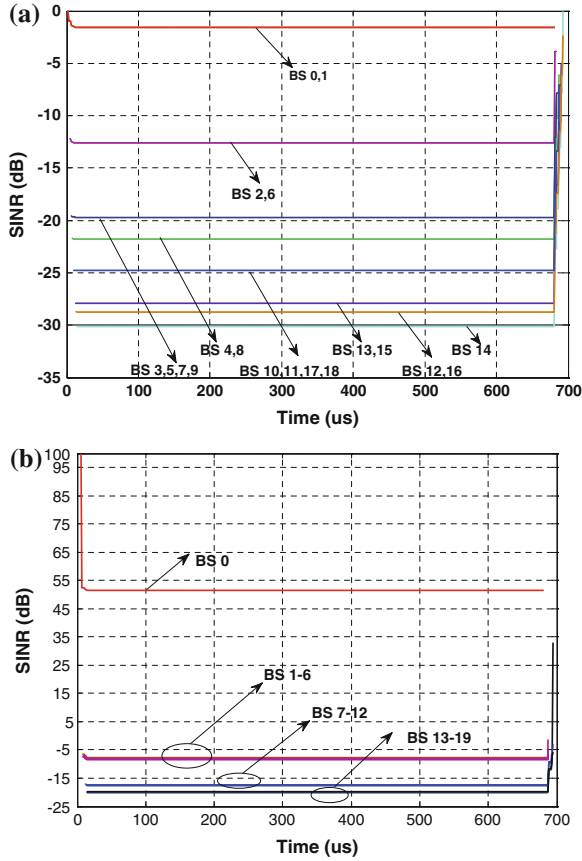


Fig. 51.6 SINR with time for cell edge and cell center UE a Cell edge UE b Cell center UE

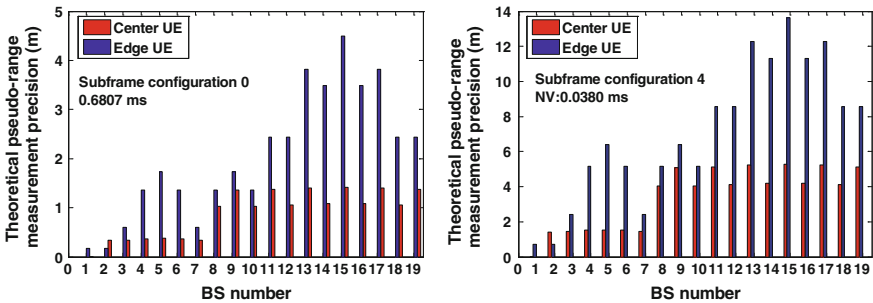


Fig. 51.7 Theoretical pseudo-range measurement precisions

51.6 Conclusion

Based on GP in the special sub-frame of TDD LTE-A system, a novel positing scheme is proposed in this paper by adding navigation signal. Under the proper design of frame structure, we showed that the proposed scheme is compatible with LTE-A protocol, and the collision with LTE-A signal can be prevented. Additionally, simulation results of DOP value and theoretical pseudo-range measurement precision showed that the scheme should be an appropriate solution for positing.

References

1. Sayed AH, Tarighat A, Khajehnouri N (2005) Network-based wireless location: challenges faced in developing techniques for accurate wireless location information. *Sig Process Mag IEEE* 22(4):24–40
2. Wigren T, Wennervirta J (2009) RTT positioning in WCDMA. In: Fifth international conference on wireless and mobile communications ICWMC'09, IEEE, pp 303–308
3. Wigren T (2007) Adaptive enhanced cell-ID fingerprinting localization by clustering of precise position measurements. *IEEE Trans Veh Technol* 56(5):3199–3209
4. Cong L, Zhuang W (2002) Hybrid TDOA/AOA mobile user location for wideband CDMA cellular systems. *IEEE Trans Wirel Commun* 1(3):439–447
5. 3GPP TS 36.211 V12.3.0 (2014-09) Evolved universal terrestrial radio access (E-UTRA) physical channels and modulation
6. Chang C-L, Juang J-C (2010) Performance analysis of narrowband interference mitigation and near-far resistance scheme for GNSS receivers. *Sig Process* 90(9):2676–2685
7. Picois AV, Samama N (2014) Near-far interference mitigation for pseudolites using double transmission. *IEEE Trans Aerosp Electron Syst* 50(4):2929–2941

Chapter 52

Research on the Pulsar Optimizing Method and the Database Construction

Xiao-Ming Bei, Ping Shuai, Liang-wei Hang and Xin-yuan Zhang

Abstract The pulsar database is the foundation of X-ray pulsar navigation. The pulsar optimizing method is the premise of the pulsar navigation database. In this paper, the three characteristic parameters of “NR coefficient”, “stability factor” and “quality factor” are put forward, which is to consider the influence of the pulsar pulse cycle, radiation flow, pulse width, pulse rate and background radiation. Then, it can realize to prioritization of pulsars, and construct the navigation database. This new method overcomes the lack of an objective evaluation of each pulsar defects in the normal method. And a preferred pulsar database, applying to X-ray pulsar navigation, is obtained.

Keywords Pulsar navigation · Stability factor · Quality facto

52.1 Introduction

X-ray pulsar navigation, which is the most promising breakthrough technology in a long-time and high-precision autonomous navigation of spacecraft, has a great value for engineering application and strategic research [1]. In 2004, X-ray pulsar navigation research projects have been put forward by DARPA. Then the feasibility, key technologies and the ground experiment have been completed, and the space flight experiments will be carried out in the international space station and high-orbiting satellites [2–4]. In addition, it has begun to X-ray pulsar navigation theory and experimental research in European Space Agency, Russia, Germany, Japan, Britain, India and Australia [5–9]. From 2005, China has begun to pay close attention to the pulsar navigation technology, carried on the thorough research of principle, database, algorithms, X-ray detector and the ground experimental system [10–26]. It is worth noting that the database is the basis for X-ray pulsar navigation.

X.-M. Bei (✉) · P. Shuai · L. Hang · X. Zhang
Qian Xuesen Laboratory of Space Technology, Beijing 100094, China
e-mail: xiaominbei@gmail.com

The construction of pulsars navigation database includes pulsar selection, catalogue and design. And the optimization of pulsars navigation is the core of building database. Considering the X-ray signal cannot penetrate the dense atmosphere. Using Large-diameter radio telescope, decades of the pulsar observation have been conducted, accumulated a large amount of observed data. Therefore, the scheme, which is the ground observation and secondly space observation, is adopted in building the database. First observation data of radio pulsars are analyzed, obtained their parameters such as angular position, timing model, etc. Then the pulsars' X-ray data is get from space-based X-ray observatory, calibrated radio profile and delay. Finally, the database applying to autonomous spacecraft navigation is obtained.

At present, there are about 2328 pulsars are observed, containing 140 X-ray pulsars. If all these X-ray pulsars are chosen as the navigation source, this will spend a lot of computing time and resources of the space-borne computer, which is not good to quickly identify the pulsar navigation and solve the navigation algorithm. Therefore, it is necessary to give pulsars navigation optimization methods and theories, which can ensure the navigation accuracy and the observation efficiency. The traditional theory of navigation optimization is based on the qualitative description, lack of objective evaluation of each pulsar. In this paper, we put forward three evaluation indexes such as stability factor, SNR coefficient and quality factor, and realize the quantitative evaluation on each pulsar. Then, by sorting and filtering of the pulsars, the pulsar database applicable to X-ray pulsar navigation is obtained.

52.2 Optimization Method

52.2.1 Stability Factor

At the time of pulsars navigation space experiment and application, stability is a core indicator of the pulsar choice, which is a premise to improve the navigation accuracy and timing accuracy. However, it is not given a quantitative evaluation of pulsars stability indicators in the traditional theory of navigation optimization. According to the stability of the atomic clocks, we put forward a pulsar quantitative indicator of stability in this paper. Unlike atomic clocks, stability of the pulsars is not calculated by observing the data but by the characteristic parameters. Considering the relative change in frequency, $\Delta f/f_0$, observed X-ray photon flux, F_X , X-ray background radiation flux, B_X , detector area, A , etc, the representation of the pulsar stability is

$$\begin{aligned}\sigma_y(T) &= 0.16 \cdot \frac{\Delta f}{f_0} \cdot \frac{1}{\text{SNR}} T^{-\frac{1}{2}} \\ &= 0.16 \cdot \frac{\Delta f}{f_0} \cdot \frac{\sqrt{(B_X + F_X(1 - p_f))d + F_X p_f}}{F_X p_f} \cdot \frac{1}{T\sqrt{A}},\end{aligned}\quad (52.1)$$

where d is the duty cycle that is the ratio of the pulse width and period, p_f is the pulsed fraction, and T is the observation time. From the above equation, of the pulsar clock observation decreases linearly with time ($\sigma_y(T) \propto 1/T$), which is different from the stability of atomic clock is inversely proportional to the square root of time ($\sigma_y(T) \propto 1/\sqrt{T}$). It is said the stability of the pulsar clock fell faster than atomic clocks with time. Noted that, for long observation, the stability of the pulsar clock is more superior to atomic clocks.

Ignoring common items of Eq. (52.1), the stability factor of the pulsar can be defined as

$$\sigma_X = \frac{\Delta f}{f_0} \cdot \frac{\sqrt{(B_X + F_X(1 - p_f))d + F_X p_f}}{F_X p_f} \quad (52.2)$$

It is shown that the stability factor is only associated with characteristic parameters of each pulsar, which is the nonpublic part of the pulsar stability (52.1). Further, using the above formula, we can calculate the stability factor for every pulsars, and present the stability factor of the top 10 X-ray pulsars in Table 52.1.

It is shown from Table 52.1 that these high quality pulsars have the following characteristics: short period, small frequency error, and low flow etc. Therefore, millisecond pulsars have an advantage to be navigation pulsars.

52.2.2 SNR Coefficient

Under the condition of the given detector area, observation time and background radiation flux, the pulse signal component of the pulsar radiation can be confirmed by measuring the number of photons. But, due to the presence of X-ray photons measurement noise, pulse profiles will deviate from the real profiles. Refer to SNR expression of radio frequency, SNR relates the pulsed component of the signal source photon count, N_{pulsed} , to the one-sigma error in detecting this signal as

$$\text{SNR} = \frac{N_{pulsed}}{\sigma_{noise}} = \frac{F_X p_f \sqrt{AT}}{\sqrt{(B_X + F_X(1 - p_f))d + F_X p_f}}. \quad (52.3)$$

Table 52.1 Navigation pulsars stability factor

Number	Name	Period (s)	Frequency (Hz)	Frequency error	Pulsed fraction (%)	Pulse width (s)	Flux (Ph/cm ² *s) @2–10 keV	Stability factor (σ_X)
1	B1937+21	0.00156	641.92826	1.3E-12	86.0	0.000021	4.99E-05	3.10E-14
2	B1821-24	0.00305	327.40566	2.0E-12	98.0	0.000055	1.93E-04	4.41E-14
3	B1957+20	0.00160	622.12203	1.2E-11	60.0	0.000080	8.31E-05	2.73E-13
4	J0437-4715	0.00575	173.68795	3.0E-12	27.5	0.00029	6.65E-05	4.22E-13
5	B1509-58	0.15023	6.61151	3.0E-12	64.6	0.0027	1.62E-02	4.40E-13
6	B0540-69	0.05037	19.80244	2.0E-11	67.0	0.0025	5.15E-03	1.68E-12
7	J0218+4232	0.00232	430.46107	3.0E-10	73.0	0.000350	6.65E-05	1.00E-11
8	B1823-13	0.10145	9.85350	2.0E-11	10	0.0018	2.63E-03	1.24E-11
9	B0531+21	0.03340	30.22544	1.0E-07	70.0	0.00167	1.54E+00	3.11E-10
10	J1124-5916	0.13531	7.38133	9.0E-09	10	0.0025	1.70E-03	9.30E-09

It will be seen that the SNR can be determined from the navigation source due to observed X-ray photon flux, and X-ray background radiation flux. During the duty cycle, background radiation and the non-pulse component contributes to the noise. Ignoring common items of Eq. (52.3), it follow immediately that the SNR coefficient may be written as

$$SNR_X = \frac{F_X p_f}{\sqrt{(B_X + F_X(1 - p_f))d + F_X p_f}}. \quad (52.4)$$

This show that the SNR coefficient is the constant term of the SNR, which is only related to X-ray photon flux, Pulsed fraction, duty cycle and background radiation flux. So it can be used as one of the evaluation standard of pulsars.

52.2.3 Quality Factor

Regardless of the pulse profile features, the TOA accuracy depends on the pulse signal quality and profile shape. Using the SNR expression (52.3), the TOA variance can be expressed as

$$\begin{aligned} \sigma_{TOA}^2 &= \frac{1}{4} \frac{W^2}{SNR^2} \\ &= \frac{W^2 B_X d}{4(F_X p_f)^2 AT} + \frac{W^2 F_X ((1 - p_f)d + p_f)}{4(F_X p_f)^2 AT}, \end{aligned} \quad (52.5)$$

Here it is assuming that the pulse shape is Gaussian distribution and pulse width is 2σ . It follows at once that quality factor given by the expression

$$\begin{aligned} Q_X &= \frac{1}{\sigma_{TOA}^2} \\ &= \frac{(F_X p_f)^2}{W^2 B_X d} + \frac{F_X p_f^2}{W^2 ((1 - p_f)d + p_f)} \end{aligned} \quad (52.6)$$

Assuming that the background photons are isotropic, the first item of above Eq. (52.6) containing background noise can be removed. For this case the formula (52.6) may be expressed in the form

$$Q'_x = \frac{F_X p_f^2}{W^2 [p_f + \frac{W}{P}(1 - p_f)]}. \quad (52.7)$$

Table 52.2 Navigation pulsars SNR coefficient and quality factor

Number	Name	Period (s)	Pulsed fraction (%)	Pulse width (s)	Flux (Ph/cm ² *s) @2–10 keV	SNR coefficient (SNR _X)	Quality factor (Q _X /Q _{crab})
1	B0531+21	0.03340	70.0	0.00167	1.54E+00	10.6484	1.0
2	B1937+21	0.00156	86.0	0.000021	4.99E-05	0.0654	0.26
3	B1821-24	0.00305	98.0	0.000055	1.93E-04	0.1384	0.17
4	B1957+20	0.00160	60.0	0.000080	8.31E-05	0.0706	0.020
5	B1509-58	0.15023	64.6	0.0027	1.62E-02	1.0321	0.0037
6	B0540-69	0.05037	67.0	0.0025	5.15E-03	0.6021	0.0014
7	J0218+4232	0.00232	73.0	0.000350	6.65E-05	0.0694	0.00099
8	J0437-4715	0.00575	27.5	0.00029	6.65E-05	0.0410	0.00052
9	B1823-13	0.10145	10	0.0018	2.63E-03	0.1632	0.00018
10	J1124-5916	0.13531	10	0.0025	1.70E-03	0.1311	0.00006

Background noise term is ignored, but the pulsar sorting by Q'_x are consistent with these by Q_x . Normalizing by the Q_{crab} value for the Crab Pulsar (PSR B0531+21), this quality factor can be dimensionless as

$$Q = \frac{Q'_x}{Q_{crab}}. \quad (52.8)$$

Using this expression (52.8) quality factor provides a means to evaluate and rank sources, given the top 10 X-ray pulsars in Table 52.2.

52.3 Conclusion

Pulsar database is the foundation of X-ray pulsar navigation. In this paper, we have deeply studied to the quantitative evaluation of Pulsars. First, three evaluation indexes, including stability factor, SNR coefficient and quality factor, have been proposed. Then, by means of these methods, we can evaluate and rank sources of X-ray pulsars, and give the top 10 X-ray pulsars.

References

1. Shuai P, Li M, Chen SL, Huang Z (2009) Principles and techniques of X-ray pulsar-based navigation system. China Astronautic Publishing House, Beijing, p 14
2. Hanson JE (1996) Ph.D. Dissertation, Stanford University, Stanford
3. Sheikh SI (2005) Ph.D. Dissertation, University of Maryland, Maryland
4. Emadzadeh AA, Speyer JL (2011) Navigation in space by X-ray pulsars. Springer, Berlin
5. Sala J, Urruela A, Villares X et al (2004) Ariadna study 03 4202
6. Neutron stars to become space guides. <http://rt.com/news/sci-tech/neutron-stars-to-become-space-guides/>. Accessed 14 Feb 2013
7. Bernhardt MG, Prinz T, Becker W (2010) In: Proceedings of high time resolution astrophysics IV, PoS(HTRA-IV)050, Agios Nikolaos, Crete, Greece, pp 1–5
8. Bernhardt MG, Becker W, Prinz T et al (2011) In: IEEE Proceedings of 2nd international conference on space technology athens, Greece, pp 1–4
9. Becker W (2009) X-ray emission from pulsars and neutron stars, in neutron stars and pulsars. Springer, Berlin, pp 91–140
10. Zhao M, Huang TY (2009) Sci China (G) 39:1671
11. Fei BJ, Sun WJ, Pan GT, Ji CX (2010) Chin J Space Sci 30:85
12. Li JX, Ke XZ 2009 Sci Sin G Physica Mech Astron 39:311
13. Huang LW, Liang B, Zhang T et al (2012) Sci China Ser G Physics Mech Astron 55(3):527–539
14. Sun SM, Zheng W, Tang GJ, Li M (2009) Acta Astron Sin 50:325
15. Mao Y (2009) Ph.D. Dissertation, PLA Information Engineering University, Zhengzhou
16. Zhou QY, Ji JF, Wang T (2013) Acta Phys Sin 62:019701
17. Huang LW, Liang B, Zhang T (2013) Sci China Ser G Phys Mech Astron 56:848–858
18. Wang WB, Zhang XD, Wang XL (2013) Acta Phys Sin 62:069701

19. Wang L, Xu LP, Zhang H, Luo N (2013) *Acta Phys Sin* 62:139702
20. Wang WB, Wang XL (2013) *Acta Phys Sin* 62:209701
21. Hu HJ, Zhao BS, Sheng LZ, Sai XF, Yan QR, Chen BM, Wang P (2012) *Acta Phys Sin* 61:019701
22. Su Z, Xu LP, Wang T (2011) *Acta Phys Sin* 60:119701
23. Hu HJ, Zhao BS, Sheng LZ, Yan QR (2011) *Acta Phys Sin* 60:029701
24. Sun HF, Xie K, Li XP, Fang HY, Liu XP, Fu LZ, Sun HJ, Xue MF (2013) *Acta Phys Sin* 62:109701
25. Sheng LZ, Zhao BS, Wu JJ, Zhou F, Song J, Liu YA, Shen JS, Yan QR, Deng NQ, Hu HJ (2013) *Acta Phys Sin* 62:129702
26. ZhengW, Sun SM, Tang GJ (2010) In: 1st China satellite navigation conference. Beijing, China, 18–22 May 2010

Chapter 53

Study on the Transportation of Electrons in the Graphene-Based X-Ray Detector

Yaojun Wu, Ping Shuai, Hengbin Zhang, Qian Zhang
and Lingzhong Fu

Abstract As a kind of two-dimensional materials, graphene has many excellent characteristics, such as zero band-gap, high electron mobility, high thermal conductivity and high mechanical strength, which make it widely applied in the field of photoelectric detection. X-rays can be detected based on the electric field effect of graphene. When the detector working, electron-hole pairs are generated in the substrate by the incident X-rays. The drift of the electrons in the substrate will change output of the detector. Therefore, it is necessary to study the transportation of electrons in the graphene-based X-ray detector. The motion of electrons in the detector is simulated using finite element method, and the influence of the electrons' motion to the detector output is analyzed. Dependence of the detector output on the incident position of X-rays is studied. Also, the influence of structure parameters to the performance of the detector is analyzed. Based on the results of this study, performance of the graphene-based X-ray detector can be optimized.

Keywords Pulsar-based navigation · Graphene · X-ray detection · Transportation of electrons · Finite element method

53.1 Introduction

The pulsars are rapidly rotating neutron stars, irradiating X-rays at an extremely stable period. The X-ray signal irradiated by the pulsar can be detected by a detector to obtain the pulse time of arrival and image information, with which the navigation parameters like position, velocity, time and attitude of spacecraft can be calculated. Therefore, the X-ray pulsar-based navigation is a promising way to long time and high precision autonomous navigation for spacecraft, which has great application

Y. Wu (✉) · P. Shuai · H. Zhang · Q. Zhang · L. Fu
Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology,
No. 104 Youyi Road, Haidian District, Beijing 100094, China
e-mail: wuyaojun03@163.com

prospect and has become the research focus of space technology [1–4]. It requires the X-ray detector with high energy resolution, high time resolution, large area, light weight, small volume, the certain ability of imaging, without cryogenic cooling and other characteristics. However, X-ray detectors such as gas proportional counter, NaI (Tl) scintillation detector, Si (Li) detector, high purity germanium detector and CCD detector is hard to be applied in the pulsar-based navigation due to some inherent defects. Micro-channel plate detector, silicon drift detector and swept charge device has good performance relatively, and can satisfy the need of pulsar-based navigation to a certain extent. But for the practical engineering applications of pulsar-based navigation in the future, a novel X-ray detector with better performance is needed. The graphene-based X-ray detector is expected to meet the requirements of application of pulsar-based navigation.

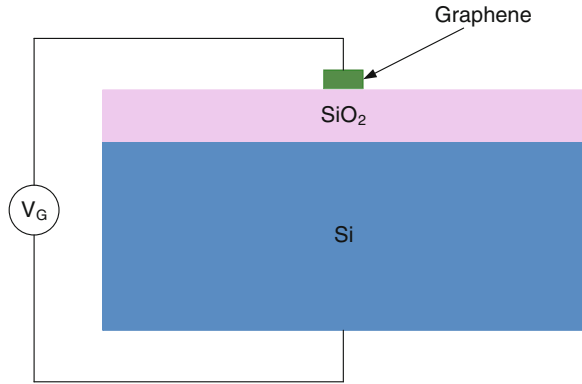
Graphene is a new material composed of monolayer carbon atoms with two-dimensional honeycomb lattice structure [5, 6]. In 2004, Novoselov and Geim prepared graphene for the first time by mechanical exfoliation of pyrolytic graphite, and therefore won the Nobel Prize in Physics in 2010 [7]. Graphene has excellent characteristics such as zero band-gap, high electron mobility, low resistivity, high thermal conductivity and high mechanical strength. It can be used for the development of graphene-based photoelectric detector for terahertz, infrared, visible light, ultraviolet and high energy radiation (X/gamma rays). Vicarelli made a THz wave detector based on graphene field effect transistor, and realized detection of 0.3 THz wave at room temperature [8]. Ryzhii studied the terahertz and infrared detector based on PIN junction of multilayer graphene, and it was very sensitive for its high quantum efficiency [9]. Son studied the ultraviolet detector based on ZnO quantum dots and graphene [10]. Koybasi fabricated graphene field effect transistor detector on SiC substrate, and studied the response of detector to light under different gate voltage [11].

Here we present a model of carrier transportation, and analyze the influence of drift motion of electrons and holes to the output signal of graphene-based X-ray detector. Meanwhile, the influence of incident position of X-rays and structure parameters of the detector to risetime of the output signal are studied. The research in this paper can be used as guidance for the design of graphene-based X-ray detector.

53.2 Theoretical Models

Figure 53.1 is the schematic diagram of architecture for graphene-based X-ray detector. The detector is composed of detection layer of graphene, insulating layer of SiO₂ and absorption substrate of Si semiconductor. When the detector is working, incident X-rays ionize the Si absorption substrate and creates electron-hole pairs. Under the electric field caused by the gate voltage V_G , the holes move downward and are collected by the back electrode, while the electrons move upward and build up under the SiO₂ insulation layer. The electrons generated at different locations by

Fig. 53.1 Schematic of the detector architecture



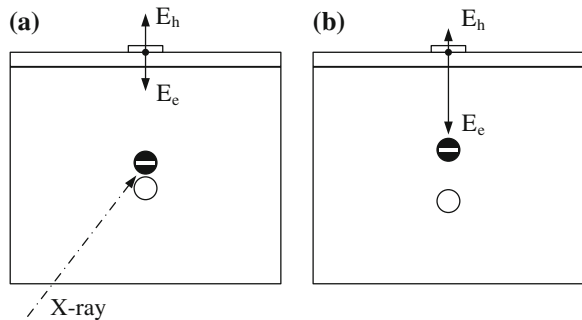
incident X-rays are then gathering at the same distance to graphene layer, which will change the electric field strength of graphene. According to the electric field effect of graphene, the resistance value of graphene will change along with the electric field strength experienced by graphene. Through the real-time monitoring of the resistance value of graphene, the detection of X-rays can be realized.

During the movement electrons and holes, the electrons are moving closer to the graphene, while the holes are moving farther away from the graphene. The drift of electrons and holes will change the electric field strength experienced by the graphene. Therefore, it is necessary to study the effects of transportation of electrons and holes to the output of detector, and analyze the characteristics of the output signal.

53.2.1 Transportation Model of Carriers

As mentioned above, the incident X-rays creates electron-hole pairs in the substrate of the detector and the drift motion of the electrons and holes changes the electric field on graphene. This process is schematically shown in Fig. 53.2. The detector

Fig. 53.2 Electric field on graphene by electrons and holes



structure in Fig. 53.2 is the same as Fig. 53.1. The electron-hole pairs are just generated in Fig. 53.2a, while the electrons and holes have drifted for a distance under the force caused by electric field in Fig. 53.2b. E_h stands for the electric field on graphene generated by holes, while E_e stands for the electric field on graphene generated by electrons. When the electron-hole pairs are just generated, the electrons and holes are almost at the same position in the detector substrate, so E_h and E_e has the same value and opposite direction. As a result, the total electric field strength is zero in Fig. 53.2a. E_e is gradually increasing as the electrons move closer to graphene, while E_h is gradually decreasing as the holes move farther away from the graphene. As a result, the total electric field strength is becoming larger and larger in Fig. 53.2b.

The electric field strength on the graphene generated by the electrons and holes can be written as:

$$\begin{cases} E_e = \frac{N_0 e}{4\pi\epsilon r_1^2} \\ E_h = \frac{-N_0 e}{4\pi\epsilon r_2^2} \end{cases} \quad (53.1)$$

where E_e is the electric field generated by electrons and E_h is the electric field generated by holes. e is the electron charge. ϵ is the dielectric constant of the SiO₂ layer. r_1 is the distance between the electrons and the graphene, while r_2 is the distance between the holes and the graphene. N_0 is the number of the electron-hole pairs generated in Si substrate by incident X-rays (here we assumed that the electrons and holes still keep together during the drift process). The downward direction is prescribed as the positive direction in Fig. 53.2.

The drift velocity of electrons and holes in the Si substrate is given by:

$$\begin{cases} v_e = \mu_e E_G \\ v_h = \mu_h E_G \end{cases} \quad (53.2)$$

where v_e represents the drift velocity of electrons, while v_h represents the drift velocity of holes. μ_e is the electron mobility, while μ_h is the hole mobility. E_G represents the electric field generated by the gate voltage.

Assuming that the distance between the position where electron-hole pairs generated at the very start by incident X-rays and graphene is r_0 . Then the distance between the electrons or holes and the graphene vary with time as:

$$\begin{cases} r_1 = r_0 - v_e t \\ r_2 = r_0 + v_h t \end{cases} \quad (53.3)$$

By Eqs. (53.1), (53.2) and (53.3), the total electric field strength on graphene can be written as:

$$\begin{aligned}
 E &= E_e + E_h \\
 &= \frac{N_0 e}{4\pi\epsilon} \left[\frac{1}{(r_0 - \mu_e E_G t)^2} - \frac{1}{(r_0 + \mu_h E_G t)^2} \right]
 \end{aligned}
 \tag{53.4}$$

53.2.2 Effects of Carrier Transportation on Output

Assuming that energy of the incident X-ray photon is 10 keV, the average electron-hole creation energy of Si is 3.61 eV, so N_0 equals 2770. The value of e is 1.6×10^{-19} C. The value of ϵ is $4.0 \times 8.854 \times 10^{-12}$ F/m. Here we assume that the thickness of the Si substrate is 400 μm . r_0 is 200 μm . μ_e is 1450 $\text{cm}^2/(\text{V s})$. μ_h is 500 $\text{cm}^2/(\text{V s})$. The gate voltage V_G is -50 V, so E_G is 1.125×10^5 V/m. Substituting these parameters into Eq. (53.4), the total electric field strength on graphene varying with time can be obtained, as shown in Fig. 53.3. When the change of electric field strength on the graphene is small, the change the resistance ΔR is approximately proportional to the change of electric field strength ΔE of graphene. Therefore, the curve in Fig. 53.3 can also be considered as the output signal of the detector varying with time. As can be seen from Fig. 53.3, when the electrons and holes just begin to drift, the total electric field strength on graphene caused by carriers is small and changes slowly. When the electrons move close enough to the graphene, the electric field strength on graphene increase sharply. At around 11 ns, the electrons reach the SiO_2 layer and no longer continue to move towards the graphene, so the electric field strength on graphene reaches the maximum value.

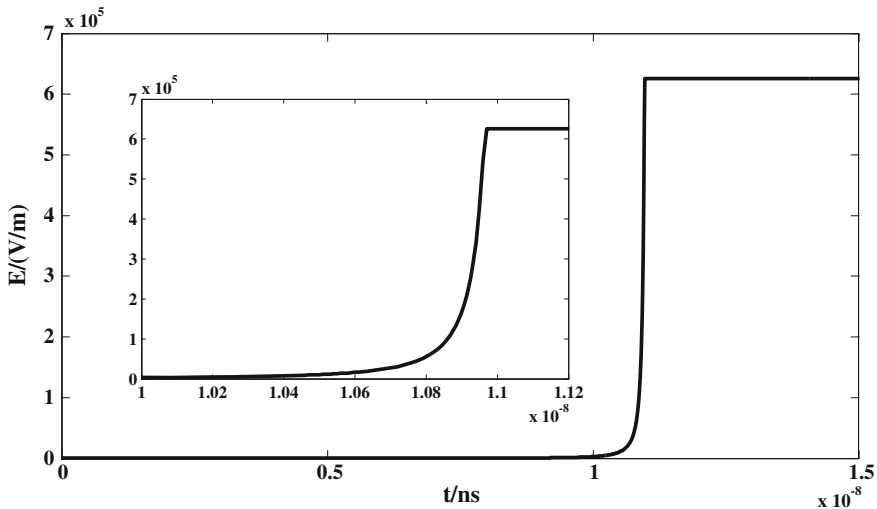


Fig. 53.3 Electric field on graphene

After the time of 11 ns, the varying of electric field strength with time is almost negligible, corresponding with the movement of the holes to the back electrode. By the Eq. (53.4), the electric field strength is inversely proportional to the square of the distance between carriers and the graphene, so the electric field strength on graphene caused by holes is becoming smaller and smaller when the holes are moving away from the graphene.

In the above calculation, the initial position of electrons and holes is assumed to be in the center of detector substrate, which makes the drift path of electrons and holes be the shortest, and the electric field strength along the drift path the largest. As a result, the drift time of electrons and holes in the detector substrate is the shortest under the above assumption, so the risetime of actual output signal is longer than the curves in Fig. 53.3.

53.3 Numerical Calculations

The gate voltage V_G applied between the graphene and the back electrode of detector generates an electric field distribution inside the Si substrate. Due to the dimension difference between the graphene and the back electrode, the electric field formed within the substrate is not uniform. The electric field strength is strong in the center region, while weak in the outer region, resulting different drift time of carriers generated at different location in the detector substrate by incident X-ray. Therefore, the risetime of the detector output signal is not constant, which will affect the results of detection. Under the non-uniform electric field in the detector substrate, the transportation question of the carriers is difficult to solve theoretically, which could only be analyzed numerically using the finite element method. The characteristics of output signals under different incident position of X-rays and the influence of structural parameters to the detector performance will be researched in the following section.

The parameters used in calculation are as follows: dimensions of the Si substrate are $400 \times 400 \times 400 \mu\text{m}^3$. The electron mobility of Si is $1450 \text{ cm}^2/(\text{V s})$. The relative dielectric constant of Si is 11.7. Dimensions of the SiO_2 insulating layer are $400 \times 400 \times 0.5 \mu\text{m}^3$. The relative dielectric constant of SiO_2 is 4. Graphene detection layer is located on the centre of SiO_2 insulating layer, and is simplified as thin sheet of $10 \times 10 \mu\text{m}^2$. The potential on graphene is set to 0 V, while potential on the back electrode of detector is set to -50 V . As the structure of the detector is symmetrical around the central axis, in order to reduce the amount of calculation, the structure of the detector will be simplified as two-dimensional in simulation.

Table 53.1 Drift time of electrons under different incident position

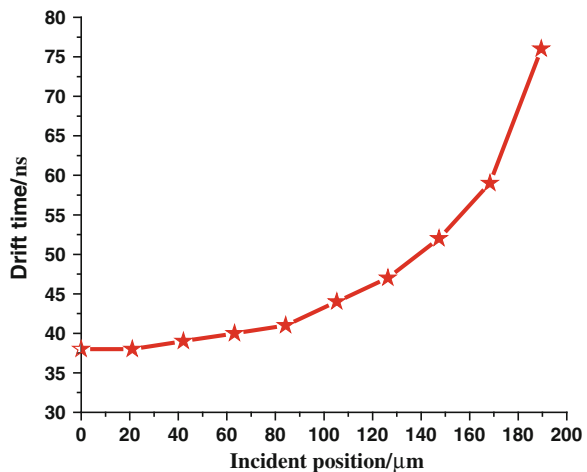
	Incident position (μm)	Drift time (ns)
1	0.0	38.0
2	21.1	38.0
3	42.1	39.0
4	63.2	40.0
5	84.2	41.0
6	105.3	44.0
7	126.3	47.0
8	147.4	52.0
9	168.4	59.0
10	189.5	76.0

53.3.1 Effects of X-Ray Incident Position on the Drift Time

When the incident X-rays generate electron-hole pairs at different positions within Si substrate, the time for the electrons drift to the graphene layer are different. To study the electron drift time under the worst case (the longest possible drift time of electrons), we assume that the electron-hole pairs are generated just as the X-rays enter into the detector substrate. Table 53.1 presents the calculated longest drift times of electrons at different incident positions of X-rays. The value of incident position of X-ray is the horizontal distance between the location where electron-hole pairs are generated and the center of the detector. As can be seen from Table 53.1, the farther the position of incident X-ray away from the detector center, the longer time the electrons drift. The drift time is 38.0 ns when the electron-hole pairs are generated in the center position, while 76.0 ns in the most outside position.

Figure 53.4 presents the varying of drift time of electrons with the incident positions of X-rays. When the distances between the incident positions of X-rays

Fig. 53.4 Drift time of electrons under different incident positions

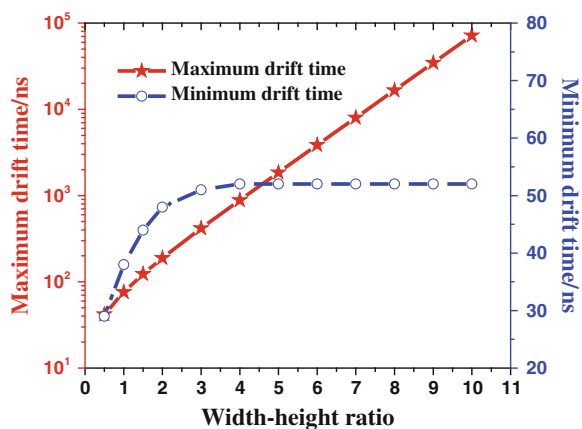


and the center of detector are from 0 to 100 μm , the drift time of electrons are between 38.0 and 41.0 ns, for the distribution of electric field in the region near the center within 100 μm is almost uniform. Therefore the electron drift times in this region are almost the same. When the distances between the incident positions of X-rays and the center are bigger than 100 μm , the electron drift times increase quickly along with the distance from the center. This is because the electric field strength becomes weaker as the incident position of X-ray moves outward, resulting the decreasing of the drift velocity of the electron. Moreover, the drift path in the outer region is longer than the center region. These two reasons together cause the rapid increase of the electron drift time. It also can be seen from Fig. 53.4 that the electron drift times increase with the incident positions as the square law.

53.3.2 Effects of Structure Parameters on the Drift Time

The structure parameters such as length, width and height will change the distribution of the electric field in the detector substrate and the length of drift path, thus affecting the electron drift time. Because the structure of the detector is symmetrical around the central axis, the length and width plays the same role in the influence on the drift time, so the detector is simplified to a two-dimensional structure in the simulation. The electron drift times under different width-height ratios are studied here. We also assume that the electron-hole pairs are generated just as the X-rays enter into the detector substrate to face the worst case. The height of Si substrate of detector is fixed at 400 μm , while the widths of the detector substrate are respectively set as 200, 400, 600, 800, 1200, 1600, 2000, 2400, 2800, 3200, 3600 and 4000 μm , corresponding to width-height ratio of 0.5, 1, 1.5, 2, 3, 4, 5, 6, 7, 8, 9 and 10. Figure 53.5 shows the calculated results.

Fig. 53.5 Drift time of electrons under different width-height ratio



In Fig. 53.5, the Y-axis on the left side represents the case of maximum electron drift time (when the incident positions of the X-rays are at the most outer position of the detector), while Y-axis on the right side represents the case of minimum electron drift time (when the incident positions of the X-rays are at the center of the detector). As can be seen from Fig. 53.5, the maximum electron drift time rises exponentially with the increment of the width-height ratio, for the electric field strength is weaker in the outer region of the detector. The minimum electron drift time increases with width-height ratio firstly, and remains almost constant after the width-height ratio reaches 4. When the width of the detector reaches a certain value, the influence of the increment of width to the electric field strength in the center area can almost be negligible, so the minimum electron drift time does not increase anymore. From Fig. 53.5, we can also see that if the maximum electron drift time of the detector is expected to keep below 100 ns, the width-height ratio of the detector should be designed less than 1.

53.4 Conclusions

The electrons and holes generated by the incident X-rays drift in the substrate under the electric field caused by the gate voltage, which will affect the output of the graphene-based detector. The theoretic formula of the total electric field strength on graphene varying with the drift of electrons and holes is derived, with which the output signal of the detector is calculated. It is found that the risetime of the output signal is mainly influenced by the drift of electrons. And the risetime of the graphene-based detector is in the order of 10 ns. The influence of incident position of X-rays and the structural parameters to the electron drift time is studied using the finite element method. It is found that the electron drift times increase with the incident positions as the square law. As the width-height ratio of the detector increases, the maximum electron drift time rises exponentially, while the minimum electron drift time increases firstly and remains almost constant after the width-height ratio reaches 4. The width-height ratio of the detector should be designed less than 1 to keep the maximum electron drift time below 100 ns. The research in this paper is helpful to design graphene-based X-ray detectors with more excellent performance.

Acknowledgments This work is supported by the National Natural Science Foundation of China (11405265).

References

1. Shuai P, Chen S, Wu Y et al (2007) Advance in X-ray pulsar navigation technology. *Chin J Space Sci* 27(2):169–176
2. Shuai P, Chen S, Wu Y et al (2007) Navigation principles using X-ray pulsars. *J Astronaut* 06:1538–1543

3. Xiong K, Wei C, Liu L (2008) Research on the autonomous navigation of satellite constellation using pulsars. *J Astronaut* 02:545–549
4. Liu J, Ma J, Tian J (2010) Integrated X-ray pulsar and doppler shift navigation. *J Astronaut* 06:1552–1557
5. Geim AK, Novoselov KS (2007) The rise of graphene. *Nat Mater* 6:183–191
6. Geim AK (2009) Graphene: Status and prospects. *Science* 324:1530–1534
7. Novoselov KS, Geim AK, Morozov SV et al (2004) Electric field effect in atomically thin carbon films. *Science* 306:666–669
8. Vicarelli L, Vitiello MS, Coquillat D et al (2012) Graphene field-effect transistors as room-temperature terahertz detectors. *Nat Mater* 11:865–871
9. Ryzhii M, Otsuji T, Mitin V et al (2011) Characteristics of p–i–n terahertz and infrared photodiodes based on multiple graphene layer structures. *Jpn J Appl Phys* 50:070117
10. Son DI, Yang HY, Kim TW et al (2013) photoresponse mechanisms of ultraviolet photodetectors based on colloidal ZnO quantum dot-graphene nanocomposites. *Appl Phys Lett* 102(2):021105
11. Koybasia O, Childres I, Jovanovi I et al (2012) Graphene field effect transistor as a radiation and photo detector. *Proc SPIE* 8373:1–8

Chapter 54

Concepts and Perspectives on Navigation Satellite Autonomous Health Management System Based on Cognitive Technology

Jun Xie, Jianjun Zhang and Ming Xue

Abstract The navigation satellite navigation autonomous health management system based on cognitive technology is an advanced intelligent spatial information processing system. It takes the working mode and mechanism of the human brain as reference, has more complex information and environment sensing ability, adaptively learns according to the information and environment statistic change, and effectively improves “behavioral effect feedback” capabilities of many navigation satellites in the system and the independent operation management capabilities of navigation satellite systems. Therefore it is an important direction for the development of future navigation satellite systems. Aimed at the building requirements of global satellite navigation system constellation configuration and navigation satellite function and based on exploring the cognitive technical background of navigation satellite system, the basic concepts and content of the autonomous health management system of navigation satellite based on cognitive technology are proposed. Through analyzing the existing information flow of navigation satellite, the idea of applying cognitive technology in cognitive technology-based navigation satellite health management system and the key technologies will be discussed and solved, and development recommendations of the cognitive technology-based navigation satellite health management system are given.

Keywords Cognitive technology · Satellite navigation system · Autonomous health management

J. Zhang (✉) · M. Xue
China Academy of Space Technology, Qian Xuesen Laboratory of Space Technology,
Beijing, China
e-mail: zhangjlij@163.com

J. Xie
China Academy of Space Technology, Beijing, China

54.1 Introduction

The satellite navigation system is a large and complex system running in the harsh complex space environment with long-term continuous and stable operation. Despite a series of measures is taken to guarantee the reliability in the design and development process, the navigation satellites in orbit may still not work appropriately inevitably, affecting the health of the entire system. Due to the special nature of the satellite navigation system tasks, when the navigation satellites in orbit has failure, if timely measures are not taken, it could lead to a substantial decline in service performance [1].

Cognitive technology as one of the four frontier science and technology in the 21st century, is in the interface of psychology, computer science, neuroscience, philosophy and other basic science, emerged as a new scientific discipline. Cognitive technology based satellite navigation self-health management system is a more intelligent spatial information processing system. It takes the working mode and mechanism of the human brain as a reference, has more complex information and environmental awareness, and adaptively learns according to information and environmental statistics change. It can also effectively improve the “behavioral feedback effects” capacity of many navigation satellites, is an important direction for future development of GNSS [2].

Based on the analysis of the information flow of navigation satellite, the basic concepts and system framework of navigation satellite self-health management system based on cognitive technology is proposed. The study gives the basic idea, discuss and resolve the key technology, and put forward the development ideas and proposals.

54.2 Model Design of Navigation Satellite Self-health Management System Based on Cognitive Technology

54.2.1 Basic Concepts

Cognitive technology based navigation satellite self-health management system is essentially intelligent information processing system with intelligent error detection, fault tolerance, error correction, based on the operating mode of the human brain. Through perception, analytical understanding, self-learning and memory of the historical and current satellite navigation data status, we get and determine the abnormal condition or fault-related information of a stand-alone device or subsystems, including fault type, fault location, fault occurrence time and fault level, etc. On the basis of fault information locate the fault, the systems is able to reason intelligently, determine decisions and planning, such as the redundant replacement, equipment reconfiguration, system degradation and other measures to ensure the normal running operation of the system, so that navigation satellites can adapt to

complex internal and external environments, efficiently and self-adjust the navigation satellite to fit the inside and external environment. In this way, many navigation satellites in the system have “behavioral effects and feedback” capability [3–5].

The concept of cognitive technology based navigation satellite self-health management system includes three basic elements:

- (1) Intelligent information processing: Its main task is to transmit information, and constantly interact with the internal state of the navigation satellite and the external space environment, then acquire and improve the awareness of navigation satellites on overall operating environment;
- (2) Failure information and related information learning, storage, memory: through failure information, a variety of internal information and space environment is accumulated, in order to improve the precision degree of navigation satellite failure cognition;
- (3) Behavioral effects and feedback: based on behavioral effects, intelligent information processing and intelligent reasoning are performed, and then is sent back to the star transaction processing system, making navigation satellite adaptive based on the fault decision information: if it is single-level fault, we can use reset or reload and other means; if we can still not solve the problem, we can use parts switch, and use the backup to solve; if it is subsystem or system-level fault, we can apply stand-alone switch and crossover recombination, and use the backup to solve; if we cannot solve the problem, the problem can be solved by switching the emergency mode, or by the use of system degradation.

54.2.2 Framework

The law of human brain information processing system provides the ideal foundation reserve for exploration and research of intelligent information processing, it has a bright future. From the level of the system, the cognitive processes of the brain receives original signals through the body senses, re-expresses the signal by the encoding and decoding, which embodies the links of cognitive things attributes and other things, and then is able to continue unchanged preservation and can be repeated used and modified.

Based on the cognitive theory of human brain, this paper presents a navigation satellite autonomous health management system framework (shown as Fig. 54.1), combining the framework model and functional model realizes advantages complementary, and simulates the simplifying system architecture of the human brain’s higher nervous system. The structure includes information perception, collaborative integration, and intelligent decision-making and dynamically reconfigurable components. Information sensing analogies various sensors of the satellite navigation

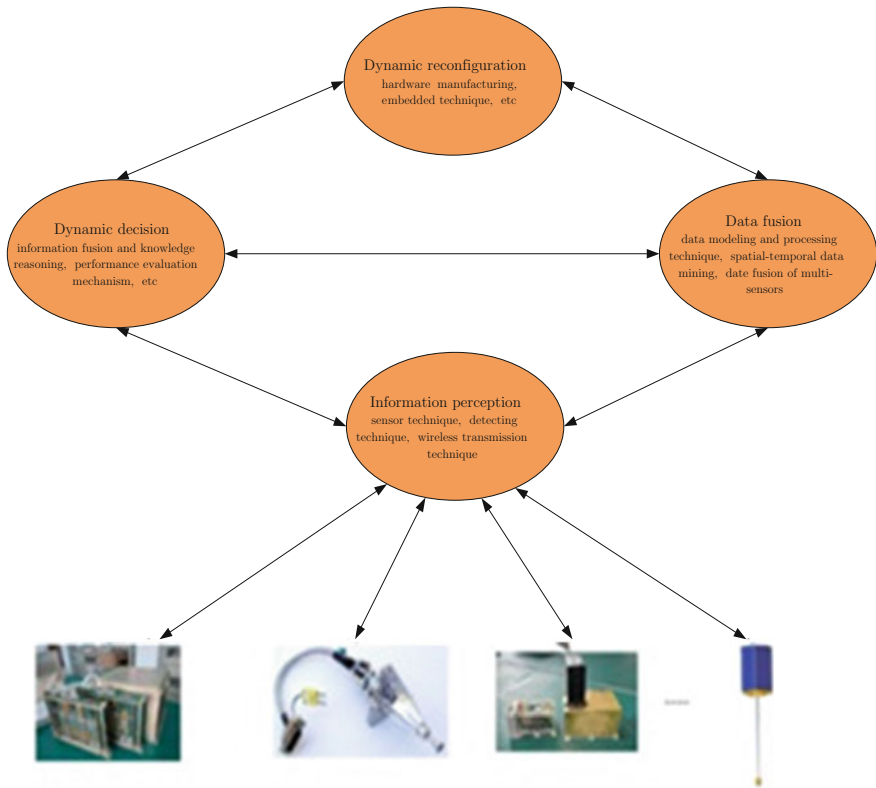


Fig. 54.1 Schematic of navigation satellite self-health management system framework based on cognitive technology

for various neuron receptors in the human body, which's collect raw information about the physical world [6–8].

Synergistic integration is similar to the human thalamus (thalamus is a processing center signals with variety of sensory information, and is responsible for temporal integration and information integration of the incoming signal external from peripheral nervous system) functionality to achieve awareness, recognition, understanding and other “intelligent perception”. Synergistic integration is a reflected process from a single attribute information to the overall information, reflecting the characteristics of a stand-alone or subsystems mutual interaction of navigation satellite, and makes information organize meaningfully.

Intelligent decision is similar to the human brain (cognitive central nervous system of the human brain is consists of interconnected left and right brain) for intelligent information processing, and mainly achieves thinking, decision-making and behavior control, and logical thinking skills, such as Lenovo, memory, learning, analysis, judgment, reasoning, and decision-making. The construction of

cognitive structure is based on object needs and related environmental, and the selection of decision requires search and reasoning based on a priori knowledge.

Dynamic reconfiguration is similar to the human cerebellum (the main function of the cerebellum is to coordinate movement and action, control the operation and posture, and maintain a stable equilibrium, i.e., by interacting with the lower central nervous system and peripheral nervous system to achieve the human body’s movement and the coordinated control of posture) for the information intelligent processing and interactive interface process, which shows behavior intelligence.

In the life evolution process of information perception, collaborative integration, intelligent decision-making, dynamic reconfiguration in the navigation satellite self-health management system is based on cognitive technology. There are a variety of dimensions of information, and different perception environment directly affects the internal information associated strong and weak degree, and has circulation cognitive characteristics. It can perceive the required information of the current navigation satellite state, and make intelligent decisions and dynamic reconfiguration achieve dynamic adaptation to the environment. The information life cycle is “physical data → single parameter information → multidimensional parameter information → multimodal fusion information → parameters associated knowledge → parameter coupling knowledge → dynamic configuration → behavioral control”. As is shown in Fig. 54.2, the navigation satellite self-health management system based on cognitive technology is an effective pass along and is fully applied constant flow direction and evolution of transformation, forming a closed-loop information lifecycle.

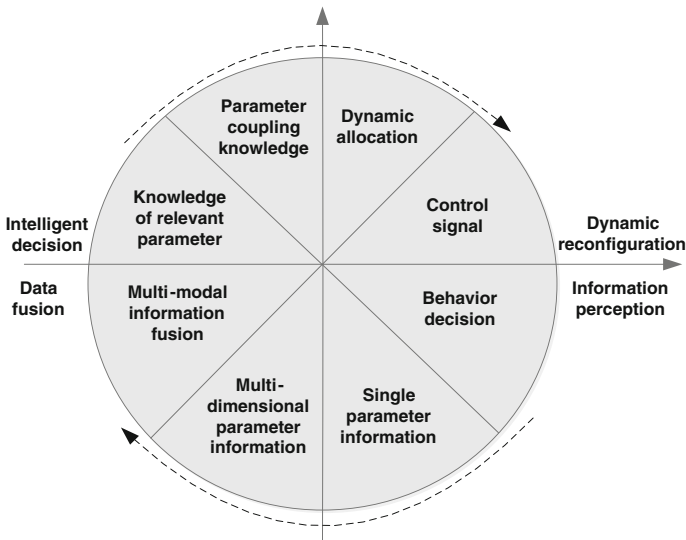


Fig. 54.2 Information life cycle diagram of navigation satellite self-health management system framework based on cognitive technology

54.2.3 Multi-parameter Correlation Modelling

In the reference of the causes, mode, damage and measures of the orbit satellite fault, treating the subsystem as a unit, we not only study a single component or subsystem fault model established methods, but also consider the correlation among fault models. Therefore, the multi-parameter correlation modeling of the satellite navigation is the core part of the self-health management system.

- (1) The node association model of the navigation satellite. The self-health management system connects a variety of stand-alone or subsystems seamlessly using wired network, forming a large network, each node is regarded as an object, as shown in Fig. 54.3. Therefore, each node in the self-health management system is modeled based on the object modeling method.

A given node object can be described by a triple dual, which is the combination of the type of objects, the content attribute of objects and behavior of objects. The object-based modeling method can be used to establish the relationship between the different classes, and classes and node objects, describing the functions and features of the network nodes according to the level of progressive relationship.

$$Object = \{A, B, O\} \tag{54.1}$$

Where $A = \{a_1, a_2, \dots, a_n\}$ represents the types of *Object*. $B = \{b_1, b_2, \dots, b_n\}$ Represents the content attributes of node objects. $O = \{o_1, o_2, \dots, o_n\}$ is the behavior of objects.

- (2) System object graph is modeled based on the cognitive techniques. The cognitive model based on the object graph is not only consistent with the nature of the navigation satellites complex dynamic system, but also meets

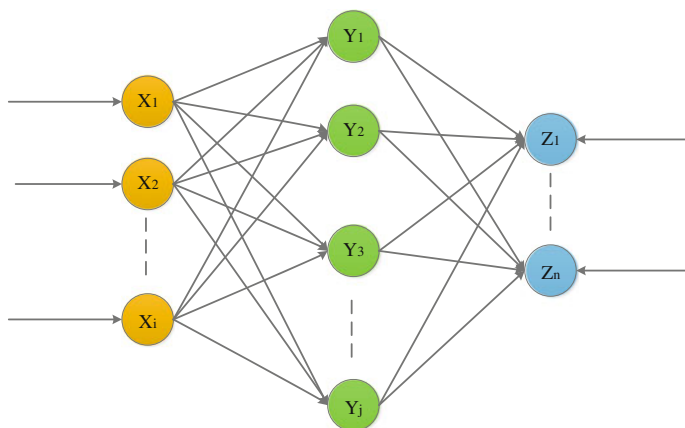


Fig. 54.3 Navigation satellite self-health management system network model

the dynamic requirements from the dynamic system to its cognitive network. Each object in the network association model is defined as 7 parameter function groups:

$$o = (c_n, o_n, OA, DA, E, g, f) \quad (54.2)$$

where c_n is the name of the node type, O_n is the name of the node, $OA = [oa_1, oa_2, \dots, oa_n]$ is the correlation combination of solution of O_n , $DA = \{da_1, da_2, \dots, da_n\}$ represents the correlation type of OA , $E = \{e_1, e_2, \dots, e_n\}$ is the other nodes which is link to O_n , function g represents the corresponding node of each association, function f represents the corresponding type of each association. The parameters in the function set form a very large amount of information, and there are associated cognition among the parameters to different degrees. These associations represent not only the coupling among parameters, but is also able to carry out the next couple of reasoning from the association, and get more implicit knowledge.

54.3 Navigation Satellite Self-health Management System Framework Design Based on Cognitive Techniques

Under the basic concept of the navigation satellite self-health management system based on cognitive techniques and with the analysis of the system framework and the information life cycle, the navigation satellite self-health management system framework based on cognitive techniques is formed, which not only achieves the large closed-loop of navigation satellite—Internal context-aware and intelligent processing—decision, but also achieve closed-loop of application—assessment—update in a knowledge base, as is shown in Fig. 54.4. Taking full account of the structure of the various subsystems, the navigation satellite system integrate all the functions that the health management needs, so that each module can not only collaborate, but also to maintain a clear interface, which is efficient and reliable.

According to the system functions, the navigation satellite self-health management system framework based on cognitive techniques is divided into five parts:

(1) Information perception and acquisition

This part obtains the working condition data and information of various subsystems and stand-alone devices belonging to the navigation satellite using a variety of sensors, as a follow-up data storage data sources.

(2) Fault dynamic database

Dynamic knowledge base provides a priori knowledge for the cognitive model of adaptive fault identification module, and takes the cognitive learning using the feedback information and updates the knowledge base dynamically.

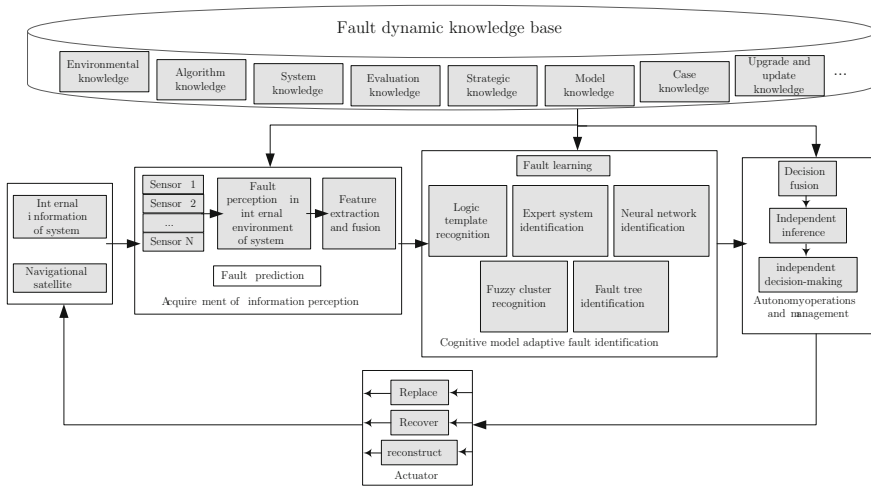


Fig. 54.4 Architecture diagram of navigation satellite self-health management system framework based on cognitive technology

(3) The cognitive model of adaptive fault identification

Combined with the results of data processing, health condition of the entire star and the subsystems is analyzed according to the related algorithms, including fault diagnosis and prediction.

(4) Autonomous operation and management

According to the results of the health assessment, decision on the equipment in which malfunction or failure may occur and subsystems are made. The results and alarm information are then passed to the actuators.

(5) Actuators

It refers to that the self-health management system can send out a fault repair, replacement or reconstruction of instruction independently including switching the machine, status setting, the primary backup switch fault repair, replacement or reconstruction of instruction, according to the result of the fault diagnosis and prediction.

54.4 The Key Issues on the Navigation Satellite Self-health Management System Based on Cognitive Technology

Currently, the issues which urgently need to be solved can be grouped into 5 aspects:

- (1) The navigation satellite self-health management system framework and associated model base on the cognitive technology is still at a preliminary stage, and there is a huge challenge to establish a comprehensive theoretical model of common sense based on the existing self-health management system structures. The reason is that the flight faults of navigation satellites in orbit are difficult to predict, and the architecture presented above is difficult to fully cover all the parameters of satellite navigation.
- (2) Cognition is the ability to explore the unknown environment. The cognitive ability of the self-health management system is driven by a navigation satellite state environmental goal.
- (3) According to current research, a similar study exists abroad, which applies the cognitive technology in a single computer system, and has achieved some success. This allows us to see favorable prospects when the recognition technology is in the use of multi-system level, but there are a lot of difficulties in the navigation satellite system to just rely on the misappropriation of cognitive model applied in stand-alone systems.
- (4) Satellite navigation is a complex system involving multi-disciplinary, multi-technology fields. The complexity of their stand-alone or multiple subsystems runtime state provides a higher demand on cognitive loop system strategy mechanisms of cognitive technology expert systems, artificial intelligence and game theory in the cognitive technology; and various states of the environment need to develop different cognitive appropriate loop strategy.
- (5) The actual work is more complex in various aspects when applying the satellite navigation self-health management technology system based on the cognitive technology. Furthermore, some factors are not repeatable and unpredictable, unless the satellite navigation self-health management technology system based on the cognitive technology is in operation, otherwise the situations in the system cannot be validated assessments one by one, which naturally limits the system design tradeoffs and evaluation of its effectiveness.

54.5 Perspectives on Further Development

First stage: analyze needs, refine requirements, set up the program.

The overall demand of the navigation satellite system in the whole process of flying missions, the navigation satellite self-health management system based on cognitive technology simulation technology should be set up. The objectives and mission requirements navigation satellite autonomous health management to achieve should be understood.

The second stage: break through the key technology, and clarify the technical development path

In the key technology of the navigation satellite self-health management system based on cognitive technology, confidence perception method, collaborative fusion

and intelligent decision are all based on the overall demand analysis. The key technology is the technological base navigation satellite self-health management function based on the cognitive technology, which is both engineering stage results and a reliable guarantee for subsequent engineering implementation.

The third stage: Building simulation system, complete system verification

Research for the navigation satellite self-health management system based on cognitive technology simulation technology is carried out, and navigation satellite ground system simulation for design verification is developed and built, fault diagnosis is achieved, satellite autonomous system is evaluated and system reconstruction and fault forecast technology demonstration are achieved.

54.6 Summary and Recommendations

- (1) The cross fusion of self-health management system and cognitive technology is the trend of the future development of GNSS

From the developing trend of both at home and abroad, the cognitive technology of space information system has made great improvement in performance. Exploration research on cross-integration of the cognitive technology and spatial information system, which foundation is the navigation satellite self-health management system based on cognitive technology, is at the preliminary stage both at home and abroad, and needs to be carried out through in-depth study.

- (2) Strengthening the theoretical model research and tamping the basis of system technology development

The complex navigation satellite self-health management system based on cognitive technology involves many new and difficult method. Although we have taken an important first step, there are still some key theoretical models that need to be broken through, and we should carry out research and development work as early as possible, so that we can break through the theoretical bottlenecks laying the solid theoretical foundation for system technology to practical application.

- (3) Timely demonstration of the ground or in orbit integration, promote system technology to the practical application

On the basis of a breakthrough in the related theoretical model and key technology, we should carry out the ground or in-orbit demonstration as soon as possible. Through the demonstration, we can maintain the continuity of the technology development, realize the rolling development of related parameter model more, timely find problems and deficiencies, accumulate experience, and for the system technology to practical lay a more solid foundation.

- (4) To promote universal model and system design for the system technology in the wide application of spatial information system.

Currently spacecraft design are considering self-health management requirements, but lacks independent and effective line of health management system technology. In the event of failure or abnormal state we only rely on the ground remote telemetry, which also restricted the development of information systems for future space and is a bottleneck in the application.

References

1. Kaplan ED, Hearty CJ (2005) Understanding GPS principles and applications. Artech House Press, Boston
2. Prasad R, Ruggieri M (2005) Applied satellite navigation using GPS, Galileo, and augmentation systems. Artech House Press, Boston
3. Kelley CW, Davis KF (2000) Hybrid GPS constellations to provide high availability of Category 1 precision approach, RAIM and continuity. In: Proceedings of the 13th international technology meeting of the satellite division of the institute of navigation. The Institute of Navigation, Salt Lake, pp 2184–2190
4. Rizos C (2007) The future of global navigation satellite systems. Technical Documents, University of New South Wales
5. Mitola J et al (1999) Cognitive radio: making software radios more personal. *IEEE Pers Commun* 6(4):13–18
6. Haykin S (2005) Cognitive radio: brain-empowered wireless communications. *IEEE J Sel Areas Commun* 23:201–220
7. Haykin S (2005) Cognitive radar networks. In: 1st IEEE workshop on computational advances in multi-sensor adaptive processing. Jalisco State, Mexico
8. Haykin S (2007) Adaptive radar signal processing. Wiley, Hoboken

Chapter 55

An X-Ray Pulsar TOA Estimation Method Considering Spacecraft Orbit Motion

Liangwei Huang, Ping Shuai and Xinyuan Zhang

Abstract Pulse time-of-arrival (TOA) estimation is the key procedure to construct the measurement in X-ray pulsar navigation. Traditional TOA estimation methods are designed assuming spacecraft to be at rest. However, spacecraft orbit motion introduces doppler effect and causes observed pulse period shift, which makes TOA estimation harder. This paper provides a method based on average cross correlation to solve the problem in orbit motion situation. Instead of direct doppler frequency estimation, this idea transplants the static cross correlation method based on epoch folding into dynamic situation by folding at the middle of the observation period to counteract the doppler effect. At certain doppler frequencies, the average cross correlation method is tested via Monte-Carlo simulation. Results show that the proposed method can output high-precision TOA estimation at small accumulated doppler phases.

Keywords X-ray pulsar · Dynamic situation · Pulse TOA estimation · Doppler frequency

55.1 Introduction

X-ray pulsar navigation is a promising technique to realize spacecraft long-time and high-precision autonomous navigation. It is a highly valuable technique appreciated by different astronomical organizations in the world [1]. The pulse time-of-arrival (TOA) is the basic measurement quantity of pulsar navigation, whose precision will decide the navigation accuracy. The pulse TOA measurement can be represented by the pulse phase in the equivalent form for the fact that the pulse TOA equals the phase multiplied by the pulse period. As a benefit, the doppler frequency can be

L. Huang (✉) · P. Shuai · X. Zhang
Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology,
100094 Beijing, China
e-mail: nostalgia@163.com

used to construct spacecraft's velocity measurement equations. Therefore, in the following text, the pulse TOA estimation problem is addressed as the pulse phase measurement problem, which concerns how to obtain high-precision pulse phase estimates and doppler frequency estimates from photon TOA data.

The pulse phase estimation methods are commonly divided into two types: one is based on the epoch folding and another directly uses photon data [2, 3]. As to epoch folding, the cross correlation [4, 5] or the nonlinear least square estimation [2, 6] can be used to compare the folded profile with the standard profile. Directly using of photon data is based on the maximum likelihood principle, which calculates the phase shift by maximizing the likelihood function (LF) [7]. The LF can be maximized using numerical algorithms such as the grid search algorithm [8] and the Newton-iteration algorithm [3] or analytic ways such as the technique based on the first-order harmonic wave approximation of the LF [9, 10]. Above mentioned methods are applicable to phase estimation in the static situation that postulates that spacecraft are at rest on orbit. When spacecraft orbit motion is considered, the observed pulse frequency departs from the pulsar's inherent rotation frequency. This departure is called the doppler shift or the doppler frequency. For pulsar navigation, the doppler frequency is unknown, which denies the pulse TOA or phase estimation. Some literature tries to directly estimate doppler frequencies that is also called observed pulse period searching. Such methods aim to simultaneously obtain the initial pulse phase and the doppler frequency of the observation period, which are called federal estimation [11]. For example, Li [12] designed the pulse period estimation algorithm based on the maximum correlated variance search, and Zhou et al. [13] proposed the improved Lomb algorithm based on fast Fourier transform (FFT) to estimate the pulse period. However, it is difficult for these methods in frequency domain to efficiently estimate the doppler frequency, for the facts holds [14, 15]: (1) the pulse signals are hidden in the background noises so that numerous photon samples are needed to increase the spectral resolution due to the fence effect; (2) the harmonic waves introduced by the sub-peaks of the pulse profiles disperse the spectral lines, which brings extra errors in doppler frequency estimation.

In this paper, an original way is proposed to sheer off direct pulse period searching, which deals with the phase estimation problem in dynamic situations as an extension of the static situations. The average effect is reached by folding at the middle of the observation period to counteract the doppler effect and to estimate the doppler phase of the midpoint. Based on this principle, the average cross correlation (ACC) algorithm is proposed, which folds the photons with inherent period at the middle of the observation to eliminate the folded pulse shift caused by the doppler effect. In this way, although the folded profile will be widened, it suffers from no shift under certain precision, so that the cross correlation can be use to estimate the pulse phase of the midpoint. The Monte-Carlo simulation is made upon the ACC algorithm. Simulation results show that at a specific doppler frequency, the ACC algorithm can output high-precision pulse phase estimation for a limited observation period.

55.2 Pulse Phase and Doppler Frequency Model

If the spacecraft orbital motion is considered, the pulse phase and frequency observed at the spacecraft are different from those observed at the pulsar, which is caused by the phase delay and the doppler effect. The phase at the spacecraft is called the observed pulse phase noted as Φ^X , whereas the pulsar's inherent rotation phase is noted as Φ^P . It takes time Δ for the pulse signal to be transmitted from the pulsar to the spacecraft. This course can be comprehended as the transmission of the same pulse phase from the pulsar to the spacecraft. Thus, the observed pulse phase is the delayed inherent phase:

$$\Phi^X(\tau) = \Phi^P(\tau - \Delta) \quad (55.1)$$

where τ is the time of arrival expressed by the spacecraft proper time. The inherent phase Φ^P is expressed by

$$\Phi^P(T) = \Phi_0^P + f_0\Delta T + 1/2f_1\Delta T^2 + 1/6f_2\Delta T^3 \quad (55.2)$$

where T is the pulsar proper time, Φ_0^P is the initial phase evaluated at $T = E_{\text{FRQ}}$ with E_{FRQ} being the frequency epoch, f_0 is the rotation frequency, f_1 and f_2 are respectively the first order and the second order time derivative of the rotation frequency, and Δ denotes the time elapse with respect to the frequency epoch, i.e., $\Delta T \equiv T - E_{\text{FRQ}}$.

The observed pulse frequency is defined to be time derivative of the observed pulse phase, i.e., $f_o \equiv d\Phi^X/d\tau$, whose reciprocal is the observed pulse period $P_o = 1/f_o$. By taking the time derivatives of both sides of Eq. (55.1), the following expression is gotten:

$$f_o = f_s + f_d \quad (55.3)$$

where f_s is called the source frequency that is the pulsar's current rotation frequency taking account of the spin-down effect:

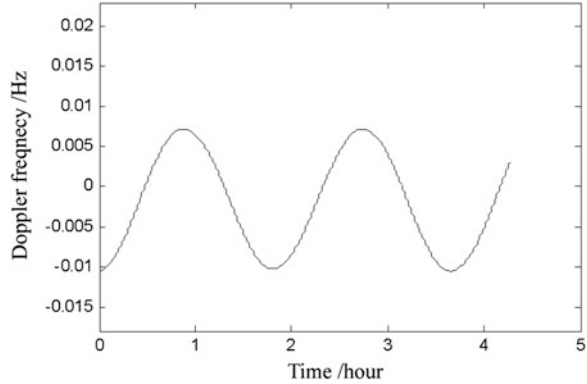
$$f_s = f_0 + f_1\Delta T + f_2/2\Delta T^2 \quad (55.4)$$

Here, $T = \tau - \Delta$, and because f_1 and f_2 are small values, ΔT can be substituted for by $\Delta\tau = \tau - E_{\text{FRQ}}$, which yields

$$f_s = f_0 + f_1\Delta\tau + f_2/2\Delta\tau^2 \quad (55.5)$$

when the pulsar's frequency epoch is old, the $\Delta\tau$ and $\Delta\tau^2$ items cannot be omitted. The quantity f_d is the doppler frequency expressed by

Fig. 55.1 The doppler frequency caused by the spacecraft and pulsar motion



$$f_d = -\left(\frac{d\Delta}{d\tau}\right)f_s \quad (55.6)$$

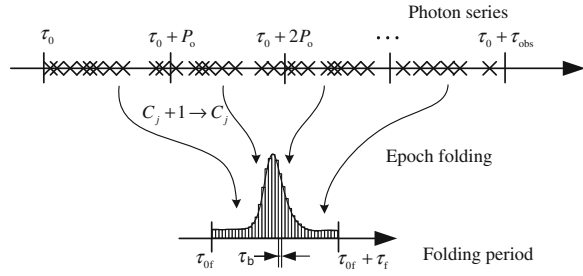
If only the first order doppler effect is considered, there is $d\Delta/d\tau = -c^{-1}v_p$, where v_p is the spacecraft's radial velocity relative to the pulsar. If the pulsar is in a binary system, v_p is composed of the spacecraft's orbital velocity and the pulsar's velocity with respect to the barycenter of the binary system. Figure 55.1 gives an example of the doppler frequency, in which the doppler frequency of a near-Earth spacecraft at the altitude of 600 km observing the binary pulsar PSR B1744-24A is plotted. The pulse period of PSR B1744-24A is 11.56 ms, whose orbital period is as short as 1.8 h. In Fig. 55.1, The doppler frequency comprises three periodic components that are the 1.8 h component of binary motion, the 1.8 h component of spacecraft orbital motion, and the 1 year component of the Earth revolution.

55.3 Cross Correlation

The cross correlation (CC) is used for pulse phase estimation in static situations. The static situation should assume that the pulsar is single and the spacecraft is at rest in the solar system so that no doppler effects exist. The CC algorithm is based on photon epoch folding, which bins the photon epochs and folds them to form a profile. The folded profile is then correlated with the standard profile to generate the phase estimation. The CC is an excellent algorithm for the static situation, because it performs well in precision and efficiency. The procedure of the CC algorithm is as follows.

Given an observation period τ_{obs} started with τ_0 , the epoch folding process is shown in Fig. 55.2. Set the folding period τ_f to be the observed pulse period P_o . Note the start point of the folding period to be τ_{0f} . Separate the folding period into N_b bins, each of which is with the length of τ_b . Calculate the position of the photon TOA in the folding period for each photon in the photon series. If the photon TOA

Fig. 55.2 The procedure of photon epoch folding



is in the j th bin, the count of this bin C_j is added by 1. Then, C_j can represent the folded profile in the discrete form, which can be normalized to be

$$\tilde{h}_j = \frac{P \left(C_j - \min_j C_j \right)}{\tau_b \sum_j \left(C_j - \min_j C_j \right)}, \quad j = 1, 2, \dots, N_b \tag{55.7}$$

where \tilde{h} is the normalized folded profile and $\tilde{h}_j = \tilde{h}[(j - 1/2)/N_b]$ denotes the normalized intensity of the j th bin of the folded profile.

The principle of the cross correlation is to maximize the cross correlation function between the folded profile and the standard profile. Here, the CC function can be expressed in discrete form [3, 16]:

$$R_D(\phi) = \frac{1}{N_b} \sum_{j=1}^{N_b} h_j \tilde{h}_{j+\phi} \tag{55.8}$$

where h_j denotes the normalized intensity of the j th bin of the standard profile, $\tilde{h}_{j+\phi}$ represents the normalized intensity of the j th bin left shifted by ϕ of the folded profile:

$$\begin{aligned} h_j &= h[(j - 0.5)/N_b] \\ \tilde{h}_{j+\phi} &= \tilde{h}[(j - 0.5)/N_b + \phi] \end{aligned} \tag{55.9}$$

The phase delay ϕ is the left shifted amount of the folded profile with respect to the standard profile. The phase delay should be adjusted to maximize $R_D(\phi)$. Suppose the start point of the folding period be the start point of the observation period. The initial pulse phase can be estimated by

$$\tilde{\phi}_0 = \arg \max_{\phi \in (0,1)} R_D(\phi) \tag{55.10}$$

Although ϕ can be searched in time domain, the way in frequency domain is more efficient. The CC function in Eq. (55.10) can be resolved by the FFT technique [17]:

$$R_D(\phi_j) = \text{IFFT}[\text{FFT}(h_j) \cdot \text{FFT}(\tilde{h}_{-j})] \tag{55.11}$$

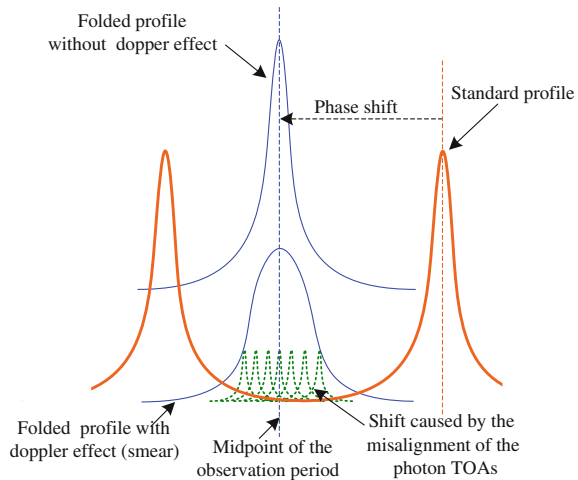
where FFT is the fast Fourier transform, IFFT is the inverse FFT, \tilde{h}_{-j} is the inverted sequence of the folded profile.

55.4 Average Cross Correlation

In dynamic situations, f_d is not zero, so that folding with the inherent period will misalign the photon TOA data and the folded profile will suffer from smear and shift. Mind that the start point of the folding period can be set to the midpoint of the observation period to counteract the profile shift caused by the doppler effect. In this way, although the folded profile is smeared, the phase estimation will be kept by a certain precision (see Fig. 55.3).

Next, an example is given to demonstrate this idea. Observing the Crab pulsar, let the detector area be $A_d = 0.5 \text{ m}^2$, the time resolution be $\tau_d = 100 \mu\text{s}$, the observation period be $\tau_{\text{obs}} = 5 \text{ s}$, the start point be $\tau_0 = 56273.0 \text{ MJD}$, and the initial pulse phase be $\phi_0 = 0.4$. Pick 21 points of f_d from -0.2 to 0.2 Hz to generate 21 groups of simulated photon TOA series. Set the folding period to $\tau_f = 1/f_s = 0.0115631 \text{ s}$, and set the start point of the folding period to $\tau_{\text{of}} = 2.5 \text{ s}$. Then, 21 phases at the midpoints are estimated via CC.

Fig. 55.3 The sketch map of average epoch folding



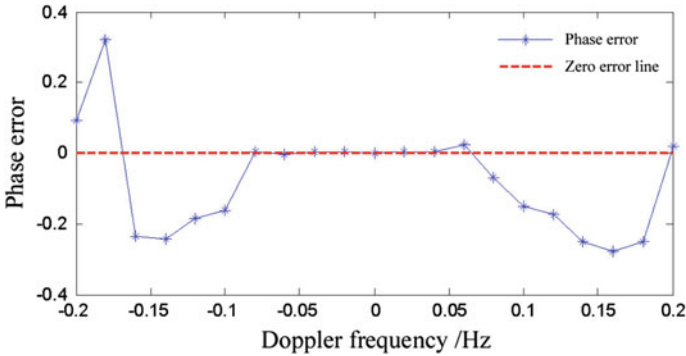


Fig. 55.4 The cross correlation phase estimation error if the epoch folding is made at the midpoint of the observation (PSR B0531+21)

The midpoint phase estimation errors are plotted in Fig. 55.4, which shows that at small doppler frequencies, the phase error is close to zero, and it diverges as the doppler frequency increases. Thus, the average cross correlation algorithm applicable to the dynamic situation is proposed. The algorithm process is as follows.

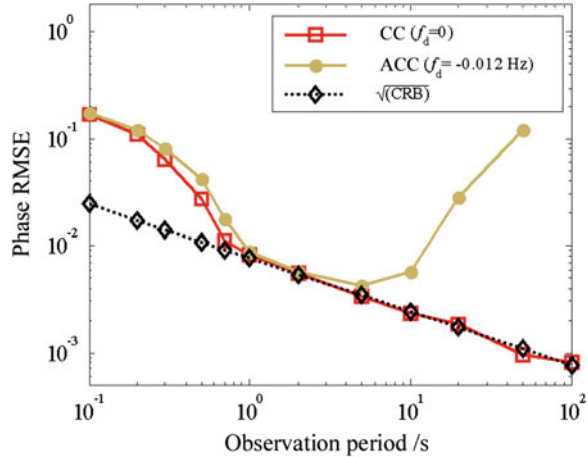
- Step 1 Set the folding period to the pulsar’s inherent rotation period, i.e., $\tau_f = 1/f_s$, set the start point of the folding period to the midpoint of the observation period, i.e., $\tau_{of} = \tau_{mid}$, and execute the epoch folding.
- Step 2 Use the CC algorithm in Sect. 55.2 to estimate the phase at τ_{of} to yield the midpoint phase estimation $\tilde{\phi}_{mid}$ of the observation period.

The ACC algorithm can be enhanced by predicting the doppler frequency. The Earth’s radial velocity and the pulsar’s velocity in the binary system can be computed according to the epochs. Thus, these two parts can be deducted from the doppler frequencies precisely. The spacecraft’s radial velocity can be taken as the current best estimation otherwise zero if no apriori information exists. With the predicted doppler frequency \tilde{f}_d , the folding period in the Step 1 of the ACC algorithm should be set to $\tau_f = 1/(f_s + \tilde{f}_d)$.

55.5 Numerical Analysis

The Monte-Carlo simulation is made that the ACC algorithm estimation error is compared with the Cramer-Rao low bound CRB. The CRB computation can be referred to Ref. [8]. The observation target is selected to be PSR B1744-24A, whose flux parameters are set that the source flux is $R_s = 1.09 \times 10^{-3}$ ph/s/cm², the background flux is $R_b = 5 \times 10^{-3}$ ph/s/cm², and the pulse fraction is $p_f = 0.6$. Set the detector area to $A_d = 10$ m², set the time resolution to $\tau_d = 100$ μ s, and set the

Fig. 55.5 Monte-Carlo simulation results of pulse phase estimation (PSR B1744-24A)



start point of the observation period to $\tau_0 = 56273.0$ MJD. The estimation is executed for 500 times to yield one phase root mean square error (RMSE).

Under the fixed doppler frequency $f_d = -0.012$ Hz, the simulation is done at different observation periods respectively being 0.1, 0.2, 0.3, 0.5, 0.7, 1, 2, 5, 10, 20, 50, and 100 s. The results are plotted in Fig. 55.5. As a comparison, the $\sqrt{\text{CRB}}$ curve and the RMSE curve at $f_d = 0$ are also plotted. It can be seen in Fig. 55.5. that as to the ACC algorithm, when $\tau_{\text{obs}} < 5$ s, the RMSE curve approaches that of $f_d = 0$ in respect that the accumulated phase is at a low level for $\tau_{\text{obs}} < 5$ s. As the observation period lengthens, the error of the ACC algorithm increases significantly.

From the Monte-Carlo simulation results, the following conclusions can be made: (1) The ACC algorithm is applicable to phase estimation in dynamic situations, however, its performance is affected by the doppler phase accumulation in the observation period. (2) The ACC algorithm keeps a high precision when the doppler phase accumulation is at a low level, and the RMSE could well approach $\sqrt{\text{CRB}}$. (3) The ACC algorithm is susceptible to the doppler frequency, whose error increases significantly at great accumulated doppler phases.

55.6 Conclusion

Aiming at the X-ray pulsar navigation mission, this paper discusses the pulse TOA (or phase) estimation techniques considering spacecraft orbit motion. The pulse phase and doppler frequency models are built for the dynamic situation. Based on the static cross correlation method, the average cross correlation method suitable to the on-orbit dynamic situation is proposed. This method elaborately takes advantage of the average effect so that it can avoid direct doppler frequency estimation

and can output the phase estimate at the midpoint of the observation period. Simulation results show that the ACC method performs well in accuracy when the accumulated doppler phase is small, which can well approach the Cramer-Rao low bound. However, the ACC method is susceptible to the doppler frequency. How to reduce the susceptibility of this method is worthy of further research.

Acknowledgments This work was supported by the National Natural Science Foundation of China (Grant No. 61403391).

References

1. Shuai P, Li M, Chen SL et al (2009) Principles and methods for X-ray pulsar navigation system. China Astronautic Publishing House, Beijing (in Chinese)
2. Emadzadeh AA, Speyer JL (2010) On modeling and pulse phase estimation of X-ray pulsars. *IEEE Trans Signal Process* 58(9):4484–4495
3. Emadzadeh AA, Speyer JL (2011) Navigation in space by X-ray pulsars. Springer, Berlin
4. Emadzadeh AA, Speyer JL (2011) X-ray pulsar-based relative navigation using epoch folding. *IEEE Trans Aerosp Electron Syst* 47(4):2317–2328
5. Emadzadeh AA, Lopes CG, Speyer JL (2008) Online time delay estimation of pulsar signals for relative navigation using adaptive Filters. In: *IEEE-ION 2008 position location and navigation symposium*, Monterey, CA, 2008, pp 714–719
6. Emadzadeh AA, Golshan AR, Speyer JL (2009) Consistent estimation of pulse delay for X-ray pulsar based relative navigation. In: *Joint 48th IEEE conference on decision and control and 28th Chinese control conference*, Shanghai, China, pp 1488–1493
7. Emadzadeh AA, Speyer JL (2009) Asymptotically efficient estimation of pulse time delay for X-ray pulsar based relative navigation. In: *AIAA GN&C conference*, Chicago, IL, pp 1–12
8. Golshan AR, Sheikh SI (2007) On pulse phase estimation and tracking of variable celestial X-ray sources. In: *ION 63rd annual meeting*, Cambridge, MA, pp 413–422
9. Sala J, Urruela A, Villares X et al (2004) Feasibility study for a spacecraft navigation system relying on pulsar timing information. *ARIADNA study* 03/4202
10. Order M, Meyr H (1998) Digital filter and square timing recovery. *IEEE Trans Commun* 36(5):605–612
11. Ashby N, Golshan AR (2008) Minimum uncertainties in position and velocity determination using X-ray photons from millisecond pulsars. In: *ION NTM 2008*, San Diego, CA, pp 110–118
12. Li JX (2008) Theoretical research on timing and autonomous positioning based on X-ray pulsar. Ph.D. Dissertation, School of automation and inf eng, Xi'an Univ of techno, Xi'an, China (in Chinese)
13. Zhou QY, Ji JF, Ren HF (2013) Quick search algorithm of X-ray pulsar period based on unevenly spaced timing data. *Acta Phys Sin* 62(1):019701: 1–8 (in Chinese)
14. Zhang H, Xu LP, Xie Q (2011) Modeling and doppler measurement of X-ray pulsar. *Sci China-Phys Mech Astron* 54(6):1068–1076
15. Xie Q, Xu LP, Zhang H et al (2012) Doppler estimation of X-ray pulsar signals based on profile feature. *Journal of Astronautics* 33(9):1301–1307 (in Chinese)
16. Jacovitti G, Scarano G (1993) Discrete time techniques for time delay estimation. *IEEE Trans Signal Process* 41(2):525–533
17. Bracewell RN, Yin QY (trans), Zhang JG (trans) (2005) *The Fourier transform and its applications*, 3rd edn. Xi'an Jiaotong University Press, Xi'an (in Chinese)

Chapter 56

The Research of X-Ray Pulsar Signals Simulation Method

Lingzhong Fu, Ping Shuai, Mengfan Xue, Haifeng Sun
and Haiyan Fang

Abstract X-ray can't be directly observed at laboratory due to the absorption of atmosphere. Furthermore, observation at space consumes a great deal of financial and material resources and is highly risky. Therefore, simulation data of X-ray pulsar is of great importance, whereas the current simulation method can only generate X-ray signals of constant period at the Solar System Barycenter (SSB), without considering the slow varying property of pulsar period and the large scale space-time effects. This paper proposes a new simulation method for X-ray pulsar signal. First, the arrival phase of the next photon at the SSB is recursively calculated according to the phase of the previous photon. Next, seek the pulsar ephemeris to obtain the phase evolution model of the current time. Then, transform the arrival phase of the photon to its arrival time at the SSB using the obtained phase evolution model. Subsequently, utilizing the solar system ephemeris and orbital parameters of the spacecraft, the arrival time of the corresponding photon at the spacecraft can be acquired by an iterative algorithm. Simulation results show the validity of the proposed methods. The simulation data of the proposed method includes the slow varying property of pulsar period and the large scale space-time effects, also is very similar to the RXTE observation data. Thus the proposed method can be applied to the validation of all the algorithms needed in the practical navigation, such as profile integrating and large scale time transformation.

Keywords X-ray pulsar navigation (XPNAV) · Signal simulation · Phase estimation model · Inverse function method

L. Fu (✉) · P. Shuai
Qian Xuesen Laboratory of Space Technology, Beijing 100094, China
e-mail: fulingzhong@163.com

M. Xue · H. Sun · H. Fang
School of Aerospace Science and Technology, Xidian University, Xi'an 710126, China

56.1 Introduction

X-ray pulsar-based navigation (XPNAV) is a new autonomous navigation method with properties of strong anti-jamming, high safety, broad coverage area and high precision, which draws more and more attention of the main space powers around the world, such as America, Japan, and Germany [1]. In 2004, The Defense Advanced Research Projects Agency (DARPA) proposed an X-ray pulsar-based navigation (XNAV) and autonomous positioning verification project to research the simulation method of XPNAV and the development prototype. At present, the simulation of XPNAV and the feasibility testing are completed, and the space flight experiment is upcoming. Due to the atmospheric absorption, the X-ray pulsar signal can be hardly captured on the ground labs. Thus the experimental verification of the pulsar-based navigation must be carried out in space. However, developing X-ray detection system is highly complicated and difficult, and will consume a great deal of financial and material resource. So it is not suitable in the early stage of XPNAV research. To reduce the cost and to save time, research work is mainly based on simulation signal generated from simulation system on ground.

At present, researchers have developed two methods to generate pulsar signals. One is the semi-physical simulation method [3, 4] and the other is the numerical simulation method [5–8]. Semi-physical method generates real photon whose intensity is modulated by pulsar's standard profile. Photon detector receives the photon signal, and the actual arrival time of the photon reach is stamped, which is used to validate algorithm, such as epoch folding, estimation of time of arrival (TOA), and navigation. Numerical simulation method uses computer to generate pulsar signals. Compared with the semi-physical method, the numerical simulation method is more simple and efficient.

At present, no matter what simulation method is used, the X-ray pulsar signal is treated as periodic. Actually the X-ray pulsar's period varies slowly [9]. Simulated pulsar signal with fixed period is not precise. In long time autonomous navigation, a slight deviation of period will lead to a great distortion of accumulative profile. In order to verify validity of the algorithm, such as period search, epoch folding, TOA estimation and autonomous navigation, it is very important to research a new simulation method to simulate the slow varying property of the pulsar period.

In this paper, a novel method for simulating the slow varying property of the pulsar period is developed. A recursion algorithm of pulsar photon phase is improved from the traditional Photon TOA recursion algorithm. The phase prediction model is used to transform the photon phase to photon TOA. Finally, validity of the simulation method is verified in aspects of period, accumulative profile and photon flux of pulsar.

By analyzing the difference between simulation data and RXTE observation data, the results prove that simulated data is very similar to RXTE observation data. Thus the simulated data can be used to develop algorithm of period search, TOA estimation and pulsar-based autonomous navigation.

56.2 Principle of Simulation Method

Basic idea of the photonic analog is to generate the photon TOAs at SSB using the pulsar standard profile and phase prediction model of SSB. Then the photon sequence is transformed from the spacecraft proper time to a barycentric standard time reference, considering the various large scale space-time effects. Figure 56.1 shows the flowchart of the photon simulation. It mainly includes the following steps:

- (1) According to the pulsar phase prediction model, calculate the pulse phase of the simulation starting time at SSB. Existing pulsar phase prediction model is based on the radio spectrum. In order to obtain the pulse arrival time of X-ray spectrum at SSB, it is necessary to generate the observed profile of X-ray pulsar using the real data, and then calculate the phase delay between the X-ray pulse profile and radio pulse delay.
- (2) Choose the corresponding pulsar standard profile according to the simulation start time of the energy range.
- (3) Generate photon TOAs of the selected energy range and the simulation time at SSB.
- (4) According to the spacecraft orbit and the ephemeris of the solar system, transform the spacecraft proper time to the TDB time of SSB, by calculating the aberration and the various large scale space-time effects. At last, output the photon TOAs at the spacecraft.

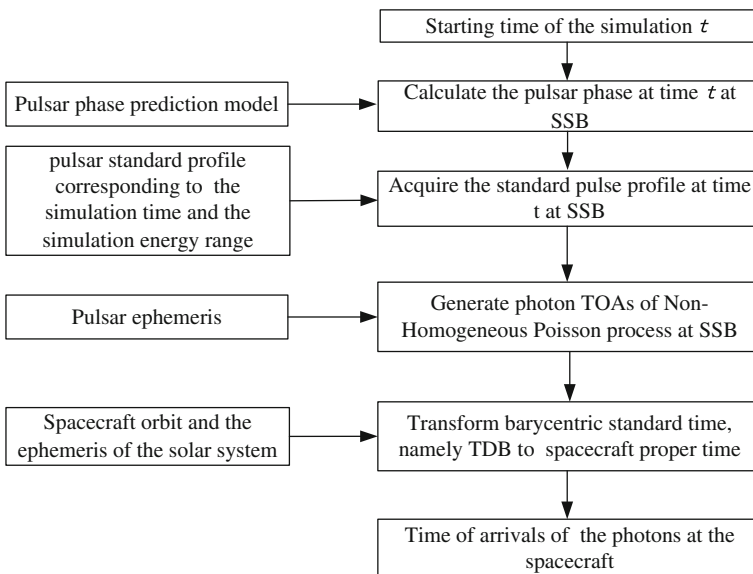


Fig. 56.1 Process of simulation

This method conforms to the propagation process of X-ray pulsar arrival photon at the spacecraft. The physical process is clear and the simulation algorithm takes advantage of the pulsar ephemeris, the spacecraft's orbit and the solar ephemeris, so the simulated photon sequence contains the pulsar cycle change, as well as the various large scale space-time effects.

56.3 Algorithm and Realization

56.3.1 Generate Photon TOAs at SSB

56.3.1.1 Recursive Simulation Method

The photon sequence emitting by the X-ray pulsar can be modeled as a non-homogeneous Poisson process (NHPP) [7]. Assuming that a photon arrives at $t_n = t$, the probability distribution of $Z = t_{n+1} - t_n > z$ can be expressed by Eqs. (56.1) and (56.2).

$$\begin{aligned} P(Z > z | t_n = t) &= P(N_{t+z} - N_t = 0) \\ &= \exp\left(-\int_t^{t+z} \lambda(t) dt\right) \\ &= \exp(-(\Lambda(t+z) - \Lambda(t))) \end{aligned} \quad (56.1)$$

where t_n is the TOA of n th photo, N_t is the number of the detected photons in the interval $(0, t)$, $\lambda(t)$ is a time-varying Poisson rate of the pulsar, and $\Lambda(t)$ is the integrated rate of the Poisson process.

$$P(Z > z | t_n = t) = 1 - F_{z|t_n=t}(z | t_n = t) \quad (56.2)$$

where $F_{z|t_n=t}(z | t_n = t)$ is the probability distribution of $Z = t_{n+1} - t_n < z$. Assuming that a photon arrives at $t_n = t$, from Eqs. (56.1) and (56.2), we can obtain

$$F_{z|t_n=t}(z | t_n = t) = 1 - e^{-(\Lambda(t+z) - \Lambda(t))} \quad (56.3)$$

Hence, the inverse function of $F_{z|t_n=t}(z | t_n = t)$ can be generated as

$$\overset{-1}{F}_{z|t_n=t}(z | t_n = t) = -t + \Lambda^{-1}(\Lambda(t) - \ln(1 - z)) \quad (56.4)$$

Given the n th photon generated at $t_n = t$, Z can be generated as

$$Z = -t + \Lambda^{-1}(\Lambda(t) - \ln(1 - U)) \quad (56.5)$$

Since 1-U is a uniform random variable in the interval (0, 1), without loss of generality, Eq. (56.5) can be expressed as

$$Z = -t + \Lambda^{-1}(\Lambda(t) - \ln U) \quad (56.6)$$

Let $E = -\ln(U)$, E is an exponential random variable. As a result, using Eq. (56.6) and giving $t_n = t$, t_{n+1} is generated as

$$\begin{aligned} t_{n+1} &= t_n + Z \\ &= \Lambda^{-1}(\Lambda(t_n) - \ln U) \end{aligned} \quad (56.7)$$

56.3.1.2 Simulation the Slow Varying Property of Pulsar Period

Using Eq. (56.7), we can generate the next photon TOA at SSB from the previous one. In Eq. (56.7), only the integrated rate of the Poisson process $\Lambda(t)$ and $\Lambda^{-1}(t)$, and the inverse function of $\Lambda(t)$ are related to the pulsar. The pulsar standard profile which is calculated in a large time scale is not a transient behavior of the pulsar and can't reflect the varying of pulsar period. But the pulsar standard profile folded by phase, expressed as $\lambda(\phi)$, is very stable. Pulsar phase is related with the pulsar photon TOA by the pulsar phase prediction mode. By this way, the pulsar photon TOA generated can reflect the varying of pulsar period.

In Eq. (56.7), substituting $\phi(t_n)$ into t_n , we can obtain the recursion equation of the pulsar photon phase

$$\phi(t_{n+1}) = \Lambda^{-1}(\Lambda(\phi(t_n)) - \ln U) \quad (56.8)$$

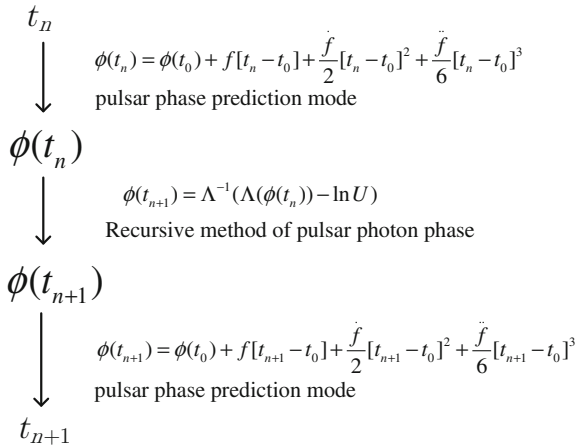
There is a consistent one-to-one match between the phase of pulsar photon and the TOA of pulsar photon. Using the pulsar phase prediction mode, we can solve t_n , whose corresponding pulsar phase is $\phi(t_n)$.

$$\phi(t) = \phi(t_0) + f[t - t_0] + \frac{\dot{f}}{2}[t - t_0]^2 + \frac{\ddot{f}}{6}[t - t_0]^3 \quad (56.9)$$

Recursion method of pulsar photon TOAs is illustrated in Fig. 56.2.

The recursive process pulsar photon arrival times contains the relationship between the photon TOAs and the pulsar period prediction model. Therefore the simulated photon sequence contains the pulsar frequency, the second and third order items of the pulsar frequency and so on. We can obtain the pulsar frequency and its derivatives from the Princeton pulsar ephemeris database. Princeton pulsar ephemeris database records the pulsar frequency and its derivatives corresponding to different time durations. The length of the time period is associated with the speed of change of the pulsar frequency. The faster the speed of change of the pulsar frequency is, the shorter the valid period of the corresponding frequencies.

Fig. 56.2 Recursive method of pulsar photon TOAs



During simulation according to the arrival time of the former photon, search the corresponding rotation frequency from the pulsar ephemeris database to obtain the phase prediction model and then generate the arrival time of next photon based on the recursive process of pulsar photons. Therefore, each simulated photon contains the change of the pulsar’s instantaneous frequency and the simulated photon sequence reflects the graded properties of the pulsar’s rotation period.

56.3.1.3 Numerical Solution for Photon Phase Recursion

Using Eq. (56.8) to generate the next pulsar photon TOA, integrated rate function of the pulsar Poisson process and its inverse function are essentially required. Because the Poisson rate of the pulsar is irregular, it is difficult to be described as an analytical function expression. Even though analytical function expression can be obtained by fitting method, it will be very complex. Thus numerical method is more efficient to solve this problem. When the Poisson rate of the pulsar is discrete, the integrated rate function of pulsar can be described by a phase-integrated rate value chart, Phase as independent variable and integral rate value as the dependent variable. Whatever a photon phase is, a corresponding integral rate value can be calculated by interpolation method.

The inverse function can be easy obtained just by exchanging the independent variable and the dependent variable of original function. Thus when an integral rate value is given, a corresponding photon phase can be calculated by the interpolation method. Figure 56.3 is the diagram of the numerical recursion solution for the photon phase.

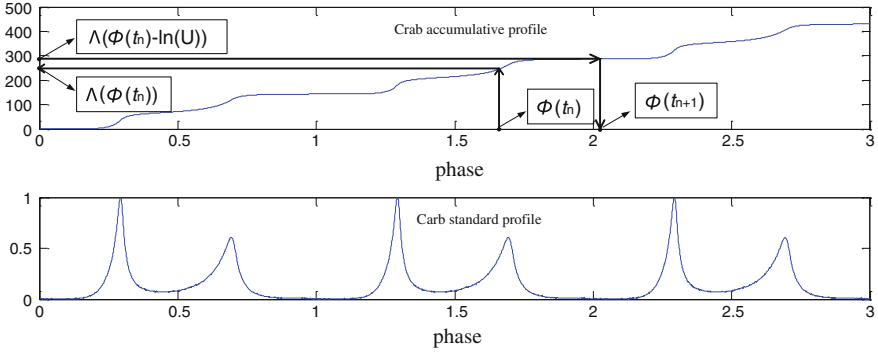


Fig. 56.3 Diagram of numerical solution for photon phase recursion

56.3.2 Time Transfer of Photon TOAs from SSB to Spacecraft

The photon sequence at the SSB solar photon uses the TDB time, while the spacecraft recorded the arrival time of photons using the natural proper time generated by atomic clock (proper time, PT), therefore, we need to correct the aberration and transform the time scales. Time scale transfer mainly includes the one of Einstein delay. The aberration correction includes the Roemer delay and the sun Shapiro delay.

56.3.2.1 Aberration Transfer

The complete aberration transfer needs integral operation, and is not practical. This paper adopts the practical simplified formula of the RXTE platform.

$$t_{SSB-TDB} - t_{SC-TDB} = \frac{\vec{n} \cdot \vec{r}_{SC/SSB}}{c} - \frac{2\mu_s}{c^3} \log\left(1 + \frac{\vec{r}_{SC/sun} \cdot \vec{n}}{\sqrt{\vec{r}_{SC/sun} \cdot \vec{r}_{SC/sun}}}\right) \quad (56.10)$$

This simplified time transformation formula calculates the time needed for the photon propagating from SSB to the spacecraft. Let $t_{SSB-TDB}$ denote the TDB time of the photon TOA at the SSB and t_{SC-TDB} denote the TDB time of the photon TOA at the spacecraft. The position of the spacecraft relative to the SSB is $\vec{r}_{SC/SSB}$, $\vec{r}_{SC/SSB} = \vec{r}_{SC/E} + \vec{r}_{E/SSB}$, the unit direction to the pulsar is \vec{n} with respect to the SSB and the position of the spacecraft with respect to the sun is $\vec{r}_{SC/sun}$.

Calculating $\vec{r}_{SC/SSB}$ and $\vec{r}_{SC/sun}$ needs to know the TDB time of the photon TOA at the spacecraft. Then by searching the solar ephemeris files, the relevant parameters can be acquired. Furthermore, the TDB time of the photon TOA at the

spacecraft is also required to get \vec{n} . Therefore, it is difficult to obtain a mathematical formulation of t_{SC-TDB} . This article uses the iteration method to solve t_{SC-TDB} . The iteration steps are as follows:

- (a) Let $k = 0$, $t_0 = 0$, and $\Delta t = 0$;
- (b) $t_{SC-TDB} = t_{SSB-TDB} - \Delta t$;
- (c) Calculate $\vec{r}_{SC/SSB}$, $\vec{r}_{SC/sun}$ and \vec{n} ; substitute them into Eq. (10) and then calculate t_{k+1} ;
- (d) Define $\Delta t = t_{k+1} - t_k$; if $|\Delta t| < \varepsilon$, then stop the iteration procedure; Else let $k = k + 1$ and go to step (b).

To ensure that the error of the time transfer does not affect the accuracy of the subsequent photon processing algorithm, the iteration stopping criterion is $\varepsilon = 10^{-12}$ (s).

56.3.2.2 Time Scale Transfer

The complete timescale transfer needs integral operation, and is not practical. This paper adopts the practical simplified formula of the RXTE platform.

$$t_{SC-TDB} = t_{SC-TT} + \frac{1}{c^2} \vec{r}_{SC/E} \cdot \vec{v}_{E/SSB} + P \quad (56.11)$$

Equation (56.11) performs the time transformation from TT to TDB. t_{SC-TT} is the TT time of the photon TOA at the spacecraft, t_{SC-TDB} is the TDB time of the photon TOA at the spacecraft, and $\vec{r}_{SC/E}$ is the position of the spacecraft with respect to Earth. $\vec{v}_{E/SSB}$ is the inertial velocity of Earth with respect to the SSB and P includes all the period correct items. The TT time of the photon TOA at the spacecraft, which can be obtained from the spacecraft's orbit file, is needed in order to calculate $\vec{r}_{SC/E}$. However, the TT time of the photon TOA at the spacecraft the exactly what we need to resolve. Therefore, we adopt the iteration method which is as that of the aberration transfer.

56.4 Numerical Simulations

Take pulsar B0531+21 for example, analyze and process the simulated photon sequence, and verify the correctness of the pulsar signal simulation. Validation contains the following three aspects: (1) validation of the graded properties of pulsar's rotation period. Search the period of the simulated photon sequence and compare the searched period value with the one of the pulsar ephemeris to verify the correctness of the simulated period characteristic; (2) Validation of the pulsar profile and flow. According to the phase prediction model of the pulsar, obtain the

pulsar profile through the epoch folding procedure and then compare it with the standard profile to verify the simulation accuracy of the profile and flow; (3) Comparison experiments using the measured data. Validate the similarity of the simulated photon sequence and the real data via comparing the simulated data with the real data.

56.4.1 Validation of Variation Property of Period

Take Pulsar B0531+21 for example. The photon data of five time durations is generated. The length of each time period is 180 s. The average radiation flow of pulsar is $1.54 \text{ ph/cm}^2 \text{ s}$ and the detector area is 1000 cm^2 .

In order to get the period of the simulated photon sequence, we need to use the cycle search algorithm, which is as follows:

- (1) Transform the arrival time of the photon sequence to the geocentric coordinate system of the solar system;
- (2) According to the Eq. (56.12) expressed as

$$\Phi_i = \frac{(t_i - t_0)}{T} \bmod 1.0 \quad (56.12)$$

where t_i is the arrival time of the i th photon and t_0 is the reference time, and using a certain rotation frequency f ($f = 1/T$), the absolute phase of each photon is calculated and only the fractional part of it is retained. The obtained phase is within $[0, 1]$;

- (3) The phase $[0, 1]$ is divided into m equal-length bins and then the photon counts in each bin are calculated. Afterwards, calculate the Pearson chi-square statistics by

$$\chi^2 = \sum_{i=1}^m \frac{(n_i - N/m)^2}{N/m}, \quad (56.13)$$

in which N is the total photon number, and $N = \sum n_i$.

Search the period the simulated photon sequences of the selected five time durations according to the above steps. The searched periods are shown in Table 56.1. It illustrates that the period of the simulated photon is very close to the cycle in the pulsar ephemeris. The error is due to the fact that the cycle of the pulsar is slowly changing, while the searched period is the average period the corresponding time duration, but the cycle of the ephemeris is an instantaneous cycle of the starting time of simulation. If the simulation duration is longer, and the second-order and the third-order of the simulated photon sequence are also searched, it is promising to achieve simulated photons of higher accuracy.

Table 56.1 Period of simulation data

Starting time of simulation (MJD)	Period in the pulsar ephemeris (f0)	Period of the simulated photon sequence
49128.000000191	29.9245763662884	29.92460
50200.000000251	29.8897220169126	29.88980
50645.000000104	29.8752812172442	29.87540
50849.000000344	29.8686661308798	29.86890
51224.000000271	29.8565144334781	29.85660

56.4.2 Validation of Profile and Flux

Take pulsar B0531+21 for example. The starting time of the simulation is 52557.1141014815 (MJD), the simulation duration of 180 s, and the average flow rate of the pulsar radiation is 1.54 ph/cm² s. The detector area is 1000 cm². Respectively validate the simulation accuracy of the profiles of the selected six energy ranges, namely 2–5, 5–9, 9–13, 13–17, 17–22 and 22–60 keV. The simulation results are shown in Fig. 56.4. It shows that the cumulative pulse profiles of the simulated photon sequence have very high similarity with the standard pulse profile, and almost coincide perfectly with the standard profile. Table 56.2 lists the correlations between the cumulative pulse profile of the simulated photons and the standard one. It can be seen that all the correlations reach more than 0.99. Table 56.3 compares the fluxes of the simulated photons with the fluxes of real data. It depicts that the flux of the simulated photons is close to the actual photon flux. If the simulation time is longer, simulation photon flux will be more close to the practical one, and the simulated flow error will be smaller.

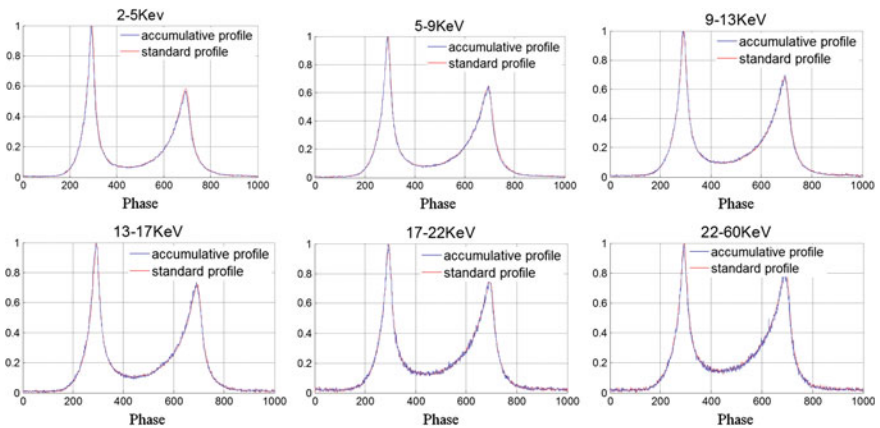


Fig. 56.4 Standard profile and accumulative profile of simulation data for B0531+21

Table 56.2 The cross-correlation coefficient between Standard profile and cumulative profile of simulation signal of B0531+21

Energy range (Kev)	2-5	5-9	9-13	13-17	17-22	22-60
Correlation coefficients of profiles	0.9989	0.9989	0.9984	0.9980	0.9978	0.9977

Table 56.3 Observation flux and simulation flux of B0531+21

Energy range (KeV)	2-5	5-9	9-13	13-17	17-22	22-60
Actual flux (ph/s)	800	540	400	200	80	60
Simulated flux (ph/s)	824	554	410	205	82	59
Flux error (%)	3.0	2.8	2.5	2.5	2.5	1.6

56.4.3 Comparison Between Simulation Signal and RXTE Observation Signal

In previous sections, slow varying property of pulsar period, a small detail of pulsar signals, has been validated to be correct. In order to validate the similarity between the accumulative profile of the RXTE observation data and that of the simulation data, observation data segment of RXTE and simulation data segment of the same pulsar, start time, stop time and orbit parameters are selected. The two selected data segments are processed into accumulative profile by the same method. The cross-correlation coefficient between accumulative profile of RXTE observation data and accumulative profile of simulation data is calculated. Also signal to noise ratios (SNRs) of the two accumulative profiles are counted separately. Verification diagram of the simulation data’s validity is illustrated in Fig. 56.5.

The 60079-01-25-00 data package of pulsar B0531+21 is selected from RXTE observation data. Then six data segments with observation time of one minute and with starting time separately of 52557.1155893744 (MJD), 52557.1162838679 (MJD), 52557.1169783609 (MJD), 52557.1176728534 (MJD), 52557.1183673454 (MJD) and 52557.119061837 (MJD), are selected from 60079-01-25-00 data package. With the help of the Heasoft software, proper time of the photon TOAs at RXTE is corrected to the TDB time at the SSB. The accumulative profile, depicted as the red curve in Fig. 56.6, is obtained using the epoch folding method. As stated above, six simulated data segments are separately folded to get their accumulative profiles, which are depicted as the blue curves in Fig. 56.6. Figure 56.6 shows that the accumulative profile of the simulated data is very similar to the accumulative

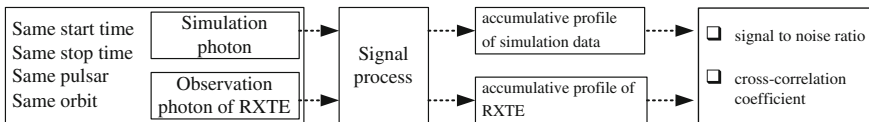


Fig. 56.5 Verification diagram of simulation data’s validity

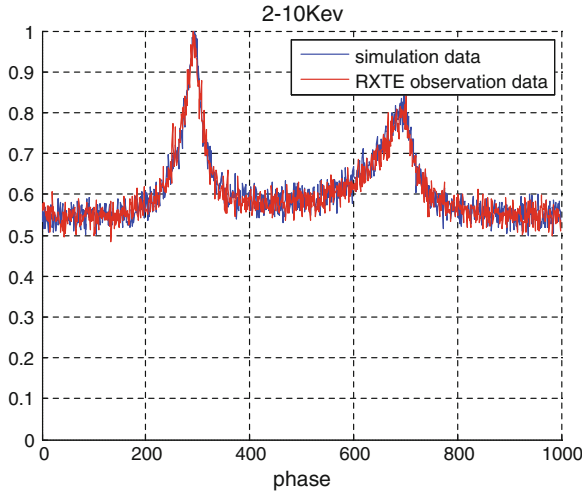


Fig. 56.6 Accumulative profile of RXTE observation data and simulation data

Table 56.4 The cross-correlation coefficient between accumulative profile of RXTE observation data and accumulative profile of simulation data

Data segments	1	2	3	4	5	6
correlation coefficient	0.90378	0.9197	0.91205	0.9402	0.9373	0.9270

Table 56.5 SNR between accumulative profile of RXTE observation data and accumulative profile of simulation data

Data segments	1	2	3	4	5	6
SNR of RXTE	34.6	37.7	34.9	39.4	37.6	36.9
SNR of simulation	36.4	35.4	37.1	37.6	37.2	36.1
Difference	-1.8	2.3	-2.2	1.8	0.4	0.8

profile of the RXTE observation data. The cross-correlation coefficients between accumulative profile of RXTE observation data and accumulative profile of the simulated data, which are all above 0.9, are presented in Table 56.4. SNR differences between the two accumulative profiles, which are presented in Table 56.5, are all within 3 dB.

56.5 Conclusion

The period of pulsar has slow changing characteristics. Photon signals generated by the fixed phase model cannot reflect the period's variation properties, have no authenticity. This article combines the pulsar phase prediction model with photonic

phase recursive formula to develop a new recursive method of photon arrival time. Finally, by simulating the pulsar's rotation cycle, pulse profile, flux and comparing these aspects with the real data detected by the RXTE, we validate the simulation method in terms of several indexes. It shows that the simulated photons not only conforms to Poisson process, but also contain the cycle variation characteristics. The correlation between the simulated pulse profile and the standard profile of real data is high up to 0.99. The flow difference between the simulated photons and the real value of RXTE satellite is less than 3 %. The SNR difference of the simulated pulse profile and test data of the RXTE satellite is less than 3 db. The simulated photon series of the proposed method are very close to the measured data of RXTE satellite, and have high authenticity. The proposed technique can be applied to validation studies of variance algorithms used in the X-ray pulsar based navigation, such as the period search, profile integration, TOA estimation and navigation filters.

References

1. Shuai P, Li M, Chen S, Huang Z (2009) Principles and techniques of X-ray pulsar-based navigation system. China Astronautic, Beijing
2. Liu J (2011) X-ray pulsar-based spacecraft autonomous navigation. Huazhong University of Science and Technology
3. Hu H, Zhao B, Sheng L, Yan Q (2011) A simulation experiment system for X-ray pulsar based navigation. *J Phys* (02):853–861
4. Sun HXK, Li X et al (2013) A simulation technique of X-ray pulsars signals with high timing stability. *J Phys* (10):518–528
5. Su Z, Xu L, Wang T (2011) A simulation experiment system for X-ray pulsar based navigation. *J Phys* (11):827–834
6. Sala J, Urruela A, Villares X, Estalella R (2004) Feasibility study for a spacecraft navigation system relying on pulsar timing information, european space agency advanced concepts team ARIADNA study
7. Emadzadeh A, Speyer JL (2010) On modeling and pulse phase estimation of X-ray pulsars. *IEEE Trans Signal Process* 58(9):4484–4495
8. Zhang H, Xu L, Song S (2014) Jiao Rong. *J Phys* 98:189–200
9. Ge M (2012) The X-ray emission of pulsars. Institute of High Energy Physics Chinese Academy of Sciences

Chapter 57

Multi-frame Visual Odometry in Image-Aided Inertial Navigation System

N.S. Gopaul, J.G. Wang and B. Hu

Abstract This paper presents a novel stereo image-based image aided inertial navigation algorithm for reducing position and orientation drifts during GNSS outages or in a poor GNSS environment. Usually, the image aided navigation based on the visual odometry uses the tracked features only from a pair of the consecutive image frames. The proposed method integrates the features tracked from all overlapping image frames towards accuracy improvement and drift reduction. The measurement equation system in this multi-frame visual odometry algorithm (MFVO) is derived from Simultaneous Localization and Mapping (SLAM) measurement equation system where the landmark position parameters in SLAM are algebraically eliminated by time-differencing the measurement at two consecutive epochs. However the resulted time-differenced measurements are time-correlated. Through a sequential de-correlation the Kalman filter measurement update can be performed sequentially and optimally. Monte Carlo simulations show that the MFVO and SLAM pose estimates are similar. Compared with SLAM, the proposed method uses less computation resources especially when the number of features in view is large. The results from a real dataset are also presented.

Keywords Mobile mapping · Inertial navigation · Kalman filter · Multi-frame · Visual odometry

57.1 Introduction

The high demand for direct-georeferencing techniques through low-cost multisensor integrated kinematic positioning and navigation systems in mobile mapping is continuously driving more research and development activities. The effective and

N.S. Gopaul · J.G. Wang (✉) · B. Hu
Department of Earth and Space Science and Engineering, York University,
4700 Keele St, M3J 1P3 Toronto, ON, Canada
e-mail: jgwang@yorku.ca

sufficient utilization of images is among the most recent scientific and high-tech research tasks. In this particular field of mapping and imaging, we are engaged in studying of the image-aided inertial integrated navigation as the natural continuation of its past research in the multisensor integrated kinematic positioning and navigation at the at the Earth Observation Laboratory of York University.

An image-aided inertial integrated navigation based on visual odometry (VO) uses the tracked features only from consecutive frames before performing the Kalman filter measurement update [4, 10, 16]. The main components of a typical stereo VO system consist of a point-feature detector, a feature matcher and pose change estimation algorithm [15]. Similarly to a stand-alone INS system, the position and orientation solution from VO drifts from the true path with time. It is important to minimize the drifts, especially in applications such as direct georeferencing and mobile mapping. There are two popular approaches to solving this problem, namely Simultaneous Localization and Mapping (SLAM) and Bundle Adjustment (BA).

The SLAM algorithm [3] with image measurements [1, 11] outputs more accurate camera position and orientation compared with the ones from VO. The estimation of INS errors using visual SLAM methods is commonly performed using the Extended Kalman filter (EKF) [14]. As the main drawback with SLAM, it requires more computation resources than VO because of the involvement of the simultaneous estimation of the camera pose and the locations of the observed stationary landmarks. Efforts to reduce the computation complexity can be achieved through approximate and suboptimal methods [9, 12]. They not only trade optimality for computation and memory usage, and also degrade or even cause the KF estimates to diverge.

Bundle Adjustment (BA) simultaneously estimates all camera pose, landmark and sometimes camera calibration parameters [17]. The adjustment problem is typically solved using non-linear least squares (LS) [17] or the Levenberg-Marquardt (LM) algorithm [6]. The computed Jacobian matrix and normal equation system can be large so that solving this linearized system is inefficient and impractical especially when there are many camera positions, orientations and landmarks [8]. In some applications, such as mobile mapping, only the camera pose parameters are required. One method was developed to algebraically eliminate the landmark position vector by using three-view constraints [7]. Given that the number of landmark states is usually much larger than the number of camera poses, an elimination of the landmarks renders a more efficient estimation. However, BA methods estimate all parameters simultaneously in the system whose dimension can be large so it becomes impractical.

The proposed method, called multi-frame visual odometry (MFVO), can reach the same position and orientation accuracy as visual SLAM does, but without having to estimate the landmark positions. This is achieved by algebraically eliminating the time-invariant landmark position parameters in the SLAM measurement equation system at two consecutive epochs. There are three advantages with the proposed algorithm. Firstly, the position and orientation solution of the MFVO is equivalent to the ones estimated by the visual SLAM. Secondly, the

measurement update of the MFVO can be performed sequentially and optimally. And finally, the proposed method uses less computation resources compared to visual SLAM, especially when the number of features in view is large.

The paper continues with the derivation and discussion of the multi-frame visual odometry. Then, test results from Monte Carlo simulations and real data comparing with visual SLAM are presented. And finally the paper ends with the conclusions.

57.2 Multi-frame Visual Odometry

This section contains the derivation of the proposed stereo-based multi-frame visual odometry algorithm. Specifically, the goal is to obtain a measurement model that contains only pose estimates such that it can be run sequentially and optimally using the Kalman filter. The term ‘visual odometry’ is used here because our focus is on the image measurements. However the algorithm can be applied to any SLAM application. The derivation here follows the feature extraction, feature matching and stereo forward spatial intersection steps applied in VO.

The EKF-SLAM algorithm uses the Extended Kalman filter (EKF) to solve the SLAM problem. The system and measurement equation at epoch k is given as [3]:

$$\begin{pmatrix} x_k \\ M_k \end{pmatrix} = \begin{pmatrix} f(x_{k-1}) + w_{k-1} \\ M_{k-1} \end{pmatrix}, w_{k-1} \sim N(0, Q_{k-1}) \quad (57.1)$$

$$z_{k,i} = h_{k,i}(x_k, m_i) + v_{k,i}, v_{k,i} \sim N(0, R_{k,i}) \quad (57.2)$$

where x_k is the vehicle’s pose state vector, $M_k = [m_1^T, \dots, m_i^T, \dots, m_N^T]^T$ is the landmark states, w_{k-1} is the process noise vector with zero mean and variance Q_{k-1} , $z_{k,i}$ is the measurement vector expressed by $h_{k,i}(x_k, m_i)$ as the non-linear function of the vehicle pose states and the landmark states, and $v_{k,i}$ is measurement noise vector with zero mean and variance $R_{k,i}$.

Consider the SLAM measurement equations (57.2) at two consecutive epochs $k-1$ and k :

$$\begin{aligned} z_{k-1,i} &= h_{k-1,i}(x_{k-1}, m_i) + v_{k-1,i}, v_{k-1,i} \sim N(0, R_{k-1,i}) \\ z_{k,i} &= h_{k,i}(x_k, m_i) + v_{k,i}, v_{k,i} \sim N(0, R_{k,i}) \end{aligned} \quad (57.3)$$

The landmark states m_i in Eq. 57.3 are time invariant and can algebraically be cancelled to obtain the following implicit measurement equation

$$\tilde{h}_k(x_k, x_{k-1}, z_k, z_{k-1}, v_k, v_{k-1}) = 0 \quad (57.4)$$

For simplicity the feature index i is dropped. Equation 57.4 is then linearized using the first order Taylor series expansion to

$$\delta\tilde{z}_k = H_{k,k}\delta x_k + H_{k,k-1}\delta x_{k-1} + \tilde{v}_k, \tilde{v}_k \sim N(0, \tilde{R}_{k,k}) \quad (57.5)$$

where $\delta\tilde{z}_k = -\tilde{h}_k(x_k, x_{k-1}, z_k, z_{k-1}, 0, 0)$, $\tilde{v}_k = M_{k,k}v_k + M_{k,k-1}v_{k-1}$, and $H_{k,k}$, $H_{k,k-1}$, $M_{k,k}$ and $M_{k,k-1}$ are the Jacobian matrices of the vectors x_k , x_{k-1} , v_k and v_{k-1} respectively. The variance matrix of \tilde{v}_k is $\tilde{R}_{k,k} = M_{k,k}R_kM_{k,k}^T + M_{k,k-1}R_{k-1}M_{k,k-1}^T$. Now consider the derived measurement equation (Eq. 57.3) at consecutive epochs $k-1$ and k :

$$\begin{aligned} \tilde{h}_{k-1}(x_{k-1}, x_{k-2}, z_{k-1}, z_{k-2}, v_{k-1}, v_{k-2}) &= 0 \\ \tilde{h}_k(x_k, x_{k-1}, z_k, z_{k-1}, v_k, v_{k-1}) &= 0 \end{aligned} \quad (57.6)$$

Notice that the measurement noise vector v_{k-1} appears in $\tilde{h}_{k-1}(\cdot)$ and $\tilde{h}_k(\cdot)$. So, the derived measurement noise vectors \tilde{v}_k and \tilde{v}_{k-1} are correlated, i.e.

$$E[\tilde{v}_k\tilde{v}_{k-1}^T] = M_{k,k-1}E[v_{k-1}v_{k-1}^T]M_{k-1,k-1}^T = M_{k,k-1}R_{k-1}M_{k-1,k-1}^T = \tilde{R}_{k,k-1} \quad (57.7)$$

The standard Kalman filter runs under the assumption that measurement noise vector is white, i.e. normally distributed with zero means, and independent. However the measurement noises here are time-correlated. The solution of the standard Kalman filter becomes suboptimal if one neglects significant time-correlation. Furthermore it can degrade the filter's performance. One method to solve this problem is to de-correlate the correlated measurements as described in [2]. Suppose that the landmark was observed at epoch $k-j$ for the very first time. Then, all the linearized measurement equations up to current epoch k can be given by

$$\begin{aligned} \delta\tilde{z}_{k-j} &= H_{k-j,k-j}\delta x_{k-j} + H_{k-j,k-j-1}\delta x_{k-j-1} + \tilde{v}_{k-j} \\ &\vdots \\ \delta\tilde{z}_{k-1} &= H_{k-1,k-1}\delta x_{k-1} + H_{k-1,k-2}\delta x_{k-2} + \tilde{v}_{k-1} \\ \delta\tilde{z}_k &= H_{k,k}\delta x_k + H_{k,k-1}\delta x_{k-1} + \tilde{v}_k \end{aligned} \quad (57.8)$$

For simplicity all the vectors and matrices corresponding to a feature tracked between frames $k-j$ and k are concatenated to obtain the following short-form equation

$$\delta\tilde{z} = H\delta x + \tilde{v}, \tilde{v} \sim N(0, \tilde{R}) \quad (57.9)$$

Using the de-correlation technique described in [2], Eq. 57.9 is multiplied by matrix C^{-1}

$$\begin{aligned} C^{-1}\delta\tilde{z} &= C^{-1}H\delta x + C^{-1}\tilde{v} \\ \delta z' &= H'\delta x + v' \end{aligned} \quad (57.10)$$

wherein C is the lower triangular square root of \tilde{R} obtained using the Cholesky factorization [5]. Measurement noise vector v' is now un-correlated and has the identity matrix as its variance matrix [2]. The de-correlated measurement equation system (57.8) at epoch k is

$$\delta z'_k = H'_{k,k} \delta x_k + H'_{k,k-1} \delta x_{k-1} + \cdots + H'_{k,k-j} \delta x_{k-j} + v'_k, v'_k \sim N(0, I) \quad (57.11)$$

The measurement vector $\delta z'_k$ can functionally be described by the matched features extracted from image frames at $k - j$ up to k . The final system and measurement model at epoch k for the feature index i become

$$\begin{pmatrix} \delta x_k \\ \delta x_{k-1} \\ \vdots \\ \delta x_{k-j} \end{pmatrix} = \begin{pmatrix} \Phi_{k-1} \delta x_{k-1} \\ \delta x_{k-2} \\ \vdots \\ \delta x_{k-j-1} \end{pmatrix} + \begin{pmatrix} w_{k-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (57.12)$$

$$\delta z'_{k,i} = H'_{k,k} \delta x_k + H'_{k,k-1} \delta x_{k-1} + \cdots + H'_{k,k-j} \delta x_{k-j} + v'_{k,i}, v'_{k,i} \sim N(0, I) \quad (57.13)$$

57.2.1 Discussion

\tilde{R} is a sparse matrix, i.e. its diagonal and off-diagonal block matrices are non-zero. Therefore, the block matrices in C epochwise can be obtained sequentially as described in [4]. Matrix C is a lower triangular matrix. Using the following blocked lower triangular inverse

$$\begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1} A_{21} A_{11}^{-1} & A_{22}^{-1} \end{pmatrix} \quad (57.14)$$

the computation of the block matrices in C^{-1} can also be done sequentially. Since both C and C^{-1} can be obtained sequentially, $\delta z'_k$ and $H'_{k,k}, \dots, H'_{k,k-j}$ can also be obtained sequentially which is practical and desirable in a Kalman filter. The value of j depends on the overlapping percentages of the image frames, which directly depends on the traveling velocity of a camera. When the vehicle is in stationary the overlapping area is 100 % and j increases without bound. For practical purposes, the authors recommend to limit j to a predefined threshold j_{\max} .

The computational complexity of the SLAM measurement update for the 3D case is $O((6 + 3N)^2)$ where N is the number of landmarks. The computational complexity of computing $\delta z'_k$ and $H'_{k,k}, \dots, H'_{k,k-j}$ is $O((3N) \times (6 + 6j))$ and MFVO measurement update is $O((6 + 6j)^2)$. This shows that MFVO is more efficient with the large number of landmarks.

57.3 Test Results

57.3.1 Simulation Test Results

A series of simulations were conducted in order to compare the performance of visual SLAM and the proposed multi-Frame visual odometry (MFVO) algorithm. A 2D scenario was used to demonstrate the performance of the proposed algorithm with the imaging and IMU measurements. The vehicle navigation states are the 2D position vector, the 2D velocity vector and the heading.

Figure 57.1 overviews the vehicle's trajectory and the landmarks and also shows the vehicle's velocity profile, heading profile and the number of the features in view.

In this example two accelerometers and one gyroscope were simulated at the data rate of 100 Hz. For simplicity and demonstration purposes, the IMU biases were omitted. The standard deviations of the used process noises of the accelerometers and gyroscope were $1.0 \text{ m/s}/\sqrt{\text{h}}$ and $4.5 \text{ deg}/\sqrt{\text{h}}$, respectively. The feature measurements are at 10 Hz and the corresponding standard deviation was set to 0.10 m.

Monte Carlo (MC) simulations were used to compare the performance of the three methods. Each algorithm was run 100 times. The true position and heading errors were computed for each run. Then the root-mean-square (RMS) errors across the 100 runs were computed at every epoch. Figure 57.2 shows the true pose RMS errors and the size of the state vector from SLAM and MFVO.

The simulations show that the MFVO and SLAM solutions are similar with the former having much less number of states.

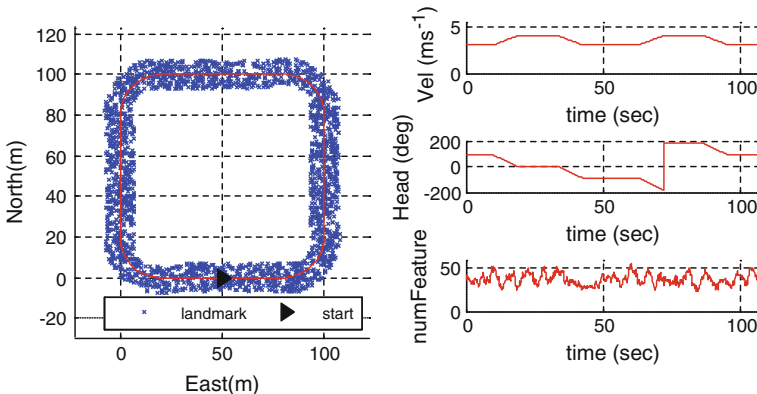


Fig. 57.1 Overview of the trajectory and the landmarks (*left*), and velocity and heading profiles and the number of the features (*right*)

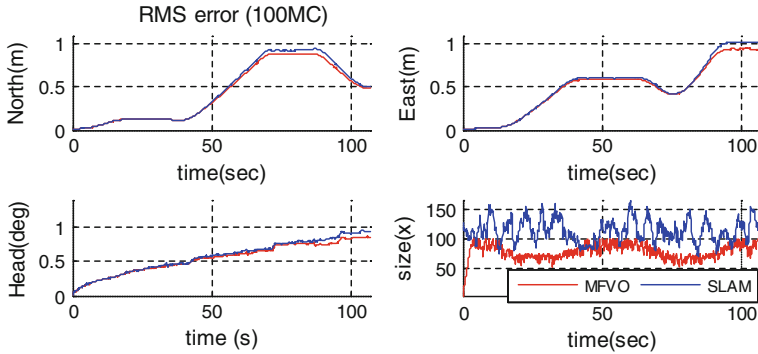


Fig. 57.2 Position RMS error, heading RMS error and size of the state vector

57.3.2 Road Test Results

This section compares the performance of visual SLAM aided-INS and MFVO aided-INS during simulated GPS outages using real data. The data was collected by the land vehicle navigation system developed at the Earth Observation Laboratory of York University [13] with two newly integrated cameras. The system consists of two NovAtel OEM GPS receivers, one Crossbow IMU440CA and two PointGrey Flea3 cameras separated by 65 cm. The resolution of the images was set to 640x480 pixels with the field of view of 50 degrees. Two GPS receivers provide the absolute heading measurements and a third GPS receiver was used as the base station so that RTK level GPS positioning accuracy have been achieved. The lever arm vectors of the GPS receivers and cameras with respect to the IMU unit were measured beforehand at the accuracy of 0.5 cm. The observation rates were set to 1.0, 100 and 7.5 Hz for GPs, IMU and cameras, respectively. Table 57.1 lists the IMU specification.

The data was collect in Vaughan, Ontario and the dataset length is 518 s. Figure 57.3 shows the top view of the vehicle’s trajectory, the velocity profile and the attitude profiles.

The full outages were simulated by tuning off the GPS position, velocity and heading measurement during the specified intervals. Three (3) GPS outages were simulated of duration 30 s located over the time intervals [130, 160 s], [240, 270 s] and [430, 460 s]. After the data were processed with the simulated GPS gaps, the solution was differenced with the one containing no gaps. The 3D position

Table 57.1 Crossbow IMU440CA specifications

Angular rate	Bias stability (deg/hour)	<10.0
	Angle random walk (deg/√hr)	< 4.5
Acceleration	Bias stability (mg)	<1.0
	Velocity random walk (m/s/√hr)	<1.0

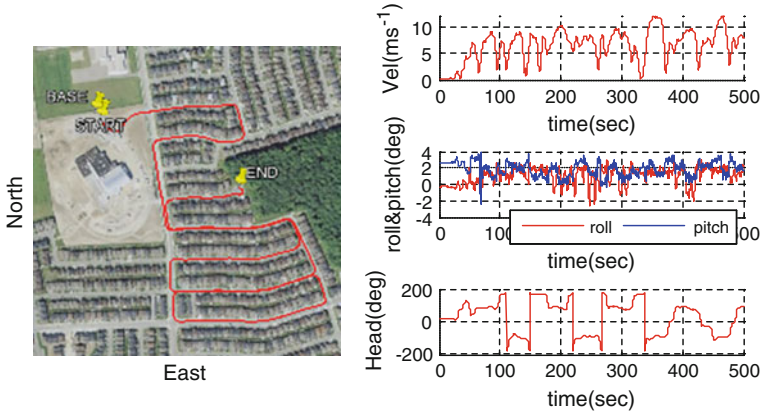


Fig. 57.3 The 2D trajectory (*left*). Velocity and attitude profiles (*right*)

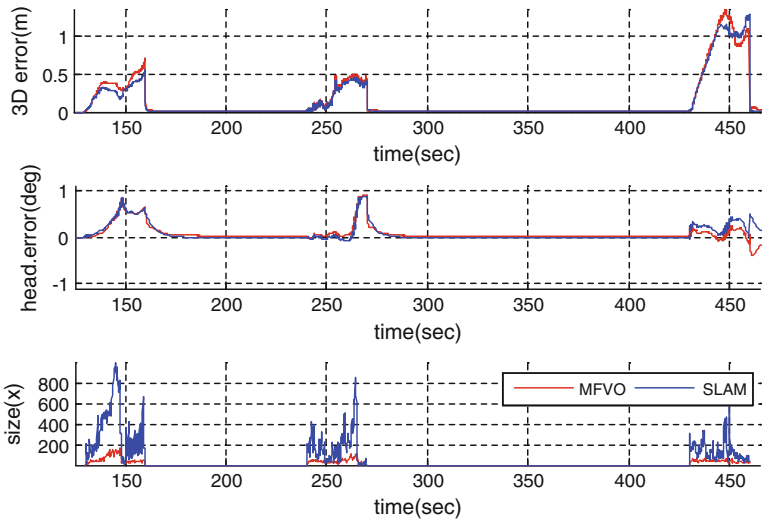


Fig. 57.4 3D position differences, heading differences and the dimensions of the state vector

difference was then computed. Figure 57.4 shows the 3D position differences, heading differences, and the dimensions of the state vector. The SLAM and MFVO solutions are comparable, but MFVO having much less number of states. The average number of states during the GPS outages in the SLAM and MFVO algorithms are 228 and 58, respectively.

57.4 Conclusions

This paper presented a novel stereo image-based multi-frame visual odometry (MFVO) for image-aided inertial navigation. The proposed method integrates the features tracked from multiple stereo image frames. MFVO measurement equation system was derived by algebraically eliminating landmark position parameters in the SLAM measurement equation system by time-differencing. However the resulted time-differenced measurements become correlated. Hence, a sequential measurement de-correlation technique has been applied so that the MFVO could be realized and implemented in a Kalman filter. Compared with the SLAM, the proposed MFVO is computationally more efficient when the number of features in view is large and their solution accuracies are quite similar. Monte Carlo simulations showed that the position and orientation solution from SLAM and MFVO are similar. The results from a road test showed that the SLAM and MFVO navigation solutions are comparable. But on average the dimension of the SLAM state vector could be as high as 4 times of the one in MFVO. The next step in the research will apply un-calibrated image measurements in the integrated system.

Acknowledgments The authors would like to acknowledge the financial support through research grants provided by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

1. Alcantarilla PF, Jos'e JY, Javier A, Bergasa LM (2012) On combining visual SLAM and dense scene flow to increase the robustness of localization and mapping in dynamic environments. In: IEEE International conference on robotics and automation RiverCentre, Saint Paul, Minnesota, USA, 14–18 May 2012
2. Bierman GJ (1977) Factorization methods for discrete sequential estimation. ISBN 0-486-44981-5, p 47
3. Durrant-Whyte H, Bailey T (2006) Simultaneous localisation and mapping (SLAM): part II State of the art. *Robot Automat Mag* 13(3):108–117
4. Gopaul NS, Wang J, Hu B (2014) Discrete EKF with pairwise time-correlated measurement noise for image-aided inertial integrated navigation. In: Joint ISPRS/IGU 2014 conference, 6–8 Oct 2014, Toronto, Canada
5. Grewal MS, Andrews AP (2001) Kalman filtering: theory and practice using MATLAB, 2nd edn, p 221
6. Hartley R, Zisserman A (2000): Multiple view geometry in computer vision. Cambridge
7. Indelman V, Melim A, Dellaert F (2013) Incremental light bundle adjustment for robotics navigation. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), Tokyo, Japan, 3–7 Nov 2013
8. Jeong Y, Nister D, Steedly D, Szeliski R, Kweon I (2012) Pushing the envelope of modern methods for bundle adjustment. *IEEE Trans Pattern Anal Mach Intell* 34(8):1605–1617
9. Julier SJ (2001) A sparse weight Kalman filter approach to simultaneous localisation and map building. In: Proceedings of the 2001 IEEE/IRJ, international conference on intelligent robots and systems Maui, Hawaii, USA, 29 Oct–03 Nov 2001

10. Konolige K, Agrawal M, Sol'a J (2007): Large scale visual odometry for rough terrain. In: International symposium on research in robotics, pp 201–212, Nov 2007
11. Lategahn H, Geiger A, Kitt B (2011) Visual SLAM for autonomous ground vehicles. In: IEEE International conference on robotics and automation shanghai international conference center, Shanghai, China, 9–13 May 2011
12. Guivant JE, Nebot, EM (2001) Optimization of the simultaneous localization and map-building algorithm for real-time implementation. *IEEE Trans Robot Autom* 17(3)
13. Qian K, Gopaul N, Wang J-G, Baoxin Hu (2012) Low cost multisensor kinematic positioning and navigation system with Linux/RTAI. *J Sens Actuator Netw* 1(3):166–182
14. Sazdovski V, Peter M, Silson G (2011) Inertial navigation aided by vision-based simultaneous localization and mapping. *IEEE Sens J* 11(8):1646–1656
15. Scaramuzza D, Fraundorfer F (2011) Visual odometry part i: the first 30 years and fundamentals. *IEEE Robot Autom Mag* 18(4):80–92
16. Tardif JP, George M, Laverne M, Kelly A, Stentz A (2010) A new approach to vision-aided inertial navigation. In: International conference on intelligent robots and systems
17. Triggs N, McLauchlan P, Hartley R, Fitzgibbon A (2000) Bundle adjustment—a modern synthesis. In: Triggs B, Zisserman A, Szeliski R (eds) *Proceedings on international workshop vision algorithms: theory and practice*, pp 298–372

Chapter 58

Integration of GNSS and MEMS-Based Rotary INS for Bridging GNSS Outages

Shuang Du, Wei Sun and Yang Gao

Abstract In a GNSS/INS integrated system, the GNSS outputs, such as position and velocity, are used to estimate the INS navigation errors and inertial sensor errors and the inertial solutions are used to bridge the GNSS outages for which a good quality IMU is needed. Since MEMS IMU outputs are corrupted by significant sensor errors, the navigation errors will accumulate quickly which degrade the navigation solution over a short time period in the presence of GNSS outages. The rotary INS technique has been proposed to reduce the MEMS-based INS navigation errors recently. A rotary INS is an inertial navigator in which the IMU is installed on a rotation platform. By modulating the significant inertial sensor errors through proper IMU rotations, the accumulation of navigation errors can be effectively mitigated without a need for external aiding. Based on the previous work on MEMS-based rotary INS, this paper proposed an integrated system of GNSS and rotary INS. Given the ability of a rotary INS to extend autonomous navigation, the integrated system can significantly improve the navigation performance to bridge GNSS outages under GNSS-challenging environments, such as vehicles travelling through tunnels. Based on a single-axis rotation table, road kinematic tests are conducted to evaluate the performance of the proposed system using two different MEMS IMUs (MTi-G from Xsens and NAV440 from Moog Crossbow). The results indicate that the IMU rotation can reduce the horizontal position errors by about 2 times for the system with MTi-G and about 3 times for the system with NAV440 during the period of GNSS outages.

Keywords Rotary INS · MEMS IMU · Integration · Error mitigation

S. Du (✉) · Y. Gao
Geomatics Engineering, University of Calgary, Calgary, Canada
e-mail: sdu@ucalgary.ca

W. Sun
Survey Engineering, Liaoning Technical University, Jinzhou, China

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_58

659

58.1 Introduction

The integrated systems of GNSS and INS are now widely based on MEMS IMU due to its small-size, low-cost, light-weight and low-power consumption. Such systems are widely used in a variety of applications including indoor pedestrian navigation, autonomous navigation embedded system for land and air vehicles, robotic navigation and precision agriculture. In a GNSS/INS integrated system, the GNSS outputs, such as position and velocity, are used to estimate the INS navigation errors and inertial sensor errors and the inertial solutions are used to bridge the GNSS outages for which a good quality IMU is needed [1, 4, 5, 7]. Since MEMS IMU outputs are corrupted by significant sensor errors, such as high frequency noise, bias, scale factors and installation errors, the navigation errors will accumulate quickly which degrade the navigation solution over a short time period in the presence of GNSS outages. For example, the position errors of a low-cost MEMS IMU may grow to a few kilometers within minutes when no external aiding is available during GNSS outages. The rotary INS technique has been proposed to reduce the MEMS-based INS navigation errors recently [6, 15]. A rotary INS is an inertial navigator in which the IMU is installed on a rotation platform. By modulating the significant inertial sensor errors through proper IMU rotations, the accumulation of navigation errors can be effectively mitigated without a need for external aiding. For example, the rotation of IMU with a constant angular rate can modulate the constant inertial bias into periodic signals and an integration of the modulated inertial data over a complete rotation cycle can eliminate the bias impact on the navigation solutions. Although this technique had been applied to high-end inertial sensors, including ring laser gyro (RLG) and fiber optic gyro (FOG), for submarines and warships, it only started to be employed for low-cost MEMS IMU recently.

Based on the previous work on MEMS-based rotary INS, this paper proposed an integrated system of GNSS and rotary INS. Given the ability of a rotary INS to extend autonomous navigation, the integrated system can significantly improve the navigation performance to bridge GNSS outages under GNSS-challenging environments, such as vehicles travelling through tunnels. This paper will first introduce the rotary INS and analyze the modulation of inertial sensor errors as well as mitigations on navigation errors. Then the integrated system of GNSS and rotary INS is described including the integration scheme and filter design. Based on a single-axis rotation table, road kinematic tests are conducted to evaluate the performance of the proposed system using two different MEMS IMUs (MTi-G from Xsens and NAV440 from Moog Crossbow). The results indicate that the IMU rotation can reduce the horizontal position errors by about 2 times for the system with MTi-G and about 3 times for the system with NAV440 during the period of GNSS outages.

The remainder of this paper is organized as follows. Section 58.2 describes the concept of rotary INS, including the mechanization algorithm and the error mitigation principle, and Sect. 58.3 presents the integrated system of GNSS and

rotary INS, with details of system flowchart and integration filter design. The conducted kinematic field tests and results analysis are given in Sect. 58.4, followed by conclusion given in Sect. 58.5.

58.2 Rotary INS

The concept of rotary INS was initially proposed for the gimballed inertial system by Geller [8]. Afterwards, this technique was applying to strapdown inertial system [9, 11]. As certain rotations of IMU can mitigate the INS navigation errors without the requirement of external information, the rotary INS was extensively employed for the worship or submarine. The rotary INS normally contains two components, an IMU and a rotational platform, as shown in Fig. 58.1.

58.2.1 Rotary INS Mechanization

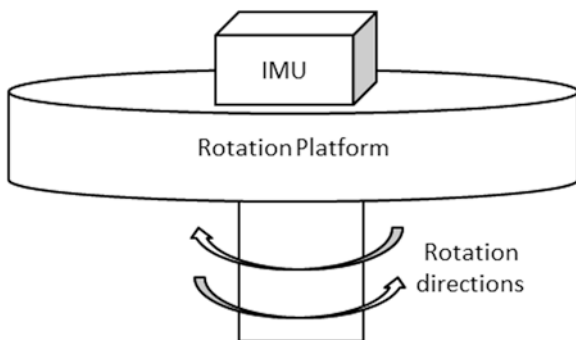
As the IMU is rotating in the rotary INS, a new frame in which the inertial readings are collected is introduced, in addition to the normally employed coordinate frames in non-rotary INS. This new coordinate frame can be referred as inertial sensor frame or IMU frame, and its axes are aligned with the sensitive axes of inertial sensors with the origin defined as the origin of IMU.

Generally speaking, the mechanization algorithm of rotary INS is very similar to what is used in non-rotary INS. As the inertial readings are collected in sensor frame, a transformation process is required as shown in Eqs. (58.1) and (58.2).

$$f_{ib}^b = C_s^b (f_{is}^s + f_{sb}^s) \tag{58.1}$$

$$\omega_{ib}^b = C_s^b (\omega_{is}^s + \omega_{sb}^s) \tag{58.2}$$

Fig. 58.1 Structure of rotary INS



where f_{ib}^b and ω_{ib}^b are the specific force and angular rate in body frame with respect to inertial frame, respectively, C_s^b is the transformation matrix from sensor frame to body frame, f_{is}^s and ω_{is}^s are the specific force and angular rate in sensor frame with respect to inertial frame, respectively, f_{sb}^s and ω_{sb}^s are the specific force and angular rate of body frame with respect to sensor frame, respectively. As the rotation of IMU does not introduce any linear movement, f_{sb}^s is a zero vector, while ω_{sb}^s is related to the IMU rotation schemes. Apparently the rotation angle between body frame and sensor frame is required in the transformation process, which is usually measured by a device installed in the rotation platform. With the transformed specific force and angular rate in body frame, the mechanization of the non-rotary INS can be used to derive position, velocity and attitude solutions. More details can be referred to Du et al. [6].

58.2.2 Rotary INS Error Mitigation

If the IMU rotates about its X axis with a rate of ω and the body frame is aligned with navigation frame, the gyro biases in the navigation frame can be described by Eq. (58.3) [2, 14]. Apparently, the biases in north and vertical directions are modulated into periodic signals, and the attitude errors caused by such errors are self-eliminated after a complete rotation cycle as shown in Eq. (58.4). As the bias in the rotation axis cannot be modulated, the attitude error in the east direction propagates in the same way as in non-rotary INS.

$$d^n = C_b^n C_s^b d^s = \begin{bmatrix} d_x^s \\ d_y^s \cos \omega t + d_z^s \sin \omega t \\ -d_y^s \sin \omega t + d_z^s \cos \omega t \end{bmatrix} \tag{58.3}$$

$$\int_0^T d^n dt = \begin{bmatrix} T d_x^s \\ 0 \\ 0 \end{bmatrix} \tag{58.4}$$

where $d^s = [d_x^s \ d_y^s \ d_z^s]^T$ and $d^n = [d_E \ d_N \ d_U]^T$ are the gyro biases in sensor frame and navigation frame, respectively, T is the time span of a complete rotation cycle, and C_b^n and C_s^b are the transformation matrices among navigation, body and sensor frames.

The accelerometer biases can also be modulated through IMU rotation, and the resulted velocity errors can be mitigated. According to Du et al. [6], the following findings were obtained: (1) the IMU rotation modulates the constant biases of inertial sensors that perpendicular to the rotation axis, and the attitude and velocity errors caused by such biases are self-eliminated after a complete rotation cycle; (2) the constant biases of inertial sensors in the rotation axis cannot be modulated, and the attitude and velocity errors caused by such errors propagate in the same way as in non-rotary INS; (3) the IMU rotation induces an extra error in the gyro of the

rotation axis due to gyro scale factor, and this error results in accumulated attitude errors in the direction of the corresponding rotation axis; (4) the IMU rotation also induces extra errors in the gyros that are perpendicular to the rotation axis due to gyro installation errors, resulting in attitude and velocity errors.

58.3 Integration of GNSS and MEMS-Based Rotary INS

Although the navigation errors can be effectively reduced by rotation of IMU, they still accumulate over time [6, 15]. Therefore, aiding from other sensors, such as GNSS, is still necessary to maintain the high accuracy navigation solutions. This section presents an integrated system of GNSS and MEMS-based rotary INS. Comparing to the integrated system with non-rotary INS, the proposed integrated system with rotary INS significantly reduces the navigation errors during GNSS outages, and therefore offers longer autonomous navigation performance without external aiding.

58.3.1 System Flowchart

The GNSS and rotary INS are loosely integrated in this research due to its simplicity of implementation. The flowchart of the integrated system is given in Fig. 58.2. The accelerometer and gyro readings collected in the sensor frame are first transformed to the body frame. Then the INS solutions are derived using mechanization algorithm based on the transformed inertial data. Next, an EKF is employed to fuse the position and velocity solutions from both GNSS and rotary INS to derive the corrections for position, velocity and attitude solutions as well as the inertial sensor errors. Finally, the system outputs the corrected INS position, velocity and attitude solutions.

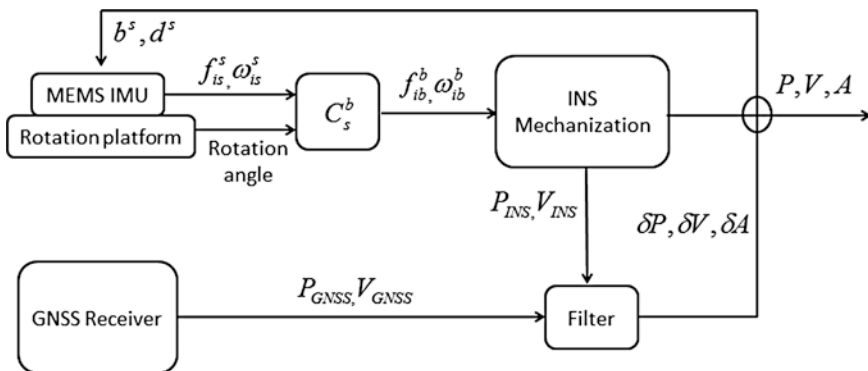


Fig. 58.2 System flowchart of integrated GNSS and rotary INS

58.3.2 Filter Design

A 15-state vector, including 3 position errors in the navigation frame, 3 velocity errors in the navigation frame, 3 attitude errors in the navigation frame, 3 accelerometer biases in sensor frame and 3 gyro biases in sensor frame, is employed to represent the system errors as shown in Eq. (58.5).

$$x = [\delta r^n \quad \delta v^n \quad \varepsilon^n \quad b^s \quad d^s]^T \quad (58.5)$$

where $\delta r^n = [\delta\varphi \quad \delta\lambda \quad \delta h]^T$ represents the position errors in navigation frame, $\delta v^n = [\delta v^E \quad \delta v^N \quad \delta v^U]^T$ and $\varepsilon^n = [\varepsilon^E \quad \varepsilon^N \quad \varepsilon^U]^T$ represents the velocity and attitude errors in navigation frame, b^s represents the accelerometer biases in sensor frame.

Although navigation errors caused by sensor errors can be mitigated through IMU rotation in rotary INS, the sensor errors remain the same which require the development of a system error model. As the rotation of IMU does not introduce any linear movement, the error model of rotary INS is the same as the one used in non-rotary INS, which can be described by a series of differential equations as follows [3, 10, 12].

$$\delta \dot{r}^n = F_{rr}\delta r^n + F_{rv}\delta v^n \quad (58.6)$$

$$\delta \dot{v}^n = F_{vr}\delta r^n + F_{vv}\delta v^n + F_{ve}\varepsilon^n + \delta f^n \quad (58.7)$$

$$\dot{\varepsilon}^n = F_{er}\delta r^n + F_{ev}\delta v^n + F_{ee}\varepsilon^n + \delta\omega^n \quad (58.8)$$

where F_{rr} , F_{rv} , F_{vr} , F_{vv} , F_{ve} , F_{er} , F_{ev} , F_{ee} are the relation matrices among position error, velocity error and attitude errors, δf^n and $\delta\omega^n$ represent the accelerometer and gyro errors, respectively.

The 1st order Gauss-Markov random process is used to model the random variation for both accelerometer and gyro biases. The IMU rotation induced gyro biases are combined with the original gyro biases and estimated in the filter. In a loose integration scheme, the GNSS position and velocity are used as measurements to estimate the INS errors, so the measurement model of the integration filter can be described by Eq. (58.9). More details of the implementation of an EKF can be referred to Shin [13] and Du [3].

$$H = \begin{bmatrix} I_{3\times 3} & 0_{3\times 3} & 0_{3\times 3} & 0_{3\times 3} & 0_{3\times 3} \\ 0_{3\times 3} & I_{3\times 3} & 0_{3\times 3} & 0_{3\times 3} & 0_{3\times 3} \end{bmatrix} \quad (58.9)$$

58.4 Kinematic Field Tests and Results Analysis

Kinematic field tests are conducted based on a land vehicle to study the navigation performance of the proposed integrated system. By using a single-axis rotation platform, two MEMS IMUs, MTi-G and NAV440, are tested. The single-axis rotation platform includes a rotation table and a console as shown in Fig. 58.3. The MEMS IMUs are firmly installed on top of the platform by using screws as shown in Fig. 58.4. Only IMU rotation about Z axis can be implemented with this rotation table, and its technical parameters are given in Table 58.1.

In addition to the single-axis rotation platform and two MEMS IMUs, the NovAtel SPAN system, which includes a high-end LCI IMU and a SE GNSS receiver, is employed to collect the data to generate the reference solutions. Other equipments include the power supply and the antennas. The equipment set-up on the test vehicle is given in Fig. 58.5. The single-axis rotation table on which the IMU installed and the power supply are placed at the trunk of the vehicle, the console and NovAtel SPAN system are placed on the back seat row. The body frame of MEMS IMU is aligned to the body frame of SPAN before the rotation of IMU. The antenna for the SPAN is placed on the top of vehicle as shown in the figure. The SPAN data including both GNSS data and inertial data are collected with data rate of 1 and 100 Hz, respectively, and stored in the SD card of GNSS receiver. The MEMS IMU data and rotation angle data are collected with data rate of 100 and 50 Hz by the rotation table console, respectively.



Fig. 58.3 Single-axis rotation platform

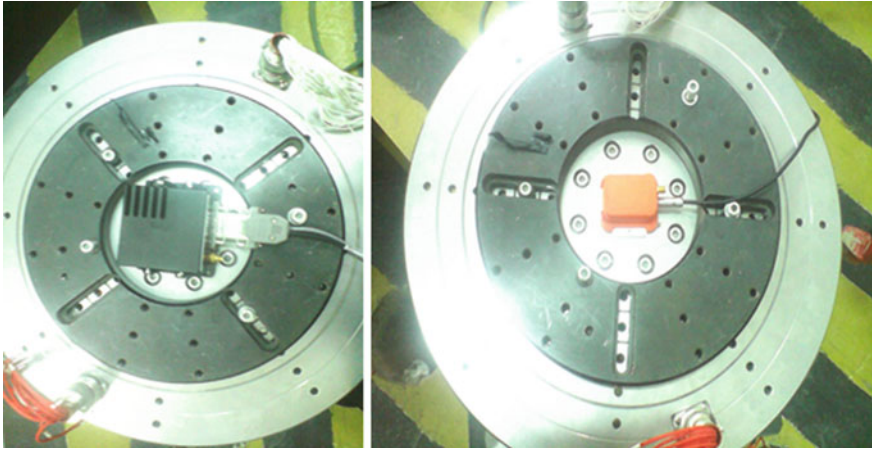


Fig. 58.4 Installations of MTi-G and NAV440 on the single-axis rotation platform

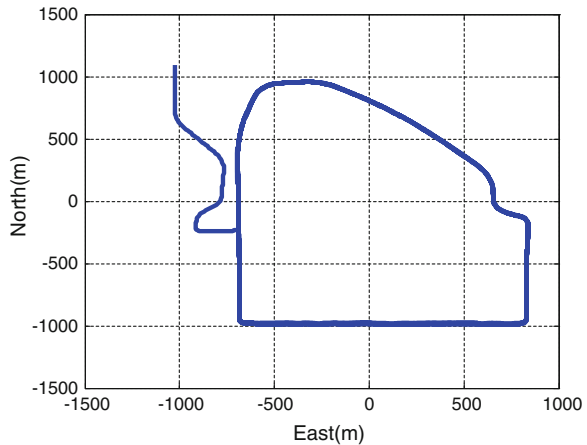
Table 58.1 Technical parameters of single-axis rotation platform

Position accuracy (°)	Rotation Rate accuracy (°/s)	Maximum rotation rate (°/s)
$8e^{-4}$	$5e^{-5}$	± 100



Fig. 58.5 Equipment set-up on the test vehicle

Fig. 58.6 Trajectory of kinematic field tests



58.4.1 Field Test Descriptions and Data Processing Strategy

The employed trajectory for the field tests includes rectangular turn, circle and straight line as shown in Fig. 58.6. Two types of field tests, namely tests of integrated system with non-rotary INS and tests of integrated system with rotary INS, are conducted and more details are given as follows.

58.4.1.1 Tests of Integrated System of GNSS and Non-rotary INS

The test starts with an initialization process during which the vehicle remains static for 5 min to solve the initial attitude solutions using SPAN system. Then the vehicle remains still for 30 s before it travels along the designated trajectory. During the whole tests, the MEMS IMU remains static relative to the vehicle (the rotating table is non-rotating). Not including the initialization time, the test lasts for about 33 min (30 s static and 32.5 min kinematic). The collected GNSS data of SPAN is processed by Waypoint GrafNav from NovAtel to derive the position and velocity information, which is loosely integrated with the MEMS IMU data. The reference solutions are generated by processing the SPAN data (including both GNSS data and LCI data) using Waypoint Inertial Explorer. MTi-G and NAV440 are employed in the tests separately. The 33-min MEMS IMU data is processed with GNSS data to derive the solutions for integrated system of GNSS and non-rotary INS, with initial attitude obtained from SPAN.

58.4.1.2 Tests of Integrated GNSS and Rotary INS

After the same initialization process, the vehicle remains still for about 30 s and then was driven along the trajectory. The MEMS IMU remains still during

initialization process and then starts rotating about Z axis with designated rotation rates after. The test also lasts for 33 min excluding the time for initialization process. The position and velocity which are loosely integrated with MEMS IMU are still obtained by processing the collected GNSS data of SPAN using Waypoint GrafNav, and the reference solutions are obtained in the same way as in 1). MTi-G and NAV440 are employed in the tests separately. The designated rotation rate for MTi-G is 10°/s, while the rate for NAV440 is 20°/s. The 33-min MEMS IMU data is processed with GNSS data to derive the solutions for integrated system of GNSS and rotary INS, with initial attitude obtained from SPAN.

58.4.2 Navigation Errors of Integrated Systems Without GNSS Outages

58.4.2.1 Navigation Errors of Integrated Systems Based on MTi-G

The attitude errors of the integrated systems are given in Fig. 58.7. With continuous GNSS measurements, roll and pitch errors are well bounded in 0.5° in most of time for both integrated systems. During the initial static period, the azimuth errors

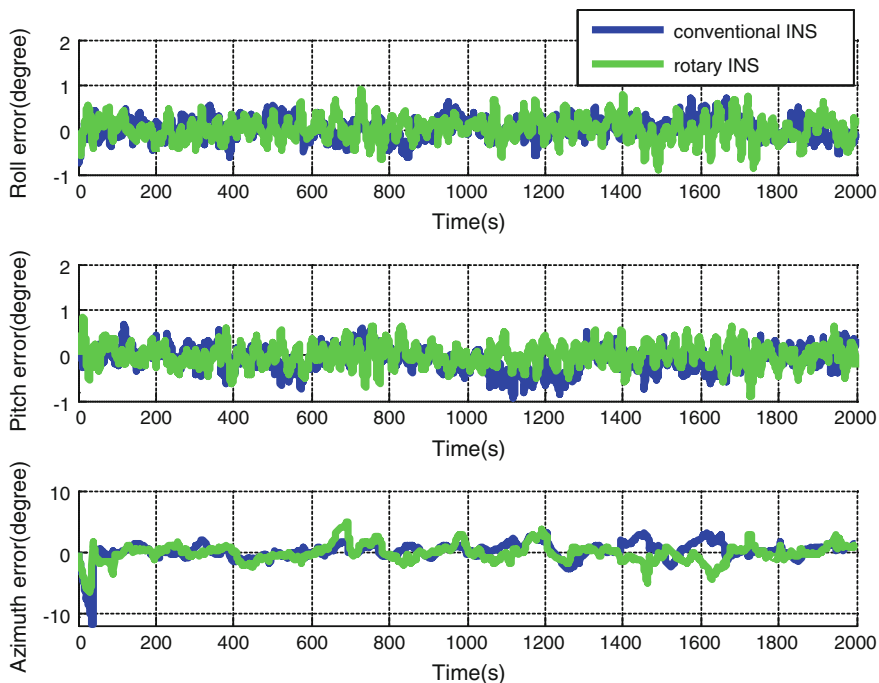


Fig. 58.7 Attitude errors of integrated systems based on MTi-G

Table 58.2 RMS errors of attitude solutions for integrated systems with MTi-G

RMS	Roll (°)	Pitch (°)	Azimuth (°)
Integrated system with non-rotary INS	0.22	0.26	1.43
Integrated system with rotary INS	0.26	0.24	1.39

Table 58.3 RMS errors of position and velocity solutions for integrated systems with MTi-G

RMS	Latitude (cm)	Longitude (cm)	Height (cm)	Ve (cm/s)	Vn (cm/s)	Vu (cm/s)
Integrated system with non-rotary INS	2.8	2.5	1.7	3.2	3.3	1.2
Integrated system with rotary INS	2.8	2.6	1.7	3.2	3.3	1.2

accumulate in both integrated systems, and they are quickly reduced to around zero after the vehicle starts moving as shown in the figure. Table 58.2 summarizes the RMS errors of attitude errors for both integrated systems based on MTi-G. As the vehicle maneuvers significantly improve the system observability, the two systems offer very similar attitude solutions.

Table 58.3 summarizes the RMS errors of position and velocity errors. As the GNSS position and velocity integrated with MEMS IMU data are derived from the SPAN data that used to generate the reference solutions, the solutions of both integrated systems are highly consistent with the reference solutions.

58.4.2.2 Navigation Errors of Integrated Systems Based on NAV440

Figure 58.8 presents the attitude errors of the integrated system based on NAV440. Similarly, the roll and pitch errors are well bounded within small values for both integrated systems. The azimuth errors are also well bounded within 1° for most of the time. Table 58.4 summarizes the RMS errors of attitude solutions. Both integrated systems offer similar results.

Given small attitude errors, the position and velocity solutions are highly consistent with the reference solutions for both integrated systems as shown in Table 58.5, which summarizes the RMS errors for position and velocity solutions.

58.4.3 Navigation Errors of Integrated Systems During GNSS Outages

The GNSS outages are simulated to study the navigation performance of the integrated systems when GNSS measurements are not available, such as vehicles travelling through tunnels. Three outages are simulated in the GNSS data collected

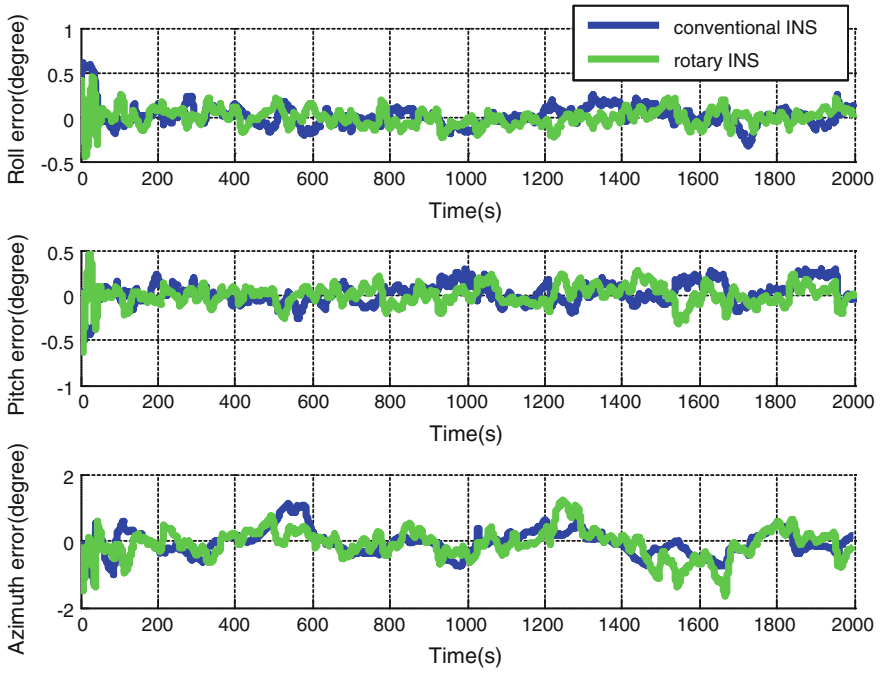


Fig. 58.8 Azimuth errors of integrated systems based on NAV440

Table 58.4 RMS errors of attitude solutions for integrated systems based on NAV440

RMS	Roll (°)	Pitch (°)	Azimuth (°)
Integrated system with non-rotary INS	0.12	0.12	0.39
Integrated system with rotary INS	0.11	0.12	0.44

Table 58.5 RMS errors of position and velocity errors for integrated systems with NAV440

RMS	Latitude (cm)	Longitude (cm)	Height (cm)	Ve (cm/s)	Vn (cm/s)	Vu (cm/s)
Integrated system with non-rotary INS	1.9	1.8	1.5	1.9	1.9	1.3
Integrated system with rotary INS	1.9	1.7	1.5	1.8	1.9	1.3

in each field test, respectively, and the three corresponding sections of the trajectory for different tests are the same. This can evaluate the navigation performance of integrated systems during GNSS outages under similar dynamic conditions.

58.4.3.1 Navigation Errors of Integrated Systems Based on MTi-G During GNSS Outages

The roll and pitch errors of both integrated systems based on MTi-G during the 1st outage are given in Fig. 58.9. As the gyro biases in X and Y axes are modulated through IMU rotation about Z axis, the roll and pitch errors of integrated system with rotary INS are kept within 0.5° most of the time, while such errors in the integrated system with non-rotary INS would drift over 1° frequently.

The velocity errors in the east-north plane during the 1st outage are given in Fig. 58.10 for both rotary and non-rotary systems. With the modulation of accelerometer biases and reduced roll and pitch errors, the velocity errors are also reduced in the integrated system with rotary INS.

The RMS errors of the PVA solutions during all outages are calculated and summarized in Table 58.6. The horizontal position and velocity errors of the integrated system with rotary INS are reduced about 2 times comparing to the errors of the system with non-rotary INS.

Figure 58.11 presents RMS of horizontal position and velocity errors versus the outage time. With IMU rotation, the rotary INS effectively damps the accumulation of position and velocity errors.

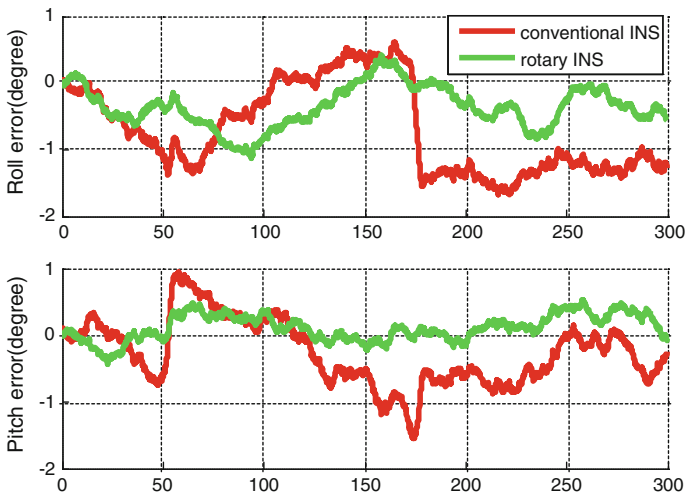


Fig. 58.9 Roll and pitch errors of the integrated systems based on MTi-G for the 1st GNSS outage

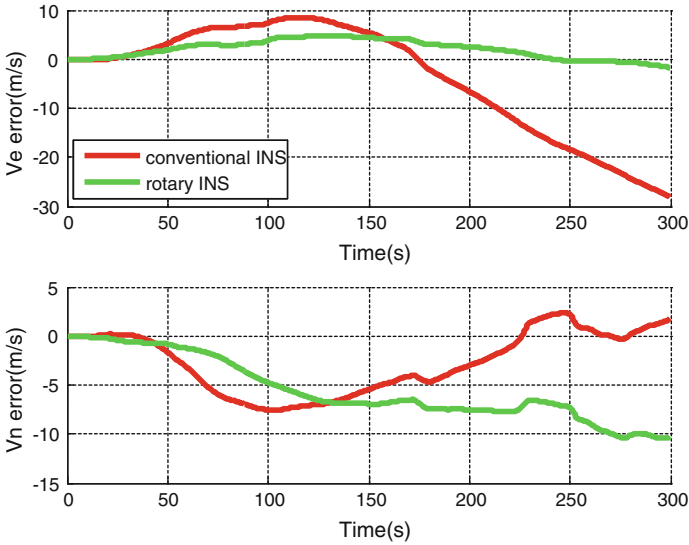


Fig. 58.10 Velocity errors in east-north plane of the integrated systems based on MTi-G during the 1st GNSS outage

Table 58.6 RMS errors of PVA solutions of integrated systems with MTi-G for all GNSS outages

RMS	HP (m)	HV (m/s)	Roll (°)	Pitch (°)	Azimuth (°)
Integrated system with non-rotary INS	3458.3	31.9	0.83	0.75	2.39
Integrated system with rotary INS	1645.3	17.3	0.59	0.57	1.67

58.4.3.2 Navigation Errors of Integrated Systems Based on NAV440 During GNSS Outages

The roll and pitch errors of both integrated systems based on NAV440 during the 1st GNSS outage are given in Fig. 58.12. As expected, the modulation of gyro biases effectively reduces the roll and pitch errors of the integrated system with rotary INS through IMU rotation.

Figure 58.13 presents the velocity errors in the east-north plane during the 1st outage for both rotary and non-rotary systems. The RMS errors of the PVA solutions during all outages are calculated and given in Table 58.7. Although the horizontal position errors are reduced from more than a thousand meters to about four hundred meters, no obvious improvements are observed in azimuth solutions. Figure 58.14 presents the RMS of horizontal position and velocity errors versus the outage time for the integrated systems based on NAV440.

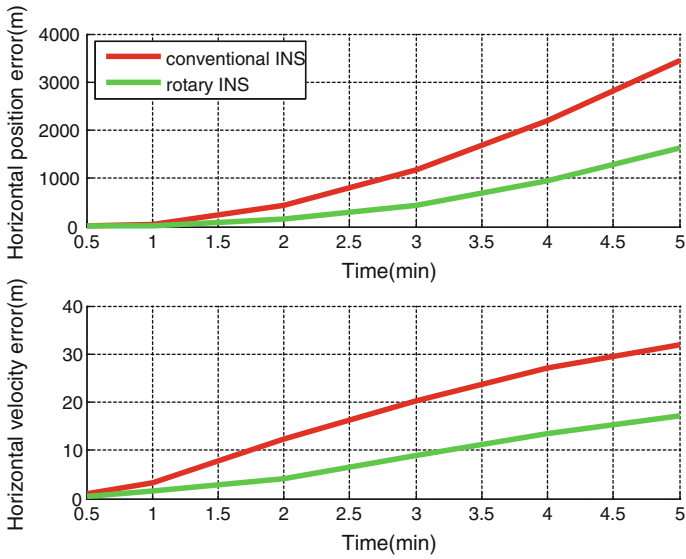


Fig. 58.11 RMS errors of horizontal position and velocity solutions of integrated systems based on MTi-G vs. time

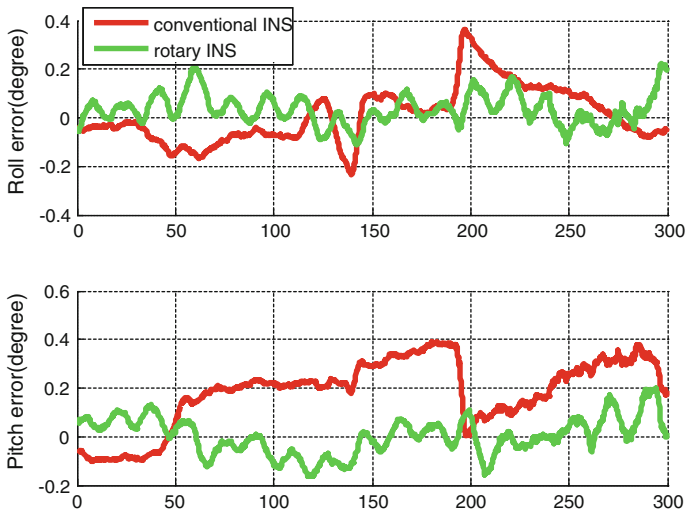


Fig. 58.12 Roll and pitch errors of the integrated systems with NAV440 for the 1st GNSS outage

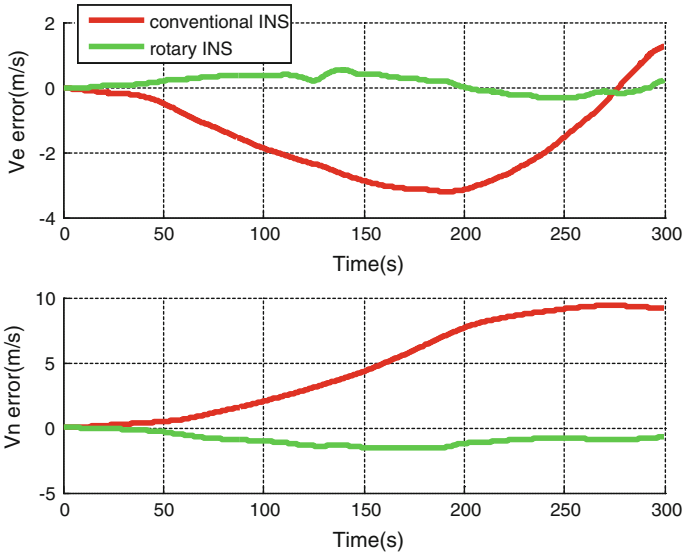


Fig. 58.13 Horizontal position errors of the integrated systems with NAV440 for the 1st GNSS outage

Table 58.7 RMS errors of PVA solutions of integrated systems based on NAV440 for all GNSS outage

RMS	HP (m)	HV (m/s)	Roll (°)	Pitch (°)	Azimuth (°)
Integrated system with non-rotary INS	1095.0	14.4	0.64	0.50	0.42
Integrated system with rotary INS	440.2	4.8	0.13	0.18	0.76

Based on the kinematic field test results, we can have following findings: (1) both integrated systems with rotary and non-rotary INS offers similarly PVA solutions with continuous GNSS measurements; (2) through sensor error modulation by IMU rotation about Z axis, the navigation errors in the east-north plane for the integrated system with rotary INS are significantly reduced during GNSS outages, and therefore it can offer longer autonomous navigation and more effectively bridge GNSS outages.

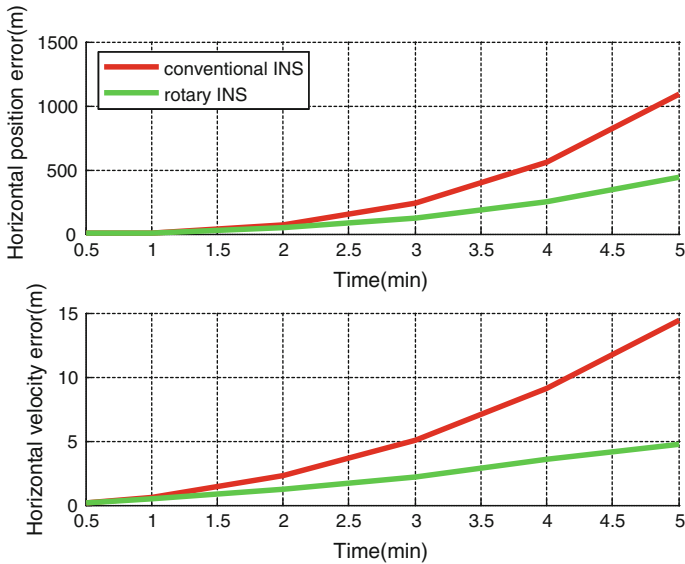


Fig. 58.14 RMS errors of horizontal position and velocity solutions of integrated systems based on NAV440 versus time

58.5 Conclusions

This paper proposed an integrated system of GNSS and MEMS-based rotary INS. After viewing the rotary INS, the integrated system is described with focus on the system flowchart and integration filter design. As IMU rotation about Z axis can effectively mitigate navigation errors in the east-north plane, such rotation scheme has been employed for the proposed integrated system. Based on a single-axis rotation table, kinematic field tests are conducted with two MEMS IMUs, namely, MTi-G and NAV440, to verify the navigation performance of the proposed integrated system. The results indicate that the IMU rotation can reduce the horizontal position errors by about 2 times for the system with MTi-G and about 3 times for the system with NAV440 during the period of GNSS outages.

References

1. Antonio A (2010) GNSS/INS integration methods. Ph.D. dissertation, Department of Geomatics Engineering, The University of Calgary, Canada
2. Ben YY, Chai YL, Gao W, Sun F. (2010), Analysis of error for a rotating strap-down inertial navigation system with fiber gyro. *J Marine Sci Appl* 9:419–424. doi: [10.1007/s11804-010-1028-z](https://doi.org/10.1007/s11804-010-1028-z)
3. Du S (2010) Integration of precise point positioning and low cost MEMS IMU. M.Sc. thesis, Department of Geomatics Engineering, The University of Calgary, Canada

4. Du S, Gao Y (2012) Inertial aided cycle slip detection and identification for integrated PPP GPS and INS. *Sensors* 12(11):14344–14362. doi: [10.3390/s121114344](https://doi.org/10.3390/s121114344)
5. Du S, Gao Y (2012). Integration of precise point positioning and the latest MEMS IMU for precise applications. In: Proceedings of ION GNSS 2012, Nashville, Tennessee, USA, 18–21 Sept 2012
6. Du S, Sun W, Gao Y (2014) An investigation on MEMS IMU error mitigation using rotation modulation technique. In: Proceedings of ION GNSS 2014, Tampa, Florida, USA, 8–12 Sept 2012
7. Fang J, Gong X (2010) Predictive iterated Kalman filter for INS/GPS integration and its application to SAR motion compensation. *IEEE Trans Instrum Meas* 59:909–915
8. Geller ES (1968) Inertial system platform rotation. In: *IEEE transactions on aerospace and electronic systems*, AES-4, July 1968, pp 627–634
9. Giovanni CS, Levinson E (1981) Performance of a ring laser strapdown marine gyrocompass. In: ION 7th annual meeting proceedings, Annapolis, Maryland, USA
10. Jekeli C (2001) *Inertial Navigation System with Geodetic Applications*. Walter de Gruyter GmbH and Co
11. Levinson DE, Giovanni SC (1980) Laser gyro potential for long endurance marine navigation. In: *IEEE position, location and navigation symposium*. Institute of Electrical and Electronics Engineers, New York, pp 115–129
12. Shin E (2001) Accuracy improvement of low cost INS/GPS for land application. MSc thesis, Department of Geomatics Engineering, The University of Calgary, Canada
13. Shin E (2005) Estimation techniques for low-cost inertial navigation. PhD thesis, Department of Geomatics Engineering, University of Calgary, Canada
14. Sun F, Sun W, Gao W, Ben YY (2009) Research on the technology of rotational motion for FOG strapdown inertial navigation system. In: Proceedings of the 2009 IEEE international conference on mechatronics and automation, Changchun, China, 9–12 Aug
15. Sun W, Wang DX, Xu LW, Xu LL (2013) MEMS-based rotary strapdown inertial navigation system. *Measurement* 46:2585–2596

Chapter 59

An Integrated PDR/GNSS Pedestrian Navigation System

Haiyu Lan, Chunyang Yu and Naser El-Sheimy

Abstract This paper is concerned with the estimation of heading information of the Pedestrian Navigation System (PNS). The MEMS inertial sensors as well as a miniature GNSS receiver are used to establish a pedestrian navigation prototype based on the Pedestrian Dead Reckoning (PDR) approach. An Extended Kalman Filter (EKF) structure is used for the estimation of the system's attitude error and the bias of the gyroscope. If there is no external acceleration, errors of pitch and roll as well as the biases of the two horizontal gyros are compensated using the aiding information from the accelerometer. When GNSS is available, its output is used for heading and heading-gyro bias estimation. Using the aiding information provided by both GNSS and accelerometer, the proposed method prevents the errors in the attitude from increasing rapidly. The proposed method for pedestrian navigation application has been well verified through real field experiments.

Keywords MEMS · Pedestrian navigation system · Pedestrian dead reckoning extended kalman filter (EKF) · GNSS · Accelerometer

59.1 Introduction

Nowadays, Micro-Electro-Mechanical System (MEMS) inertial sensors and miniature Global Navigation Satellite System (GNSS) receivers can be easily found in various kinds of smart-devices for consumers. Those devices including smart-phones, smart-watches, tablet computers, etc., are now widely used for Pedestrian Navigation System (PNS) applications.

H. Lan (✉)

College of Automation, Harbin Engineering University, Harbin 150001, China
e-mail: hlan@ucalgary.ca

H. Lan · C. Yu · N. El-Sheimy

Department of Geomatics Engineering, The University of Calgary,
2500 University Dr. N.W. Calgary, AB, Canada

© Springer-Verlag Berlin Heidelberg 2015

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015*

Proceedings: Volume III, Lecture Notes in Electrical Engineering 342,

DOI 10.1007/978-3-662-46632-2_59

Since MEMS inertial sensors still have relative high-noise ratio, the use of Inertial Navigation System (INS) for PNS is very limited. MEMS sensors cannot yet accurately track the position of a person to sub-meter level for any period greater than 10 s using INS mechanization alone.

Considering the unique features of human walking steps, an effective approach for PNS using MEMS inertial sensors is the Pedestrian Dead Reckoning [1]. The main idea of PDR is to use accelerometers to detect the pedestrian's travelling steps, to estimate step length, and to propagate the user's position using a measured heading between every two consecutive steps. Till now, calculating an accurate long-term heading remains the most challenging problem in PDR. Heading calculation using magnetometers is not a good choice because magnetometers are subject to strong magnetic disturbances such as power lines, metal structures, computers, etc., [2]. On the other hand, heading can be calculated using gyroscopes [3]. However, due to the large drift rate of MEMS gyroscope, such an operation is only active when the PNS works for short period of time and no severe swing or swaying movements of PNS occur.

In this paper, we introduce the concept of using the aiding information from accelerometers (when there is no external acceleration) and GNSS (when available) for the heading estimation of PNS. The advantages of using accelerometers is that the calculation of roll and pitch is more accurate when aided by accelerometers instead of using stand-alone gyroscopes. To demonstrate the superiority of this approach, we consider the situation of one of the bias-uncompensated horizontal gyros [4]. Due to the existing bias of the gyro, an angle error proportional to time arising from the gyro's angular-rate integration process will be introduced in roll or pitch estimation. Thus, a misalignment of the INS will be generated that in turn will incorrectly project the vertical gyro output (used for heading calculation) to the local-level coordinate. This wrong projection will introduce an accumulated error in the heading calculation. When the errors of pitch and roll as well as the biases of the two horizontal gyros are compensated by the aiding information from accelerometers. The error during the gyro's angular-rate integration will be eliminated and therefore error in roll or pitch is not proportional to time. Moreover, the advantage of using GNSS is that the heading error as well as the heading-gyro's bias could be estimated given the fact that PDR-based PNS is usually mounted on the pedestrian's upper body, which does not block the GNSS signals. In addition, PDR/GNSS could provide a PNS with superior navigation performance in comparison with either a stand-alone GNSS receiver or a PDR.

The remainder of this paper is organized as follows: First, the PDR mechanization is discussed. The step detection is based on the accelerometer output, the heading is updated by the gyroscope, and the travelled distance is computed by merging the signal from the accelerometer with a physiological model. Then the complete EKF-based attitude estimation method is shown in detail in Sect. 3. Finally, experiments were done to test the proposed method.

59.2 PDR Concept

Pedestrian Dead Reckoning (PDR) is implemented through exploiting the kinematic features of a pedestrian's gait with the traveled distance and heading information. Essentially, the PDR is the determination of a new position utilizing the knowledge of a previous known position together with the current travelled distance and heading information. Generally, PDR consists of three important components: (1) step detection; (2) stride length estimation; and (3) determination of the pedestrian's heading during every two consecutive steps. The new pedestrian's positioning coordinate with respect to a previously known coordinate can be recursively updated as follows:

$$\begin{cases} E_{k+1} = E_k + S_k \sin \psi_k \\ N_{k+1} = N_k + S_k \cos \psi_k \end{cases} \quad (59.1)$$

in which E_{k+1} is the PDR navigation estimation along local east at time step $k + 1$, N_{k+1} is the PDR navigation estimation along local north at time step $k + 1$, S_k is the estimated stride length from time step k to $k + 1$, ψ_k is the estimated heading at time step k . Details of the PDR mechanization are discussed in the following parts.

59.2.1 Step Detection

Step detection is an essential procedure to finally provide an accurate stride length estimation. Steps are usually counted by detecting peak and valley values of the output of accelerometers [5]. And the selection of the acceleration signal could be different depending on the techniques being used. In this paper, we simply set up the following two rules to detect one step:

Rule 1:

Each step includes a peak (A_{\max}^{norm} , local maximum) and a valley (A_{\min}^{norm} , local minimum) from the acceleration amplitude. In other words, if we detect one peak and one contiguous valley, one step can be counted.

Rule 2:

The difference between an effective pair of peak and valley shall exceed a predefined threshold, which can be determined by offline training. That is to say, before entering navigation mode, a PNS user is asked to walk for a fixed number of steps several times to determine the threshold value.

59.2.2 Stride Length Estimation

According to [6], the stride length S is proportional to the vertical movement of the human hip. The vertical movement is obtained from the largest acceleration difference in vertical direction at each step. Refer to [6] the stride length S is calculated by taking the following formula,

$$S = K (A_{\max}^{\text{vert.}} - A_{\min}^{\text{vert.}})^{1/4} \quad (59.2)$$

where $A_{\max}^{\text{vert.}}$ and $A_{\min}^{\text{vert.}}$ refer to the maximum and minimum values of the vertical acceleration during a stride period, and K is a coefficient that can be trained when GNSS is available.

59.2.3 Heading Estimation

Another important factor for implementing the PDR mechanization is the calculation of heading. Typically, from the recently published literature, the heading of a pedestrian can be determined mainly through the following three schemes: (1) gyro-based scheme [3], (2) magnetometer-based scheme [2], and (3) GNSS-based scheme [7]. Gyro-based scheme provides a relative heading estimate, which indicates that an initial heading should be a prior known (normally provided by GNSS or personally). The gyro-derived heading is reliable only in a short period of time due to the accumulated error as a function of time. The magnetometer could provide long-term accurate heading. However, the main problem of this method is that magnetometer is subject to strong external magnetic disturbances such as power lines, metal/steel structures. Meanwhile, due to the signal blockage, GNSS-derived heading is not always available. So, it is imprudent to depend on a single heading-estimation technique. In this paper, the heading of the pedestrian is estimated through both gyro-based scheme and the GNSS-based scheme (when GNSS is available). According to [3], if the heading-gyro is well aligned with the body (e.g., the device is held levelly, the forward-axis is aligned with the pedestrian's walking direction), then heading can be estimated by integrating the heading-gyro's output during time t_1 to t_2 by

$$\psi = \int_{t_1}^{t_2} \omega_{\text{heading-gyro}}^n(t) dt \quad (59.3)$$

For arbitrary placement of the device, the heading-gyro should be in the direction of the accelerometer which is used in stride length estimation of Eq. (59.2). For instance, if z-axis accelerometer is used for stride length estimation, the z-axis gyro should then be used as the heading-gyro. However, there is no way

to guaranty that the device is perfectly levelled as well as placed aligned with the pedestrian's traveling direction, so the device is tilted on the human body. Consequently, there always exists a set of misalignment angles between the measured heading-gyro signal in the body-coordinate and the true attitude in the navigation coordinate, i.e., local-level coordinate. Those misalignment angles should be estimated and accounted for [3]. The transformation matrix (attitude matrix) from body coordinate to navigation coordinate is expressed as \mathbf{C}_b^n [8]. The measured triad-gyro vector can be transformed into navigation coordinate by

$$\boldsymbol{\omega}^n = \mathbf{C}_b^n \boldsymbol{\omega}^b \quad (59.4)$$

in which $\boldsymbol{\omega}^b$ and $\boldsymbol{\omega}^n$ are respectively the vectors of angular rates about the body coordinate and the navigation coordinate. The detailed expressions of \mathbf{C}_b^n , $\boldsymbol{\omega}^b$, and $\boldsymbol{\omega}^n$ are illustrated in the following Sect. 59.3. Therefore, it can be depicted from Eqs. (59.3) and (59.4) that the accuracy of the heading is determined by the accuracies of \mathbf{C}_b^n and $\boldsymbol{\omega}^b$. If the initial $\mathbf{C}_b^n(0)$ is given, it can be recursively updated by integrating the output gyroscopes. However, the estimation error of \mathbf{C}_b^n will diverge quickly due to the gyro bias and numerical integration errors [9]. On the other hand, according to [9, 10] the errors in \mathbf{C}_b^n and $\boldsymbol{\omega}^b$ can be compensated using external outputs of accelerometer and GNSS (when available). At this time, the fundamental issue in heading estimation is how to combine gyro, accelerometer, and GNSS output. In this paper, an indirect EKF-based approach is proposed. First, the triad-gyro output is integrated to determine a preliminary estimate of \mathbf{C}_b^n . The errors in \mathbf{C}_b^n and the bias in $\boldsymbol{\omega}^b$ are then estimated and compensated through the proposed EKF-prototype, when both the outputs of accelerometer and GNSS are available and reliable.

59.3 Attitude Estimation Using Accelerometer and GNSS

An EKF-based prototype is proposed to further improve the heading estimation accuracy. The attitude error and gyro bias can be well bounded through the output of accelerometer and GNSS. We first give the detailed derivation of the EKF system model in Sect. 59.3.1, followed by the measurement model in Sect. 59.3.2.

59.3.1 EKF System Model

The coordinate transformation from the navigation-coordinate to the sensor body-coordinate (x-y-z) is realized via the ordered Euler-angle rotations of yaw (ψ), pitch (θ), and roll (ϕ). In this paper, the first rotation is ψ about z, which is denoted by

$\mathbf{C}_z(\psi)$; the second rotation is $\mathbf{C}_y(\theta)$, θ about y; finally, the third rotation is $\mathbf{C}_x(\phi)$, ϕ about x. According to [11], $\mathbf{C}_z(\psi)$, $\mathbf{C}_y(\theta)$, and $\mathbf{C}_x(\phi)$ are respectively expressed as

$$\mathbf{C}_z(\psi) = \begin{bmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (59.5)$$

$$\mathbf{C}_y(\theta) = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix} \quad (59.6)$$

$$\mathbf{C}_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{bmatrix} \quad (59.7)$$

Then the transformation \mathbf{C}_n^b from navigation-coordinate to body-coordinate is calculated as the multiplication of the above three single-axis rotations,

$$\mathbf{C}_n^b = \mathbf{C}_x(\phi) \mathbf{C}_y(\theta) \mathbf{C}_z(\psi) \quad (59.8)$$

Expanding Eq. (59.8), we have

$$\mathbf{C}_n^b = \begin{bmatrix} \cos \theta \cos \phi & \cos \theta \sin \psi & -\sin \theta \\ \sin \phi \sin \theta \cos \psi - \cos \phi \sin \psi & \sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi & \sin \phi \cos \theta \\ \cos \phi \sin \theta \cos \psi + \sin \phi \sin \psi & \cos \phi \sin \theta \sin \psi - \sin \phi \cos \psi & \cos \phi \cos \theta \end{bmatrix} \quad (59.9)$$

Then the transformation \mathbf{C}_b^n from body-coordinate to navigation-coordinate is the transposition of \mathbf{C}_n^b , namely,

$$\mathbf{C}_b^n = \begin{bmatrix} \cos \theta \cos \phi & \sin \phi \sin \theta \cos \psi - \cos \phi \sin \psi & \cos \phi \sin \theta \cos \psi + \sin \phi \sin \psi \\ \cos \theta \sin \psi & \sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi & \cos \phi \sin \theta \sin \psi - \sin \phi \cos \psi \\ -\sin \theta & \sin \phi \cos \theta & \cos \phi \cos \theta \end{bmatrix} \quad (59.10)$$

The angular rates about the sensor body axes x, y, and z are denoted by ω_x^b , ω_y^b , and ω_z^b respectively, and their combined expression can be regarded as the vector,

$$\boldsymbol{\omega}^b = [\omega_x^b \quad \omega_y^b \quad \omega_z^b]^T \quad (59.11)$$

Similarly, the vector of the triad-angular rates along the navigation coordinate is expressed as

$$\boldsymbol{\omega}^n = [\omega_x^n \ \omega_y^n \ \omega_z^n]^T \tag{59.12}$$

The relationship between sensor body’s angular rates ($\omega_x^b, \omega_y^b, \omega_z^b$) and Euler angles (ψ, θ, ϕ) is given by

$$\begin{bmatrix} \omega_x^b \\ \omega_y^b \\ \omega_z^b \end{bmatrix} = \mathbf{C}_x(\phi) \mathbf{C}_y(\theta) \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix} + \mathbf{C}_x(\phi) \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} + \begin{bmatrix} \dot{\phi} \\ 0 \\ 0 \end{bmatrix} \tag{59.13}$$

Expanding the above Eq. (59.13), we have,

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \frac{1}{\cos \theta} \begin{bmatrix} 1 & \sin \phi \sin \theta & \cos \phi \sin \theta \\ 0 & \cos \theta \cos \theta & -\sin \phi \cos \theta \\ 0 & \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \omega_x^b \\ \omega_y^b \\ \omega_z^b \end{bmatrix} \tag{59.14}$$

Through the above Eq. (59.14) the attitude can be determined by integrating the angular rate measurements from gyroscopes, if the initial attitude is known. However, due to the reason that the gyro’ measurement containing gyro bias, directly integrating the gyro-derived measurement can yield large attitude error. We assume that the triad of gyro bias is expressed as $\mathbf{b} = [b_x \ b_y \ b_z]^T$, and the triad of gyro measurement is $\tilde{\boldsymbol{\omega}}^b = [\tilde{\omega}_x^b \ \tilde{\omega}_y^b \ \tilde{\omega}_z^b]^T$, at this time the true angular rates shown in Eq. (59.11) can be represented by,

$$\boldsymbol{\omega}^b = \begin{bmatrix} \omega_x^b \\ \omega_y^b \\ \omega_z^b \end{bmatrix} = \begin{bmatrix} \tilde{\omega}_x^b \\ \tilde{\omega}_y^b \\ \tilde{\omega}_z^b \end{bmatrix} - \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} \tag{59.15}$$

When the MEMS gyroscopes are used during a short time, shorter than the bias correlation time of the MEMS sensors, the sensor bias error can then be modeled as the random constant. In this paper, we model the triad of gyro biases as random constants which have the following continuous-time differential equation [59.8]:

$$\begin{cases} \dot{b}_x = 0 \\ \dot{b}_y = 0 \\ \dot{b}_z = 0 \end{cases} \tag{59.16}$$

Thus the model of the random constant bias can be implemented as a state-space in the EKF system model, namely we have the following state-space model,

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \\ \dot{b}_x \\ \dot{b}_y \\ \dot{b}_z \end{bmatrix} = \begin{bmatrix} \frac{1}{\cos \theta} \begin{bmatrix} 1 & \sin \phi \sin \theta & \cos \phi \sin \theta \\ 0 & \cos \theta \cos \theta & -\sin \phi \cos \theta \\ 0 & \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \tilde{\omega}_x^b - b_x \\ \tilde{\omega}_y^b - b_y \\ \tilde{\omega}_z^b - b_z \end{bmatrix} \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{59.17}$$

and the state vector is expressed as $\mathbf{x} = [\phi \ \theta \ \psi \ b_x \ b_y \ b_z]^T$.

59.3.2 EKF Measurement Model

According to [12], when there is no external acceleration, roll (ϕ) and pitch (θ) errors resulting from gyro biases can be mitigated by using the accelerometer output, namely, a triad of specific force, which is expressed as follows

$$\mathbf{f}^b = \begin{bmatrix} f_x^b & f_y^b & f_z^b \end{bmatrix}^T \tag{59.18}$$

Referring to [11], the specific force \mathbf{f}^b satisfies

$$\dot{\mathbf{v}}^n = \mathbf{C}_n^b \mathbf{f}^b - (2 \boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n) \times \mathbf{v}^n + \mathbf{g}^n \tag{59.19}$$

in which \mathbf{v}^n is the velocity vector expressed in navigation frame; $\dot{\mathbf{v}}^n$ is the translational acceleration vector in navigation frame; \mathbf{g}^n is the local gravity expressed as $\mathbf{g}^n = [0 \ 0 \ g]^T$. Neglecting the earth rotational rate $\boldsymbol{\omega}_{ie}$ and multiplying \mathbf{C}_n^b on both sides of Eq. (59.19) we have,

$$\dot{\mathbf{v}}^b = \mathbf{f}^b - \boldsymbol{\omega}_{eb}^b \times \mathbf{v}^b + \mathbf{g}^b \tag{59.20}$$

As we know, $\tilde{\boldsymbol{\omega}}^b = \boldsymbol{\omega}_{ib}^b = \boldsymbol{\omega}_{ie}^b + \boldsymbol{\omega}_{eb}^b$, neglecting the earth rotational rate, $\tilde{\boldsymbol{\omega}}^b$ can be written as $\tilde{\boldsymbol{\omega}}^b = \boldsymbol{\omega}_{eb}^b$, also when there is no external acceleration, namely $\dot{\mathbf{v}}^b = 0$, then Eq. (59.20) can be changed into,

$$\mathbf{f}^b = \tilde{\boldsymbol{\omega}}^b \times \mathbf{v}^b - \mathbf{g}^b \tag{59.21}$$

in which

$$\mathbf{g}^b = \mathbf{C}_n^b \mathbf{g}^n = g \begin{bmatrix} -\sin \theta \\ \sin \phi \cos \theta \\ \cos \phi \cos \theta \end{bmatrix} \tag{59.22}$$

and \mathbf{v}^b is the velocity vector along the sensor body frame and has the form,

$$\mathbf{v}^b = [v_x^b \ v_y^b \ v_z^b]^T \tag{59.23}$$

in which v_x^b is the pedestrian’s average walking speed during the last detected step, v_y^b and v_z^b are assumed to be zero. When there is no external acceleration, accelerometer output could provide accurate measurements to estimate the roll (ϕ) and pitch (θ) errors. When there is external acceleration, this is not the case. Thus smaller weights should be given to the accelerometer output with respect to gyroscope output when there is external acceleration. In [9], if the norm of accelerometer measurement is not near the gravitational acceleration g , it is determined that there is external acceleration and smaller weights are given to the accelerometer output by increasing the corresponding measurement noise covariance. Thus, the non-linear measurement can be written as,

$$\begin{bmatrix} f_x^b \\ f_y^b \\ f_z^b \end{bmatrix} = \begin{bmatrix} \tilde{\omega}_x^b - b_x \\ \tilde{\omega}_y^b - b_y \\ \tilde{\omega}_z^b - b_z \end{bmatrix} \times \begin{bmatrix} v_x^b \\ v_y^b \\ v_z^b \end{bmatrix} - g \begin{bmatrix} -\sin \theta \\ \sin \phi \cos \theta \\ \cos \phi \cos \theta \end{bmatrix} \tag{59.24}$$

Accelerometer measurement could adequately help to estimate the gyro-derived pitch and roll as well as the x- and y-axis gyro biases. However, according to [12], they provide no observability of the heading and heading-gyro bias. A feasible method for estimating the heading and heading-gyro bias is to use the GNSS-derived heading, when GNSS is available, which can be given as

$$\psi^{gnss} = \arctan\left(\frac{v^{gnss, e}}{v^{gnss, n}}\right) \tag{59.25}$$

in which $v^{gnss, e}$ and $v^{gnss, n}$ are respectively the east and north GNSS-derived velocities. Then, at this time the measurement model can be augmented into

$$\begin{bmatrix} f_x^b \\ f_y^b \\ f_z^b \\ \psi^{gnss} \end{bmatrix} = \begin{bmatrix} \tilde{\omega}_x^b - b_x \\ \tilde{\omega}_y^b - b_y \\ \tilde{\omega}_z^b - b_z \end{bmatrix} \times \begin{bmatrix} v_x^b \\ v_y^b \\ v_z^b \end{bmatrix} - g \begin{bmatrix} -\sin \theta \\ \sin \phi \cos \theta \\ \cos \phi \cos \theta \end{bmatrix} \tag{59.26}$$

As can be seen from Eq. (59.25) that the accuracy of ψ^{gnss} is totally depends on the accuracy of the GNSS-derived velocity. According to [5], typically only when the GNSS-derived velocity is larger than 1 m/s^2 , it can then provide reliable GNSS-derived heading ψ^{gnss} . The overview of the EKF flow-chart for optimal attitude estimation and be seen in the following Fig. 59.1.

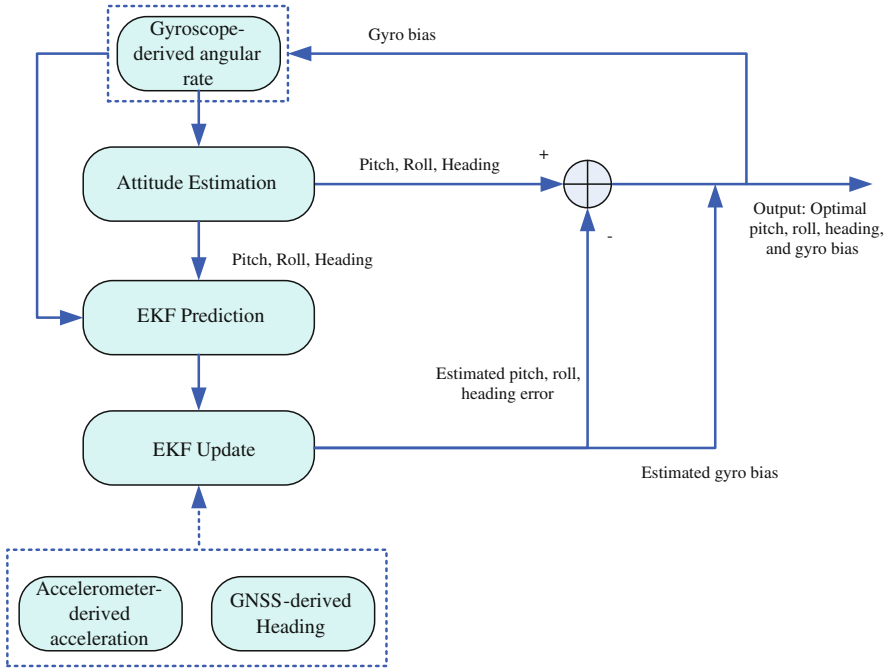


Fig. 59.1 Overview of the EKF flow-chart for optimal attitude estimation

59.4 Experiments and Results

A MEMS IMU (Samsung Note 2 smart-phone including GPS receiver) was used to conduct the experiment to evaluate the performance of the proposed PNS approach. The smart-phone was kept levelly in the user’s hand during the experiment. The x-axis of the sensor-body coordinate is along the smart-phone’s forward direction, y-axis is along the right, and z is orthogonal to the x-y plane and along the down direction. For detecting an effective pair of peak and valley during one step, the time difference between the peak and the valley shall exceed 100 ms (the predefined threshold). 100 ms time difference implies that the pedestrian will not have a step rate greater than 5 steps per second. The stride length coefficient K for Eq. (59.2) was set as 0.2431, which was calibrated on-line by GNSS data.

The experiment was conducted for 8 min on a standard football field (half size) within the University of Calgary’s campus area. The experiment was carried out along a rectangle-shaped line. The side of the rectangle was 61 m in length and 50 m in width, resulting in a total walking distance of 222 m. The sides of the rectangle were either parallel or vertical to the local true north and east. The Ground truth, walking direction, and start and end points (which are the same) can be seen in Fig. 59.2.

Fig. 59.2 Ground truth of the experiment

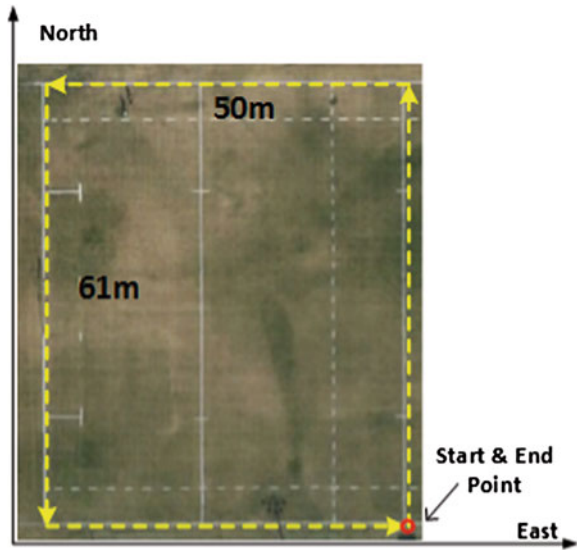


Figure 59.3 shows that the step detection method given in this paper could detect the steps successfully. It is noteworthy that there is no red circle in the end because the user stood still during that time.

Four different methods are used to estimate the heading and navigation results of the PDR-based PNS. For method 1, heading is derived directly through the integration of z-axis gyro in body frame; for method 2, heading is derived through the integration of z-axis gyro in navigation frame, pitch and roll are calculated from x-axis and y-axis gyros in body frame; for method 3, heading is derived through the integration of z-axis gyro in navigation frame, pitch and roll errors are compensated by the accelerometer output; for method 4, heading is derived through the

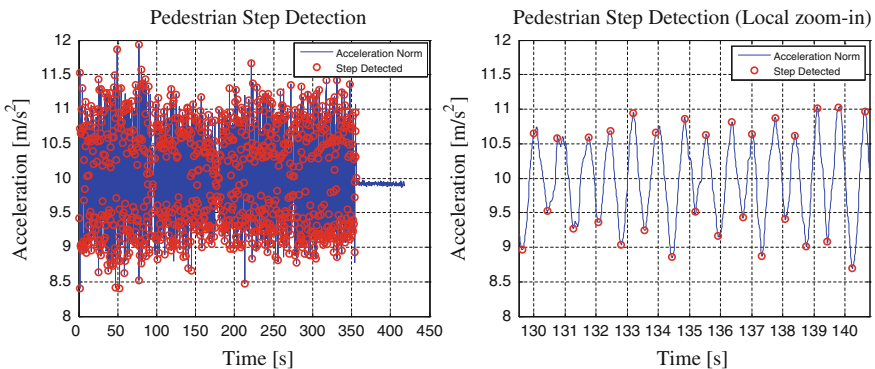


Fig. 59.3 Step detection result and its partial Local zoom-in view (130–140 s)

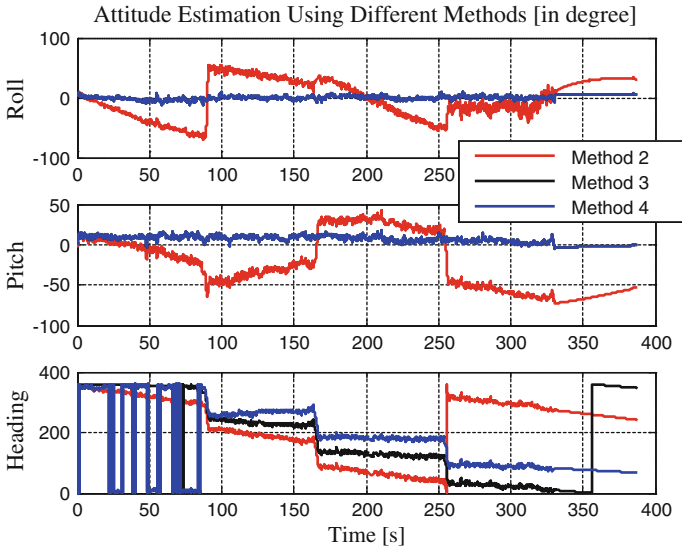


Fig. 59.4 Results of head estimation using different methods

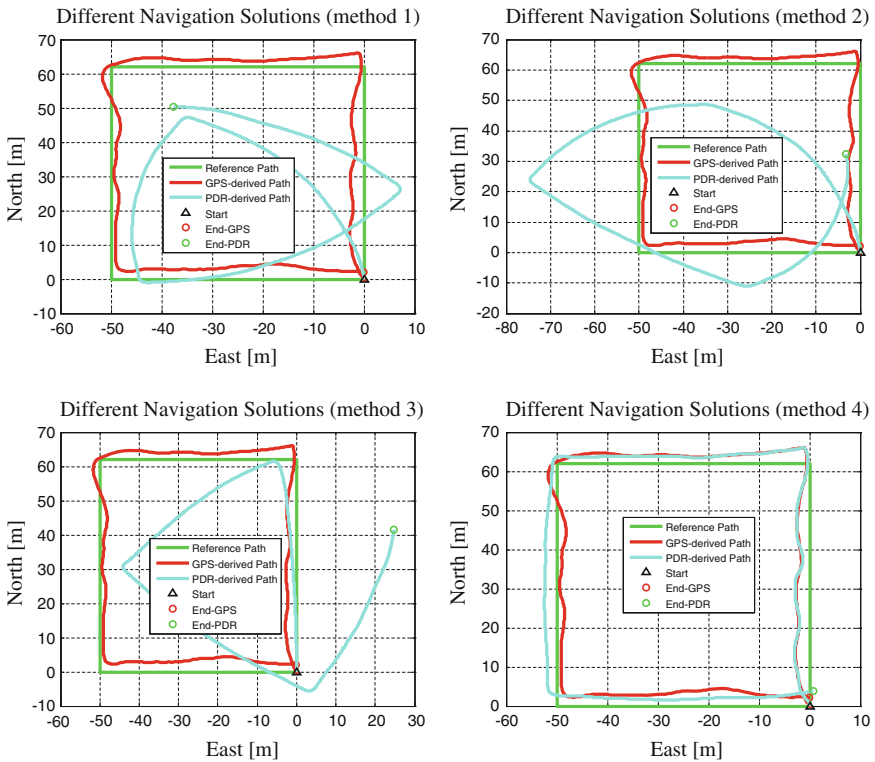


Fig. 59.5 Different navigation results

integration of z-axis gyro in navigation frame; pitch and roll errors are compensated by the accelerometer output, and during the initial 2 min (1/4 of total time) the GPS is available, after that the GPS is simulated unavailable.

It can be seen from Fig. 59.4 that the pitch and roll errors can be well mitigated by methods 2 and 3, which means the proposed accelerometer-aided method could well bound the gyro-derived pitch and roll errors as well as estimate the biases of the horizontal gyros. Moreover, correct heading information can be estimated through using method 4, which means the bias of the heading-gyro could be estimated when GNSS is available. The corresponding positioning results by using methods 1, 2, 3, and 4 can be seen respectively in Fig. 59.5. It is depicted that method 1-derived navigation result is the worst, in which the heading is derived directly through the integration of z-axis gyro in body frame. And the optimal navigation result can be obtained through method 4 which is presented in this paper, with the maximum position error is only about 5 m.

59.5 Conclusions

As can be seen from the above experimental results, the proposed attitude estimation method using accelerometer and GNSS for PDR in the paper can provide accurate and continuous navigation results for pedestrian navigation usage. Although the proposed PDR algorithm is implemented on a consumer smart-phone, they are also applicable on the platform of other smart-devices. The future work will be concentrated on the combination of indoor Wi-Fi positioning technique with the proposed PDR prototype introduced in this paper.

Acknowledgments The first author Haiyu Lan is sponsored by both Dr. Naser El-Sheimy and the China Scholarship Council (CSC) for his PhD program study at the University of Calgary, Calgary, Canada.

References

1. Jirawimut R, Ptasincki P, Garaj V, Cecelja F, Balachandran W (2003) A method for dead reckoning parameter correction in pedestrian navigation system. *IEEE Trans Instrum Meas* 52:209–215
2. Jimenez A, Seco F, Prieto C, Guevara J (2009) A comparison of pedestrian dead-reckoning algorithms using a low-cost MEMS IMU. In: *Proceedings of the IEEE international symposium on intelligent signal processing*
3. Saeedi S, Moussa A, El-Sheimy N (2014) Context-aware personal navigation using embedded sensor fusion in smartphones. *Sensors* 14:5742–5767
4. Noureldin A, Karamat TB, Georgy J (2013) *Fundamentals of inertial navigation, satellite-based positioning and their integration*. Springer, Berlin
5. Zhuang Y, Chang HW, El-Sheimy N (2013) A MEMS multi-sensors system for pedestrian navigation. In: *China satellite navigation conference (CSNC) 2013 proceedings*, pp 651–660

6. Weinberg H (2002) Using the ADXL202 in pedometer and personal navigation applications. Analog Devices AN-602 application note
7. Gabaglio V, Ladetto Q, Merminod B (2001) Kalman Filter Approach for augmented GPS pedestrian navigation. GNSS, Sevilla
8. Shin E-H (2005) Estimation techniques for low-cost inertial navigation. UCGE report, 20219
9. Suh YS (2010) Orientation estimation using a quaternion-based indirect Kalman filter with adaptive estimation of external acceleration. *IEEE Trans Instrum Meas* 59:3296–3305
10. Bachmann ER, Yun X, Brumfield A (2007) Limitations of attitude estimation algorithms for inertial/magnetic sensor modules. *IEEE Robot Autom Mag* 14:76–87
11. El-Sheimy N (2003) Inertial techniques and INS/DGPS integration. Engo 623-Course notes, pp 170–182
12. Barton JD (2012) Fundamentals of small unmanned aircraft flight. Johns Hopkins APL Tech Dig 31:132–149

Chapter 60

Moving Target Autonomous Positioning Based on Vision for UAV

Long Zhao and Pengfeng Chen

Abstract In this paper, we propose a systematic framework for the ground moving target autonomous positioning with a Unmanned Aerial Vehicle (UAV). The proposed framework consists of the ground moving target positioning design and algorithm. Therein the camera mounted on the UAV functions as a mission sensor targeted for the ground moving target positioning. In this paper, We design an algorithm for the vision based mission system by using the navigation information outputted by Attitude and Heading Reference System (AHRS) and GNSS/Strapdown Inertial Navigation System (SINS) integrated navigation system, including motion detection and compensation, target detection and tracking, as well as target positioning. The experiments of the ground moving target autonomous positioning have been carried out. The test results show that the system can provide the positioning information of ground moving target for UAV.

Keywords Systematic framework · Moving target positioning · Target detection · Target tracking

60.1 Introduction

Unmanned Aerial Vehicles (UAVs) are low-risk and cost effective in comparison with human operators. UAVs have become an integral solution for civilian applications and modern military such as remote sensing, aerial surveying, post-disaster assessment, pipeline and power line inspection, and ground target positioning [1–3]. A report in 2011 by the Teal Group, which specializes in the aerospace market

L. Zhao (✉)
Science and Technology on Aircraft Control Laboratory, Beihang University,
Beijing, China
e-mail: buaa_dnc@buaa.edu.cn

P. Chen
Digital Navigation Center, Beihang University, Beijing, China

analysis, forecasts that UAV expenditures worldwide will double within the next 10 years to reach \$11.3 billion US dollars annually in 2020 [4]. By 2018, analysts expect that over 15,000 UAVs could be operating within United States Airspace [5].

Due to the payload, size, and price limitations, low-cost and light-weight Strapdown Inertial Navigation System (SINS) based on micro-electro-mechanical systems (MEMS) are widely applied to the navigation of medium- or small-sized UAVs. Low-cost INS is subject to large measurement biases and high measurement noises. Hence, pure SINS drifts rapidly. In practice, SINS usually is aided by the Global Navigation Satellite System (GNSS) to realize drift-free state estimations. But GNSS signals may be lost due to various reasons such as environment occlusion or intentional jamming. If the GNSS signal for some reason becomes unavailable or corrupted, the state estimation solution provided by the SINS alone drifts in time and will be unusable after a few seconds. For this reason, plenty of research has been dedicated to finding alternative navigation schemes to cope with short and long term GNSS outages. The research community is making a great effort to solve this problem in different ways. One potential solution is a vision navigation system using a suitable camera, for almost every UAV already has a video camera as a standard sensor in its payload package. A comprehensive of various sensors of the payload package such as Inertial Measurement Units (including gyroscopes and accelerometers), magnetometer, barometer, airspeed meter, GNSS, and camera, enables both the autonomous navigation and the mission execution.

Although UAV is more and more widely applied in the modern military and civilian applications, there are several problems which have to be resolved before UAV is introduced in practical applications. One of them is the navigation system and it's integrity, principle and implementation of the navigation system can be found in [6].

After guaranteeing the stability of UAV control during flight, it is necessary to accomplish some specific missions. One of them is the detection, tracking and localization of a moving target for the aerial surveillance. And as a result, ground target tracking as an important application of UAVs has been widely studied by researchers [7–9]. Despite the great progress has been made, moving ground target tracking and positioning is still an important and challenging topic in practical applications of UAVs. In order to achieve this goal, we divide this problem into three sub problems, i.e. motion detection and compensation, moving target detection and tracking, as well as moving target positioning. The motion of the camera due to the motion of the UAV increases the background activity and the tasks of target detection, tracking and positioning become even more daunting. In order to extract the foreground target areas in this case, the global motion between two consecutive frames must be estimated and compensated for a stable video [10].

The paper is organized as follows. The System framework is presented in Sect. 60.2. Section 60.3 focuses on the ground target tracking and positioning system. Finally, we draw some conclusions and shed light on future work in Sect. 60.4.

60.2 Systematic Framework and Problem Statements

The camera mounted on a UAV can be used as a navigation sensor for navigation and positioning, or it can be used as a mission sensor for some specific missions such as target detection, target tracking, target positioning, remote sensing survey, etc.

The work presented in this paper originates from a research project on UAV applications in the Digital Navigation Center (DNC) at Beihang University. The primary goal of the project lies in the development of an integrated hardware/software UAV platform for ground moving target detection, tracking and positioning. The system framework of the ground moving target tracking and positioning system, as shown in Fig. 60.1, depicts an aircraft equipped with a gimbaled optical camera pointing to the ground moving target, where the geographic data are stored in Secure Digital (SD) card. The system can fulfill the target positioning in 3D space after the target detection and tracking. In order to achieve this goal, the transformation between coordinate systems is a key factor for the target positioning in 3D space because different sensors return data with respect to different coordinate frames, such as Earth coordinate frame (e frame), Geographic coordinate frame (g frame), Navigation coordinate frame (n frame), Body coordinate frame (b frame), and Camera coordinate frame (c frame). The definition of these coordinate systems can be found in [6].

According to the definition of the coordinate frames, the problem of ground moving target tracking and positioning based on autonomous navigation systems for UAV is formulated as follows

$$\mathbf{R}_g^t = \mathbf{R}_g^b + \mathbf{R}_b^c + \mathbf{R}_c^t \tag{60.1}$$

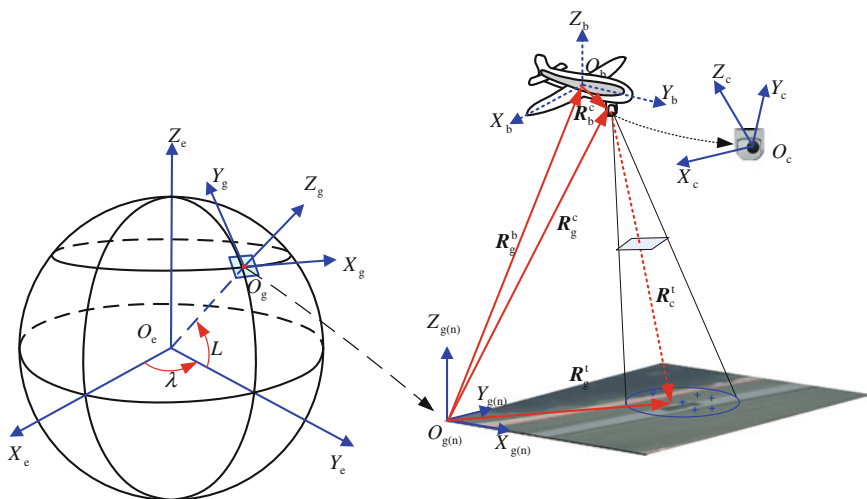


Fig. 60.1 The system framework of ground moving target tracking and positioning

where R_g^t , R_g^b , R_b^c and R_c^t denote the distance vectors from g frame to target point, from g frame to b frame, from b frame to c frame, from c frame to target point, respectively. These parameters calculation process can be found in [6].

60.3 The Target Detection, Tracking and Positioning

60.3.1 Motion Detection and Compensation

The motion of the camera due to the motion of the UAV results in that the background is moving. In contrast to a static background, detecting, tracking and positioning target in a scene with a moving background is much more complex. In order to extract the foreground target areas in this case, the global motion between two consecutive frames must be estimated and compensated.

The motion detection and compensation algorithm involves two stage. The previous and current frames are processed. The dominant motion between them is estimated and is assumed to be the background motion. This motion is then used to align the previous frame to the current frame in order to detect a motion of the current frame. The second stage performs a comparison between the prediction and the current frame. If the motion estimation adequately describes the true background motion, then any unmatching regions is marked as having independent motions.

The global motion model (eight parameters) [11], namely the combination of the affine transformation and projective warp, is described as follows

$$u' = \frac{a_1u + a_2v + a_3}{a_7u + a_8v + 1} \quad (60.2)$$

$$v' = \frac{a_4u + a_5v + a_6}{a_7u + a_8v + 1} \quad (60.3)$$

where (u, v) and (u', v') denote the feature point and its correspondence point in two consecutive frames, and a_i ($i = 1, 2, \dots, 8$) denotes motion parameters.

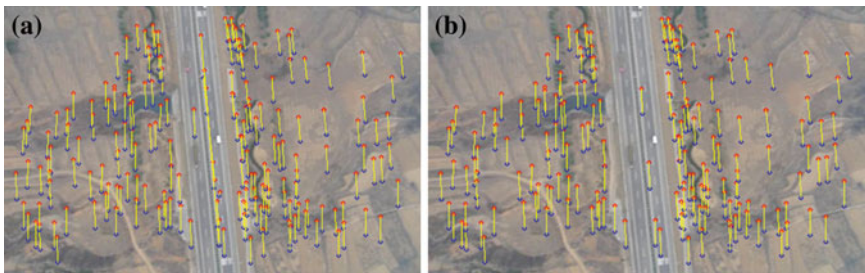


Fig. 60.2 The feature detection and registration, **a** the set of features obtained with the SURF algorithm; **b** the outlier feature set has been identified and removed using the RANSAC algorithm

The motion parameters are estimated based on the SURF matching [12] and RANSAC [13] algorithms. The SURF algorithm is used to process the previous and current frames to extract feature points and match them. The RANSAC algorithm is used to discard outliers so as to ensure that the estimation of the motion parameters is not affected by wrong feature correspondences, as illustrated in Fig. 60.2.

The motion parameters can be estimated using the direct linear transformation with a minimum number of four feature points (the features must be noncollinear). In practice, the motion parameters are estimated using a larger number of

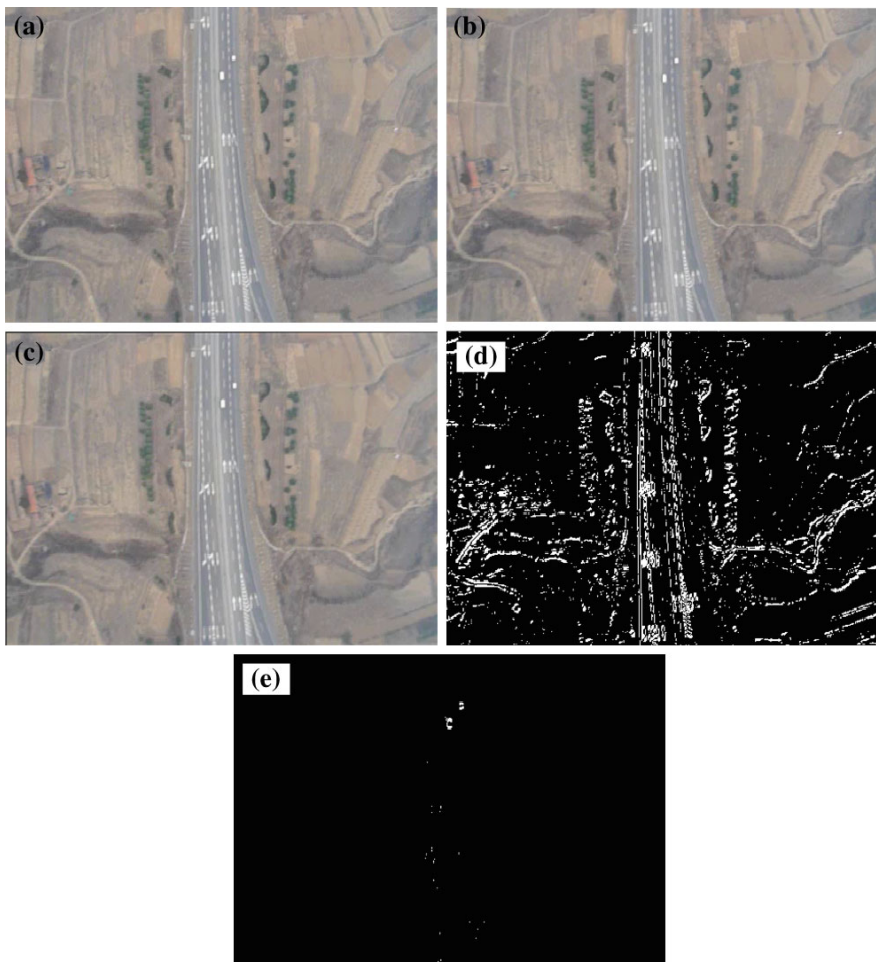


Fig. 60.3 The results of the motion detection and compensation, **a** the previous frame, **b** the current frame, **c** the motion compensation result, **d** the difference binary image between (a) and (b) before motion compensation, and **e** the difference binary image between (b) and (c) after motion compensation

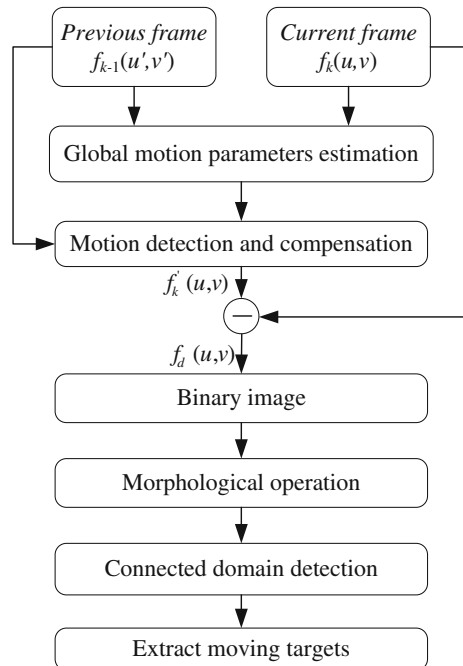
corresponding feature points. Once the motion parameters are obtained, they can be used to reconstruct background pixels in the interpolated frames by using the bidirectional interpolation [14].

An experiment of the motion detection and compensation is implemented by using practical flying video data from the DNC, in which the image resolution is 480×320 pixels. The experimental results are shown in Fig. 60.3.

60.3.2 Target Detection and Tracking

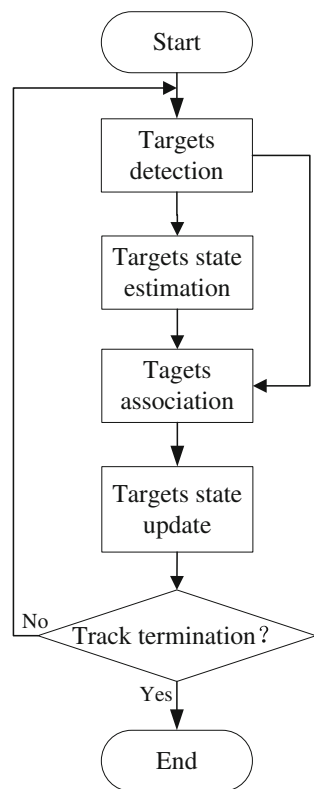
To extract the moving targets from a video frame, background subtraction and frame difference are the most widely used approaches. When the video camera is stationary, the background scene does not change, so it is convenient to construct a background model and extract the moving targets by using the background subtraction. When the camera exhibits motion, however, in order to extract the moving target in this scene where the background changes with each new frame, the global motion between two consecutive frames must be estimated and compensated, so hence convenient to extract the foreground target areas by using the frame difference, and then the moving targets can be detected by the morphological operation and connected domain detection; the structure is shown in Fig. 60.4.

Fig. 60.4 The structure of moving target detection



Once a moving target is detected, the track initialization event can be implemented, such that the moving target can be continuously tracked by the tracking algorithm in the living period of a track (this period starts from its initialization to its termination [15]). The termination of a track occurs when a target can no longer be detected, because it leaves the field of view, it stops and becomes static, or it can no longer be distinguished from the background. Detected objects are not confirmed to be true moving targets until they have been consistently tracked for a period of time before their target tracks are initialized. We create a dynamic list of potential tracks using all detected objects. Associations will be established between objects detected in a new image frame and potential tracking targets. When a potential target is tracked in several continuous frames, it is recognized as a true moving target and a track will be initialized. During target tracking, the target tracking algorithm, i.e., a particle filter based on integral channel features is used to continuously track the target [16]. With real-world video sequences, it is reasonable to take target splitting and merging into account, or detect multiple targets. In order to deal with these issues, an improved nearest neighbor data association algorithm [17] is developed for the multiple targets tracking association task; its structure is in Fig. 60.5.

Fig. 60.5 The flow diagram of targets tracking



An experiment of the motion detection and tracking is implemented by using a practical flight video data from the DNC with the resolution of 480×320 pixels. The experimental results are shown in Fig. 60.6.

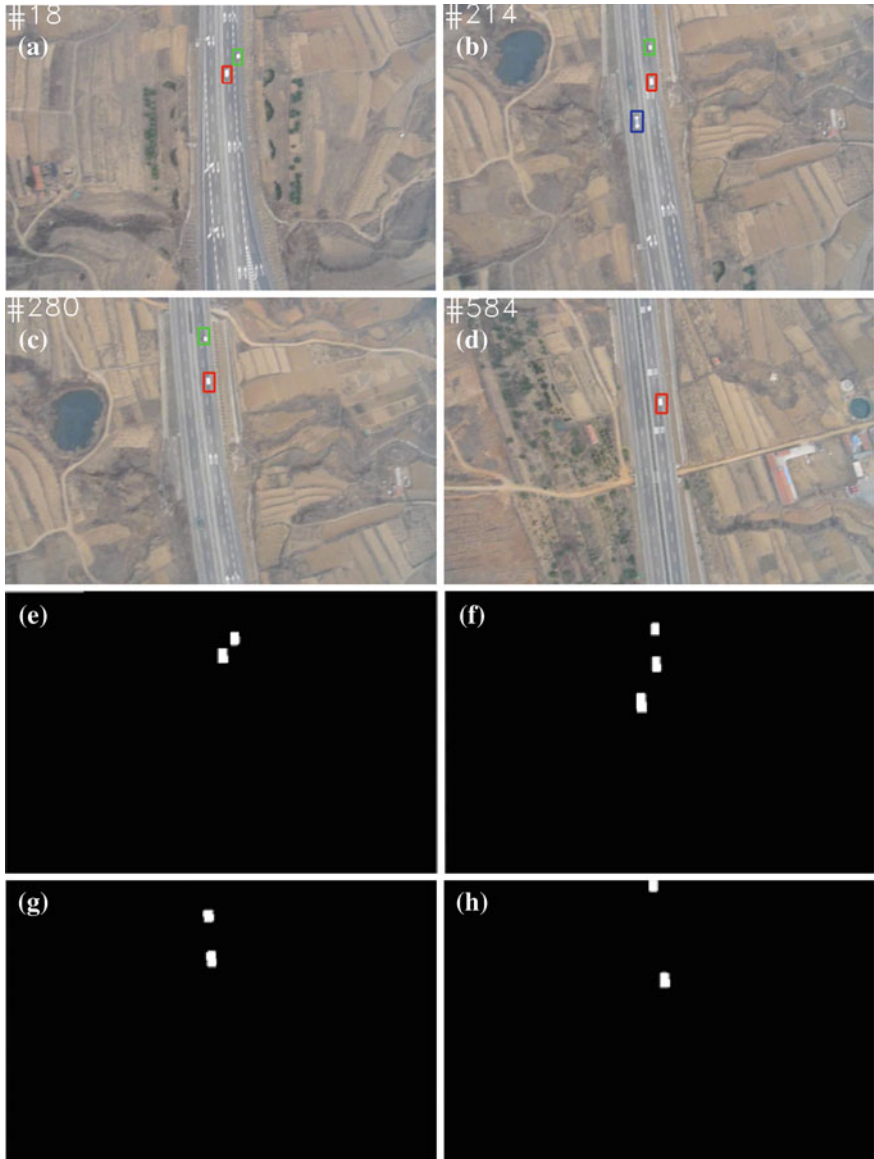


Fig. 60.6 The results of the targets detection and tracking, **a** frame 18, **b** frame 214, **c** frame 280, **d** frame 584, **e** binary image of targets detection in **(a)**, **f** binary image of targets detection in **(b)**, **g** binary image of targets detection in **(c)** and **h** binary image of targets detection in **(d)**

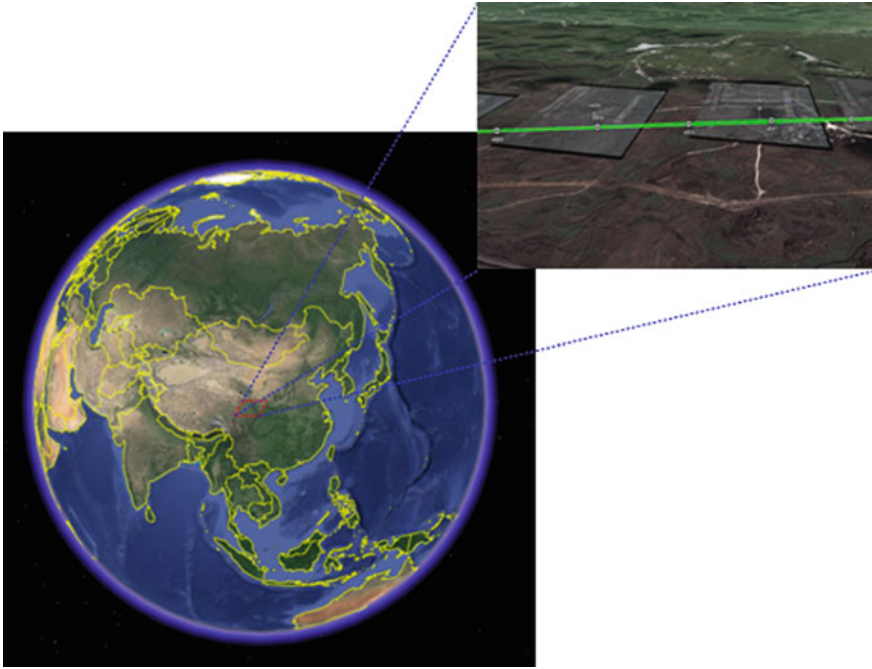


Fig. 60.7 The results displaying the UAV's track and image from the onboard camera

60.3.3 Target Positioning

After targets are continuously and reliably being tracked, any interested target can be selected by a system operator through a joystick or mouse that steers the onboard camera. Once the interested target is selected, the UAV and the camera automatically lock onto the target and continuously track it, as illustrated in Fig. 60.1. Then the system can provide the estimate of its position according to Eq. (60.1).

The system continuously computes the position of the interested target, and the UAV's track can be displayed on the Digital Earth (DE) at the ground station. In addition, the image containing the interested target from the onboard camera can be displayed on the DE at the ground station through the image mosaic technology; the diagrammatic sketch is shown in Fig. 60.7.

60.4 Conclusion and Future Work

In this paper, the comprehensive design and implementation of the ground moving target positioning for UAV was presented, including the system framework, motion detection and compensation, target detection and tracking, as well as target

positioning. The flight test results of target detection and tracking were reported. Although the development of the target positioning based on a camera functioning as the navigation sensor and mission sensor is generally challenging, we believe that the developed avionics and mission system can be widely applied to UAVs and extended to other unmanned vehicles or robots in future.

Acknowledgments This project is supported by the National Natural Science Foundation of China (Grant No. 41274038), the key program of the National Natural Science Foundation of China (Grant No. 61039003), the Aeronautical Science Foundation of China (Grant No. 2013ZC51027), and the Fundamental Research Funds for the Central Universities.

References

1. Cox TH, Nagy CJ, Skoog MA, Somers IA (2006) Civil UAV capability assessment. The National Aeronautics and Space Administration, pp 1–10
2. Wachter Z (2013) A cost effective motion platform for performance testing of MEMS-based attitude and heading reference systems. *J Intell Robot Syst* 70(411):1–9
3. Rodríguez-Canosa GR, Thomas S, Cerro J, Barrientos A, MacDonald B (2012) A real-time method to detect and track moving objects (DATMO) from unmanned aerial vehicles (UAVs) using a single camera. *Remote Sens* 4(4):1090–1111
4. Zaloga SJ, Rockwell D, Finnegan P (2011) World unmanned aerial vehicle systems market profile and forecast. Teal Group Corporation, pp 1–15
5. Schneiderman R (2012) Unmanned drones are flying high in the military/aerospace sector. *IEEE Signal Process Mag* 29(1):8–11
6. Zhao L, Wang D, Huang BQ, Xie LH (2015) Distributed filtering-based autonomous navigation system of UAV. *Unmanned Syst* 3(1):1–18
7. Dobrokhodov VN, Kaminer II, Jones KD, Ghabcheloo R (2006) Vision-based tracking and motion estimation for moving targets using small UAVs. In: The 2006 American control conference, pp 1428–1433
8. Gomez-Balderas JE, Flores G, García Carrillo LR, Lozano R (2012) Tracking a ground moving target with a quadrotor using switching control. In: International conference on unmanned aircraft systems, pp 1–10
9. Zhu SQ, Wang DW, Low CB (2013) Ground target tracking using UAV with input constraints. *J Intell Robot Syst* 69(1–4):417–429
10. Knowles MJ, Spann M (2005) Motion compensated background filtering for real time tracking in moving background sequences. In: IEEE international conference on systems, pp 1059–1065
11. Rawat P, Singhai J (2011) Review of motion estimation and video stabilization techniques for hand held mobile video. *Int J Signal Image Process* 2(2):159–168
12. Bay H, Ess A, Tuytelaars T et al (2008) Speeded-up robust features (SURF). *Int J Comput Vis Image Underst* 110(3):346–359
13. Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun Assoc Comput Mach* 24(6):381–395
14. Liu S, Yan ZD, Kim JW, Kuo CJ (2003) Global/local motion-compensated frame interpolation for low bitrate video. In: SPIE image and video communications and processing, pp 223–234

15. Mao HW, Yang CH, Abousleman GP, Si J (2010) Automated multiple target detection and tracking in UAV videos. In: SPIE airborne intelligence, surveillance, reconnaissance (ISR) systems and applications, vol VII, 77680J1-9
16. Zhang H, Zhao L (2013) Integral channel features for particle filter based object tracking. In: Fifth international conference on intelligent human-machine systems and cybernetics, pp 190–193
17. Banitalebi B, Amiri H (2008) An improved nearest neighbor data association method for underwater multi-target tracking. In: IEEE workshop & exhibition on new trends for environmental monitoring using passive systems, pp 1–4

Chapter 61

An Orbit Determination Algorithm for Spacecrafts Navigated by a Single X-Ray Pulsar

Rong Jiao, Luping Xu, Hua Zhang and Yan Ma

Abstract Commonly, X-ray Pulsar Navigation should observe three pulsars to determine the position of the detector. It leads to at least three detectors should be equipped. In this paper, an initial orbit determination algorithm using one detector to observe a single X-ray pulsar is presented in this paper. The algorithm combines the two-body motion equations and least squares iterative method to obtain initial orbits of spacecrafts, in which the phase comparison method of weighted FFT is used to process simulation signal of an X-ray pulsar. Then, the feasibility of the orbit determination algorithm and the phase comparison method are verified on Matlab environment. As expected, the results show that the orbit determination algorithm is feasible and has a good accuracy. As well, it can be widely used on earth and other planets' orbits.

Keywords Orbit determination algorithm · Single X-ray pulsar · Two-body motion equations · Least squares method

61.1 Introduction

In 1981, the NASA first described the idea that applying the pulsar X-ray source to the spacecraft navigation, which provided a new thought for navigation and positioning technology [1]. As a natural celestial body, the pulsar can provide a good service for the space navigation and positioning technology and its moving characteristics are not easily destroyed by the external force and interference. The parameters of the pulsar in a certain period of time are fixed, so it can realize the function of spacecraft autonomous navigation, reduce the stress of ground stations and have a good application prospect [2, 3]. However, X-ray detector is expensive and can only observe a pulsar at some particular time.

R. Jiao (✉) · L. Xu · H. Zhang · Y. Ma
Xidian University, Xi'an Xifeng Road 266, Xi'an 710126, China
e-mail: zhanghua@mail.xidian.edu.cn

Considering the immature navigation technology of the current X-ray pulsar and the limitation of laboratory fees, it is difficult for us to make one spacecraft carry multiple X-ray detectors [4–6]. The traditional orbit determination method depends on a variety of optical observation data, which cannot work without the support of the ground station. Meanwhile, different observation devices may cause certain observation error. In the determination of initial orbit spacecraft, some methods should be used to correct the observation error, therefore, they increases the workload of spacecraft orbit calculation and decreases the accuracy of the initial orbit. The method in this paper don't rely on optics observation data, we can get the spacecraft initial orbit parameters using the pulsar arrival time parameter. Compared with the original initial orbit determination method, this method need less observation and has a smaller observation error; This method is not rely on the ground service, so it can be used in deep space and space navigation research; Single pulsar orbit determination only need one pulsar, and it only needs an X-ray detector on the spacecraft, which saves the cost and is easy to implement.

61.2 Least Square Method to Determine the Initial Orbit

The initial orbit determination of X-ray pulsar single satellite by the Least Square Method is based on the two-body theory of spacecrafts and the Least Square Method. The detailed process is: according to the signal models of X-ray pulsar and the interrelated knowledge of the Gauss fitting about its profiles, a simulated signal of the X-ray pulsar is generated; an analysis of the two cases is made to calculate the initial state vector of spacecrafts, as the initial value of the estimated orbit of three pulsars at the same place or at two or three different places. With the initial state vector of spacecrafts on the estimated orbit, the two-body equation of motion on elliptical orbit and Newton iterative method, the estimated orbit is produced. Likewise, choosing a initial state vector of a reference orbit can make a reference orbit; According to the two-body motion of spacecrafts in the orbit coordinate system, the mathematical model of the two-body motion and using the Least Square Method to make the estimated orbit to get close to the reference orbit, we can get the initial orbit of spacecrafts.

61.2.1 Calculating the Initial Magnitude of Spacecrafts on the Estimated Orbit

Literature [7] provides a fast simulation method for the fast X-ray signal, and simulation data of three pulsars can be got by this method. Then some phase value of the photon sequence of each pulsar is obtained by using a 1024-point method on 900 s cumulative simulation data of three pulsars respectively.

We observe three pulsars one by one, that is to say, one detector can only observe one pulsar at a moment, which is called orbit determination for spacecrafts navigated by a single X-ray pulsar. Difference of the place of observation, the initial orbit determination of the single star of the X-ray pulsar can be divided into two kinds of situations: observing the three pulsars in the same place or at two or three different place.

The first case: three stars produce initial value of estimates track in the same positions.

When we observe three pulsars at the same place, the positions of spacecraft are the same results. We can consider locating the spacecraft by the geometric method. With the measured phases, we can ensure the accurate position of the spacecraft.

Supposing the pulsar signal emission time of the i pulsar is t_{T_i} , the arrival time to the spacecraft is t_{SC_i} , \mathbf{D}_i and \mathbf{r}_{SC} are respectively the vector of mass center of earth to the i pulsar and the place of the spacecraft. Their relationship follows

$$(t_{SC_i} - t_{T_i}) = 1/c \cdot \hat{\mathbf{n}}_{SC}(\mathbf{D}_i - \mathbf{r}_{SC}) + 1/c \cdot \text{Rel}_i \quad (\hat{\mathbf{n}}_{SC} = (\mathbf{D}_i - \mathbf{r}_{SC})/|\mathbf{D}_i - \mathbf{r}_{SC}|) \quad (61.1)$$

where Rel_i is the i pulsar baseline in direction of the relativistic effect, $\hat{\mathbf{n}}_{SC}$ is the unit vector of the i pulsar to spacecraft direction, c is speed of light.

By Eq. (61.1), the ρ_i (The distance of the pulsar to the spacecraft) can be expressed as

$$\rho_i = c(t_{SC_i} - t_{T_i}) = \hat{\mathbf{n}}_{SC}(\mathbf{D}_i - \mathbf{r}_{SC}) + \text{Rel}_i = |\mathbf{D}_i - \mathbf{r}_{SC}| + \text{Rel}_i \quad (61.2)$$

Distance measurements contain signal timing errors δt_{SC} , distance measurements noise η_{SC_i} , and the pulse signal transmitting time errors δT_i . We can rewrite Eq. (61.2) as

$$\begin{aligned} \rho_i &= c(t_{SC_i} - t_{T_i}) + c\delta t_{SC} + c\delta T_i + \eta_{SC_i} \\ &= |\mathbf{D}_i - \mathbf{r}_{SC}| + \text{RelEff}_i + c\delta t_{SC} + c\delta T_i + |\delta \mathbf{D}_i| + |\delta \mathbf{r}_{SC}| + \delta \text{Rel}_i + \eta_{SC_i} \end{aligned} \quad (61.3)$$

where $\delta \mathbf{D}_i$, δRel_i and $\delta \mathbf{r}_{SC}$ respectively show the pulsar position error, the relativistic effect error and the spacecraft position error.

If the entire cycle phase of the pulsar to spacecraft is represented as N_i , Eq. (61.3) can be expressed as

$$\begin{aligned} \lambda_i \Phi_i &= \lambda_i(\varphi_i + N_i) = \rho_i = c(t_{SC_i} - t_{T_i}) + c\delta t_{SC} + c\delta T_i + \lambda_i \beta_{SC_i} \\ &= |\mathbf{D}_i - \mathbf{r}_{SC}| + \text{Rel}_i + c\delta t_{SC} + c\delta T_i + |\delta \mathbf{D}_i| + |\delta \mathbf{r}_{SC}| + \delta \text{Rel}_i + \lambda_i \beta_{SC_i} \end{aligned} \quad (61.4)$$

where λ_i is the pulse signal wavelength, β_{SC_i} is the phase measurement error.

Because the distance between the pulsar and spacecraft is too far, the visual angle's variation caused by the change of the distance is very small. We cannot determine the pulsar signal transmission time and the phase integer cycle. In addition, we cannot accurately obtain the geometric distance of the pulsar. Thus, we generally use difference method in the practical application.

Single difference refers to using group difference between pulsar time prediction model phase and the spacecraft pulse phase measurements. Differential observation can reduce errors, but it needs more observation time and observation sources. According to Eq. (61.5), position vector $\Delta \mathbf{x}$ which is spacecraft relative to the mass center of earth.

$$\begin{aligned} \Delta \rho_i &= \rho_{E_i} - \rho_{SC_i} = c(t_{E_i} - t_{T_i}) - c(t_{SC_i} - t_{T_i}) \\ &= [|\mathbf{r}_E - \mathbf{D}_i| + \text{Rel}_{E_i}] - [|\mathbf{r}_{SC} - \mathbf{D}_i| + \text{Rel}_{SC_i}] = \hat{\mathbf{n}}_i \cdot \Delta \mathbf{x} \end{aligned} \quad (61.5)$$

where t_{E_i} is the time pulse arriving at the mass center of earth, r_E is the distance between the mass center of earth and the mass center of sun. ρ_{E_i} is the distance between pulsar and the mass center of earth, and ρ_{SC_i} is the distance between pulsar and spacecraft. Rel_{E_i} is the relativistic correction of ρ_{E_i} , Rel_{SC_i} is the relativistic correction of ρ_{SC_i} .

According to Eqs. (61.4) and (61.5), the single difference measurements can be set as

$$\Delta \rho_i = \lambda_i \Delta \Phi_i = \lambda_i (\Delta \varphi_i + \Delta N_i) = \hat{\mathbf{n}}_i \cdot \Delta \mathbf{x} \quad (61.6)$$

where $\Delta \varphi_i$ is the phase between the spacecraft and earth, ΔN_i is phase ambiguity between the spacecraft and earth.

If we observe three pulsars at the same time, a simplified measurement equation is

$$\begin{aligned} [\hat{n}_1 \quad \hat{n}_2 \quad \hat{n}_3]^T \mathbf{x} &= \mathbf{H} \mathbf{x} \\ &= [cP_1(\Delta\phi_1 + \Delta N_1) \quad cP_2(\Delta\phi_2 + \Delta N_2) \quad cP_3(\Delta\phi_3 + \Delta N_3)]^T \end{aligned} \quad (61.7)$$

where \hat{n}_i is the i pulsar direction unit vector, \mathbf{x} is the spacecraft's position vector relative to the vector time model reference point. In this paper, Time model reference point is the mass center of the earth, c is speed of light, P_i is the i Pulsar pulse period; ΔN_i is the single difference phase ambiguity. $\Delta\phi_i$ is a transformed single difference phase observation.

Because the three pulsars are observed in the same location, we think they are observed at the same time. The pulsar distance is very far, Eq. (61.7) is simplified as

$$[\hat{n}_1 \quad \hat{n}_2 \quad \hat{n}_3]^T \mathbf{x} = \mathbf{H} \mathbf{x} = [cP_1\Delta\phi_1 \quad cP_2\Delta\phi_2 \quad cP_3\Delta\phi_3]^T \quad (61.8)$$

Assuming that each pulsar's photon sequence can be divided into N segments, and using Eq. (61.8), we can calculate n positions vector of spacecraft.

Some \mathbf{x}_i corresponding to t_i are known, we can get the formula of spacecraft initial position vector \mathbf{x}_0 and velocity vector $\dot{\mathbf{x}}_0$

$$\mathbf{x}_i = f_i \mathbf{x}_0 + g_i \dot{\mathbf{x}}_0 (i = 1, 2, \dots, n) \quad (61.9)$$

The solution with the least square method is

$$\begin{cases} \mathbf{x}_0 = [\mathbf{g}^2][\mathbf{f}\mathbf{x}] - [\mathbf{f}\mathbf{g}][\mathbf{g}\mathbf{x}]/[\mathbf{f}^2][\mathbf{g}^2] - [\mathbf{f}\mathbf{g}][\mathbf{f}\mathbf{g}] \\ \dot{\mathbf{x}}_0 = [\mathbf{f}^2][\mathbf{g}\mathbf{x}] - [\mathbf{f}\mathbf{g}][\mathbf{f}\mathbf{x}]/[\mathbf{f}^2][\mathbf{g}^2] - [\mathbf{f}\mathbf{g}][\mathbf{f}\mathbf{g}] \end{cases} \quad (61.10)$$

where $[\mathbf{x}] = \sum_{i=1}^n x_i$.

The second case: three stars produce initial value of estimates track in three or two positions.

When the three stars are in the three or two positions, calculating the spacecraft's projection of the pulsar direction in the initial time by using least square method for each pulsar, we can obtain simultaneous equations about the position vector and velocity vector, the solution of the spacecraft orbit position vector and velocity vector can be estimated.

$$\mathbf{x}_i \vec{n}_i = f_i \mathbf{x}_0 \vec{n}_i + g_i \dot{\mathbf{x}}_0 \vec{n}_i = cP_i \phi_i (i = 1, 2, \dots, n) \quad (61.11)$$

where P_i is the period of the i pulsar, c is speed of light, ϕ_i is pulsar photon sequence phase value at t_i moment. The solution with the least square method is

$$\begin{cases} \mathbf{x}_0 = [\mathbf{g}^2][\mathbf{f}\mathbf{x}] - [\mathbf{f}\mathbf{g}][\mathbf{g}\mathbf{x}]/[\mathbf{f}^2][\mathbf{g}^2] - [\mathbf{f}\mathbf{g}][\mathbf{f}\mathbf{g}] \\ \dot{\mathbf{x}}_0 = [\mathbf{f}^2][\mathbf{g}\mathbf{x}] - [\mathbf{f}\mathbf{g}][\mathbf{f}\mathbf{x}]/[\mathbf{f}^2][\mathbf{g}^2] - [\mathbf{f}\mathbf{g}][\mathbf{f}\mathbf{g}] \end{cases} \quad (61.12)$$

where $x = cP\phi$, $[\mathbf{x}] = \sum_{i=1}^n x_i$.

In the above two cases, f , g are calculated according to the approximate circular orbit two-body motion equations. If the initial position vector and velocity vector of the spacecraft are unknown, we considered orbit is an approximate circular orbit. Selecting a reasonable initial value r_0 , f and g can be obtained by this formula

$$\begin{cases} f = \cos\left((t - t_0)/r_0^{3/2}\right) + O(\epsilon_0) + \cos\left(\delta/r_0^{3/2}\right) + O(\epsilon_0) \\ g = r_0^{3/2} \sin\left(\delta/r_0^{3/2}\right) + O(\epsilon_0) \end{cases} \quad (61.13)$$

where $\delta = t - t_0$, $O(\epsilon_0)$ is higher order infinitesimal.

Equation (61.13) can be deformed into

$$\begin{cases} f = 1 - \delta^2/2r_0^3 + O(\delta^4) \\ g = \delta - \delta^3/6r_0^3 + O(\delta^5) \end{cases} \quad (61.14)$$

Equations (61.13) and (61.14) above mentioned, commonly used in practical work, are two kinds of approximation formulas for f, g . If the initial value r_0 gets close to the actual value, the eccentricity e_0 is small. No matter how great the magnitude of δ in Eq. (61.13) is, the calculation accuracy of f, g is good. but when the value of δ in Eq. (61.14) is larger, the calculation accuracy error of f, g is also larger.

61.2.2 The Estimation of Initial Orbits of Spacecrafts

(1) The generation of the estimate and reference orbits

According to initial value of estimate orbits and appropriate initial value of reference orbits, f, g, f' and g' can be worked out through elliptical orbit two-body motion equations. The equations can be represented as

$$\begin{cases} f = 1 - a_0/r_0 \cdot (1 - \cos \Delta E) \\ g = \sqrt{a_0}r_0 \sin \Delta E + a_0(r_0\dot{r}_0)(1 - \cos \Delta E) \\ f' = -\sqrt{a_0}/r_0r \cdot \sin \Delta E \\ g' = 1 - a_0/r \cdot (1 - \cos \Delta E) \end{cases} \quad (61.15)$$

where f' and g' can be gotten by taking the derivative of f and g .

The equation about ΔE can be expressed as

$$\Delta E = \Delta M - r_0\dot{r}_0/\sqrt{a_0} \cdot (1 - \cos \Delta E) + (1 - r_0/a_0) \sin \Delta E \quad (61.16)$$

where, $\Delta M = M - M_0 = n(t - t_0)$, $n = \sqrt{\mu/a^3}$ is the average velocity of spacecrafts, $\mu = GM$ is gravitational constant and a is the semi-major axis of the ellipse.

Because of its similarity to the Kepler's equation, the Eq. (61.16) is called Generalized Kepler equation, which can be solved with the Newton iterative method.

The steps to solve this equation can be described as

$$\text{Step 1 : } \begin{cases} f(\Delta E) = \Delta E - (1 - r_0/a_0) \sin \Delta E + r_0\dot{r}_0/\sqrt{a_0} \cdot (1 - \cos \Delta E) - \Delta M \\ f'(\Delta E) = 1 - (1 - r_0/a_0) \cos \Delta E + r_0\dot{r}_0/\sqrt{a_0} \cdot \sin \Delta E \end{cases} \quad (61.17)$$

Here, ΔE can be obtained by Newton iterative method, there is

$$\Delta E \approx \Delta E_{i+1} = \Delta E_i - f(\Delta E_i)/f'(\Delta E_i) \quad (61.18)$$

Step 2 : choose $\Delta E_0 = \Delta M$ and solve Eq. (61.18) for ΔE_1

$$\Delta E_1 = \Delta M + ((1 - r_0/a_0) \sin \Delta M - r_0 \dot{r}_0 / \sqrt{a_0} \cdot (1 - \cos \Delta M)) / (1 - (1 - r_0/a_0) \cos \Delta M + r_0 \dot{r}_0 / \sqrt{a_0} \cdot \sin \Delta M) \quad (61.19)$$

Then substitute ΔE_1 for ΔE_i in Eq. (61.18) and calculate ΔE_2 . And by this analogy, calculate $\Delta E_3, \Delta E_4, \dots$, until $|E_{i+1} - E_i| < \varepsilon$.

Based on the value of f, g, f' and g' on elliptical orbit and the position vector x_0 and velocity vector \dot{x}_0 , the position and velocity vector of spacecraft corresponding to several times of the estimate orbit and reference orbit can be got by Eq. (61.20).

$$\begin{cases} \dot{x} = fx_0 + g \dot{x}_0 \\ \dot{x} = f'x_0 + g' \dot{x}_0 \end{cases} \quad (61.20)$$

(2) The calculation on the initial position and velocity vector of spacecraft

According to the two-body motion of spacecraft in orbital coordinate system, we can construct a mathematical model of the two-body motion of spacecraft. Based on the theoretical knowledge of Chap. 2, if the matrix A and \tilde{H} are already known, the initial orbit can be estimated by Least Squares Method.

It is assumed that the spacecraft is only affected by the center attraction, without regard to the influence of other inertia force. It is assumed that the surface-based observing station (X_s, Y_s, Z_s) and gravitational constant is known. Thus, the state vector can be represented as

$$X = [x \ y \ z \ u \ v \ w]^T \quad (61.21)$$

Here, (x, y, z) is the position vector of spacecraft at one point, (u, v, w) is the velocity vector of spacecraft at one point. According to the two-body motion equation, we can obtain $\ddot{x} = -\mu x/r^3, \ddot{y} = -\mu y/r^3, \ddot{z} = -\mu z/r^3$.

By Eq. (61.22), we can obtain

$$\begin{aligned} \dot{X} &= [\dot{x} \ \dot{y} \ \dot{z} \ \dot{u} \ \dot{v} \ \dot{w}]^T = [F_1 \ F_2 \ F_3 \ F_4 \ F_5 \ F_6]^T \\ &= [u \ v \ w \ -\mu x/r^3 \ -\mu y/r^3 \ -\mu z/r^3]^T \end{aligned} \quad (61.22)$$

According to Eq. (61.22)

$$A(t) = \frac{\partial F(X, t)}{\partial X} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -\frac{\mu}{r^3} + \frac{3\mu x^2}{r^5} & \frac{3\mu xy}{r^5} & \frac{3\mu xz}{r^5} & 0 & 0 & 0 \\ \frac{3\mu xy}{r^5} & -\frac{\mu}{r^3} + \frac{3\mu y^2}{r^5} & \frac{3\mu yz}{r^5} & 0 & 0 & 0 \\ \frac{3\mu xz}{r^5} & \frac{3\mu yz}{r^5} & -\frac{\mu}{r^3} + \frac{3\mu z^2}{r^5} & 0 & 0 & 0 \end{bmatrix} \quad (61.23)$$

The matrix H is defined as

$$\tilde{H} = \partial\rho/\partial X = [\partial\rho/\partial x \quad \partial\rho/\partial y \quad \partial\rho/\partial z \quad \partial\rho/\partial u \quad \partial\rho/\partial v \quad \partial\rho/\partial z] \quad (61.24)$$

where $\rho = [(x - X_s)^2 + (y - Y_s)^2 + (z - Z_s)^2]^{1/2}$.

According to Eq. (61.24)

$$\tilde{H} = [(x - X_s)/\rho \quad (y - Y_s)/\rho \quad (z - Z_s)/\rho \quad 0 \quad 0 \quad 0] \quad (61.25)$$

Known matrix A, $\Phi(t, t_0)$ can be obtained by the characteristic value and eigenvector. Then we can estimate the state vectors of the spacecraft by Least-squares iteration method mentioned in literature [8].

Estimation process of spacecraft initial state is the process of iterative convergence. The condition of iterative convergence is: the mean square error of estimated orbit observation and the reference orbit observation is less than 1 km.

(3) The initial orbit determination of the spacecraft

Given the position and velocity vector of the spacecraft in the initial state, we can get six orbital elements ($a_0, e_0, i_0, \Omega_0, \omega_0, M_0$) of the elliptic orbits. For eccentricity $e_0=0$, perigee has no definition, similar to ω_0 and M_0 . When eccentricity e_0 get close to 0, it is difficult to get accurate solutions of ω_0 and M_0 . For $i_0 = 0$, ascending node has no definition, and neither do Ω_0 and ω_0 . When Orbital inclination i_0 get close to 0, it is difficult to get accurate solution of Ω_0 and ω_0 . So, in the special cases such as a small eccentricity and small angle, the solution variable is no longer six orbital elements of Kepler elliptic orbit, and the calculation method will change, too. This section will be emphasis on some calculation formulas in general situations.

For $x_0 = (x, y, z)$ and $\dot{x}_0 = (u, v, w)$, we can get

$$\begin{cases} r_0 = (x^2 + y^2 + z^2)^{1/2} \\ v_0^2 = (u^2 + v^2 + w^2) \end{cases} \quad (61.26)$$

According to the dynamic formula

$$v_0^2 = \mu(2/r_0 - 1/a_0) \quad (61.27)$$

Elliptical orbit semi-major axis of a spacecraft is

$$a_0 = (2/r_0 - v_0^2/\mu)^{-1} = \mu r_0 / (2\mu - r_0 v_0^2) \quad (61.28)$$

For $e \neq 0$, by the elliptical motion relationship, we obtain

$$\begin{cases} e_0 \cos E_0 = 1 - r_0/a_0 \\ e_0 \sin E_0 = r_0 \dot{r}_0 / \sqrt{a_0} \end{cases} \quad (61.29)$$

Here

$$r_0 \dot{r}_0 = x_0 \cdot \dot{x}_0 = xu + yv + zw \quad (61.30)$$

According to Eq. (61.29), there is

$$e_0 = \sqrt{(1 - r_0/a_0)^2 + (r_0 \dot{r}_0 / \sqrt{a_0})^2} \quad (61.31)$$

$$\tan E_0 = (r_0 \dot{r}_0 / \sqrt{a_0}) / (1 - r_0/a_0) \quad (61.32)$$

With E_0 and e_0 , we can get M_0 , it is

$$M_0 = E_0 - e_0 \sin E_0 \quad (61.33)$$

For $i \neq 0^\circ$ and $i \neq 180^\circ$, we can get the right ascension of ascending node formula

$$\tan \Omega_0 = (yw - zv) / (-uz + xw) \quad (61.34)$$

The orbit inclination is

$$\cos i_0 = (xv - yu) / \sqrt{a_0(1 - e_0^2)} \quad (61.35)$$

By the fundamental formulas of elliptic orbit, the true anomaly f_0 is

$$\tan f_0 = \sqrt{1 - e_0^2} \sin E_0 / (\cos E_0 - e_0) \quad (61.36)$$

The true latitude u_0 is

$$\tan u_0 = z / (x \cos \Omega_0 + y \sin \Omega_0) \cdot \sin i_0 \quad (61.37)$$

With the true latitude, we can get that the argument of perigee is

$$\omega_0 = u_0 - f_0 \tag{61.38}$$

61.3 Comparisons and Analyses of Simulation Results

The real orbit parameters setting refer to the orbit of GPS(GPS BIIA-10) during the simulation whose parameters are shown in Table 61.1.

61.3.1 Simulation Results of Estimating the Initial Value of the Orbit

At different observation time, the initial states of the orbit are different. Assuming an approximate and reasonable initial value of circular orbit is 3.8999×10^8 m. Observation time starts from 20 s and gradually increases by 5 s until 105 s. Estimation results of the height and speed for initial state of the orbit are shown in Tables 61.2 and 61.3.

Table 61.1 Orbit parameters setting during the simulation

Orbit name	Semi-major axis (km)	Right ascension of ascending node	Orbit eccentricity	Argument of perigee	Orbital inclination
GPS BIIA10	4164.3	224.67°	0.0116	338.24°	54.39°
MEGSAT-1	2254.5	341.25°	0.0049	281.70	64.56°

Table 61.2 Estimation results of the height and speed for initial state of the orbit

	25 s	35 s	45 s	55 s	65 s
Altitude (m)	3.8849×10^8	3.8849×10^8	3.8849×10^8	3.8849×10^8	3.8849×10^8
Velocity (m/s)	1.0390×10^4	7.4218×10^3	5.7725×10^3	4.7229×10^3	3.9963×10^3

Table 61.3 Estimation results of the height and speed for initial state of the orbit

Time	25 s	35 s	45 s	55 s	65 s
Data					
Altitude (m)	3.8849×10^8	3.8849×10^8	3.8849×10^8	3.8849×10^8	3.8849×10^8
Velocity (m/s)	2.0781×10^3	1.9242×10^3	1.7915×10^3	1.6759×10^3	1.5743×10^3

From the Tables 61.2 and 61.3, it can see that the initial velocity value of estimated orbit is unreasonable when the observation time is less than 35 s. Then, with the observation time becomes longer ranging from 35 s to 105 s, the initial height of estimated orbit remains unchanged and the initial velocity decreases. Meanwhile, the estimated initial height and velocity are reasonable. So the reasonable observation time setting should range from 35 s to 105 s.

61.3.2 Simulation Result of LSM

Assuming that the initial height value between the estimated orbit and the reference orbit differs 10 km, that is, the height value of the reference orbit is 3.9×10^8 m. In the process of iterative for determining the initial spacecraft orbit, the starting observation time lasting for 3 s, increases by 3 s each time, until reach the total length of 30 s. Calculate the reference track and estimate track matrix y value respectively, then compute mean square by calculating the difference of y value, the mean square value can be used as the convergence conditions. Calculation results are shown in Tables 61.4 and 61.5, the first column represents observation time for estimating the initial state of orbit, the first line says the observation time for least square method to determine the initial orbit.

From the Tables 61.4 and 61.5, it can see that when the estimated initial value of orbit at 105 s of the initial value, the error between the reference trajectory and the estimated orbit reaches its minimum value. In this case, as the time observed in the process of iterative becomes longer, the error between the estimated orbit and the reference trajectory becomes greater. In another word, the estimated orbit goes off the reference orbit further and further, and when the time observed in the process of iterative is 3 s, the error between the estimated orbit and the reference trajectory

Table 61.4 Simulation results of observing three pulsars in the same places

	3 s (m)	9 s (m)	15 s (m)	21 s (m)	27 s (m)
60 s	371.4927	371.6545	379.2448	400.8785	412.4463
80 s	194.7183	194.7773	195.5015	197.1876	291.9386
90 s	195.3095	195.1225	194.9779	195.1719	222.8665
105 s	181.8157	181.6162	181.4276	182.5941	181.7273

Table 61.5 Simulation results of observing three pulsars in different places

	3 s (m)	9 s (m)	15 s (m)	21 s (m)	27 s (m)
60 s	371.4877	371.6707	380.4308	400.8802	412.4560
80 s	194.7168	194.7776	195.4965	197.1913	291.9417
90 s	195.3143	195.1257	195.0066	195.1621	222.8742
105 s	181.8228	181.6193	181.4195	182.5988	181.7215

Table 61.6 Six orbital elements for the approximate time of observation on different elliptic orbits

	a_0	e_0	i_0	Ω_0	ω_0	M_0
35 s	7.6201e + 006	72.0540	14.0572	1.0310	5.0302	70.4832
45 s	1.2928e + 007	42.4784	14.0701	1.0310	4.5086	40.9076
55 s	1.9968e + 007	27.5112	14.0868	1.0310	4.0810	25.9404
65 s	2.9075e + 007	18.9075	14.1075	1.0310	3.7130	17.3367
75 s	4.0729e + 007	13.5148	14.1328	1.0310	3.3846	11.9440
85 s	5.5632e + 007	9.9173	14.1634	1.0310	3.0829	8.3465
95 s	7.4832e + 007	7.4026	14.2003	1.0310	2.7990	5.8318
105 s	9.9948e + 007	5.5814	14.2447	1.0310	2.5259	4.0106

reaches its minimum. As a result, the initial value of estimated orbit can be regarded as the spacecraft's initial orbit. With the same estimated orbit initial value and the time in the process of iterative, the error stem from the observation of three pulsars in the same places becomes greater than that in different places. In this case, as seen from the Table 61.5, the error between the estimated orbit and the reference trajectory reaches 134.5958 m. In the process of simulation, a difference in height by 5 km between the estimated orbit and the reference trajectory are taken into consideration. Simulation results show that the error between the reference trajectory and the estimated orbit is the same as the case of a difference in height by 10 km (Table 61.6).

Simulation results indicate that: with the same initial value of estimated orbits, six orbital elements are not affected by the changeable time of observation during iteration. With different initial value of estimated orbits, it is easy to get that the initial heights remain stable with longer time of observation for circular orbit. However, changing initial velocities of the orbit brings about the variation on six orbital elements on elliptic orbits. The semi-major axis of elliptic orbits increases with longer time of observation, and this variety and the order of magnitude go hand in hand. The orbital eccentricity, the argument of perigee and the mean anomaly has a decreasing tendency. The orbital inclination increases slowly. The alternations of the right ascension of ascending node are unlikely ever to happen.

61.4 Conclusions

The position and velocity vector of spacecraft are presented based on the application of the two-body motion equations, spacecraft orbit determination theory and least squares iterative method. The six orbital elements of spacecraft on elliptic orbits can be obtained by means of spacecraft orbit determination. This paper makes an analysis of the error sources of the orbit determination methods of X-ray pulsar single satellite with the help of the pulsar arrival time of observation and two-body

problem. The determination of spacecraft initial orbit is an iterative process, which is based on the initial estimation orbit and the reference orbit as standard. It is the convergence of the iterative conditions that the mean square error of observations between the estimation orbit and the reference orbit is less than 1 km.

References

1. Chester TJ, Butman SA (1981) Navigation using X-ray pulsars NASA. The telecommunications and data acquisition progress report, TDA PR 42–63
2. Bernhardt MG, Becker W, Prinz T, Breithuth FM, Walter U (2011) Autonomous spacecraft navigation based on pulsar timing information. In: 2011 2nd international conference on space technology (ICST), pp 1–4
3. Sheikh SI (2005) The use of variable celestial X-ray sources for spacecraft navigation. University of Maryland, College Park
4. Graven P, Collins J, Sheikh S, Hanson J, Ray P, Wood K (2008) XNAV for deep space navigation. In: 31st annual AAS rocky mountain guidance and control conference, vol 131. pp 349–364
5. Luo N, Xu L, Zhang H, Xie Q (2013) Feasibility analysis for attitude estimation based on pulsar polarization measurement. *J Zhejiang Univ Sci C* 14(6):425–432
6. Zhang H, Xu L, Xie Q (2011) Modeling and Doppler measurement of X-ray pulsar. *Sci China Phys Mech Astron* 54(6):1068–1076
7. Zhang H, Xu L, Song S, Jiao R (2014) A fast method for X-ray pulsar signal simulation. *Acta Astronaut* 98:189–200
8. Lee DJ (2005) Nonlinear bayesian filtering with applications to estimation and navigation. Ph.D thesis, Texas A&M University, Texas

Chapter 62

Research on the Architecture of Cloud GNSS Based on Hadoop

Linyang Li, Zhiping Lu, Lihui Fan and Jian Li

Abstract According to the challenges of storage and computation faced by massive, multi-source and heterogeneous GNSS data, the design objective of cloud GNSS is analyzed, then the architecture of cloud GNSS from infrastructure, data management, service management to application is designed, the deployment model including service management platform, Web server cluster and multiple Hadoop clusters is provided, and its' characteristics such as strong expansibility, high reliability, loose coupling, are summarized. Cloud GNSS platform is built in the experiment, the storage model of massive GNSS data and the parallel computing model of GNSS network are built, distributed storage, parallel retrieval, sub-network division distributed computing and data publication are achieved. The result shows that the architecture proposed by the paper can be applied in the storage, processing and service publication of large-scale GNSS network.

Keywords GNSS · Hadoop · Cloud storage · Distributed computing

62.1 Introduction

With the construction of Continuously Operating Reference Station (CORS) in the worldwide, national and regional level, and the combination of combined CORS, the network scale of Global Navigation Satellite System (GNSS) is larger and larger. Facing with these challenges of storage and processing of massive GNSS

L. Li (✉) · Z. Lu · J. Li
School of Surveying and Mapping, PLA Information Engineering University,
Zhengzhou 450052, China
e-mail: lilinyang810810@163.com

L. Fan
Surveying and Mapping Information Center, Chengdu 610000, China

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_62

data, University NAVSTAR Consortium officially provided the conception of GNSS data cloud storage in 2014 [1]. Based on Hadoop Distributed File System (HDFS), Li et al. put forward a storage architecture of massive CORS data [2]. Xiong et al. proposed a small spatio-temporal files management scheme combining user access and data features [3]. The first group of International Association of Geodesy (IAG) efficiently process the data of global reference frame by distributed computing [4]. Enrico Serpelloni et al. calculated 80 CORS Stations in the way of distributed computing, and required accuracy was obtained [5]. Shorock et al. advanced a distributed processing strategy based on grid computing [6]. After that, LI Jian et al. established an architecture of large-scale CORS network [7].

Compared with GFS, MapReduce and BigTable in the Google, Hadoop is an open-source cloud platform which helps researchers study from the bottom layer. In recent years, owing to its' reliability, escalating property and economy, it develops into the most widely used platform. Based on Hadoop, according to the above challenges and the deficiency of existed theories which have not systematically researched on massive GNSS data, this paper systematically researches the architecture of cloud GNSS. Finally the experiment validates that the architecture proposed in this paper is reasonable and feasible.

62.2 Main Challenges Faced by GNSS

62.2.1 Storage and Management

There is about 20,000 GNSS reference stations in the world. It remains impossible to store and manage all GNSS data according to ideal standard at this stage.

There are two major GNSS data management ways, one is File Transfer Protocol (FTP), and another is Relational Database Management System (RDBMS). The International GNSS Service (IGS) usually use the former, although it is mature, it lacks unified management mechanism and concurrency control and recovery, and is lower in real-time performance, inconvenient in data retrieval. Beijing CORS applies the latter. It is flexible and convenient, but the retrieval efficiency reduces with the increase of data amount, the concurrent sharing and disaster tolerance are poor. And poor data transfer has a bad effect on the practicality. Thus, the main challenges faced by storage and management are listed as follows.

1. First, the problem of storage capacity. Facing with TB, PB and even ZB level of data, the current file system and RDBMS are hard to bear.
2. Second, current method can't meet the requirements of concurrent data read, write and access.
3. Finally, current method can't satisfy the needs of high expansibility and availability.

Centralized storage can't meet the needs of massive storage application. Seeking more efficient storage method has become one of the hot research areas [8].

62.2.2 Data Processing

Massive data provides rich computing resource, but it also brings challenges to the data processing. Figure 62.1a shows the amounts of estimated parameters of 30 GPS satellites. They reach up to 15,000 when 200 stations is processed. If other systems are considered simultaneously, the estimated parameters will be more.

Precise Point Positioning (PPP) and double-difference (DD) are the two main processing ways. For n stations, PPP needs the time of $O(n)$, but its' accuracy is lower than DD, the accuracy in the east component can be improved through restoring and fixing the integer ambiguity [9]. Besides, PPP ignores the correlation between baseline vectors, which is unbeneficial to the network adjustment.

However, DD needs the time of $O(n^3)$, the GAMIT processing time of different network scale is shown as Fig. 62.1b. With the increase of the station number, the computation time increases in the trend of geometric times. Most high accuracy software can only solve fewer than 100 stations, such as GAMIT 10.5, 99 stations are the upper limit. Therefore, large network is usually divided into several sub-networks, but it will cause the loss of some information, the robustness will be influenced by the reutilization of common points.

Limited by the computer memory and capacity, many researchers focus on reducing the computing expenditure, such as compressing data storage [10], eliminating and restoring estimated parameters [11], constraining PPP resolutions by double-difference of dependent baseline [12]. Seeking fast processing method of massive GNSS data also becomes one of the hot research areas [13].

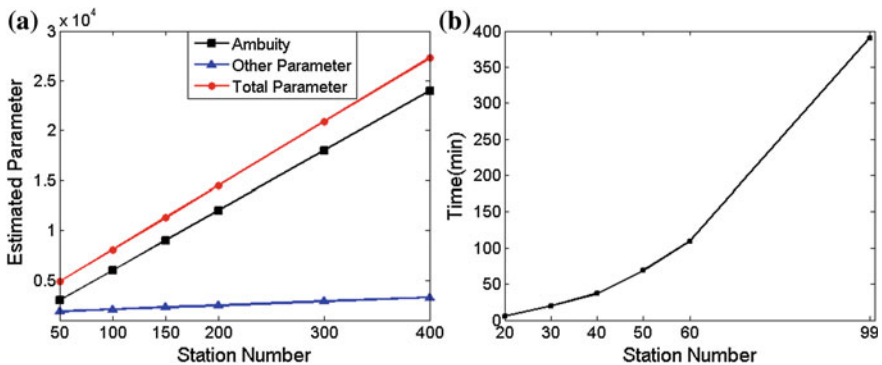


Fig. 62.1 Estimated parameter and computation time under different size of network. **a** The estimate parameter. **b** The computation time

62.3 The Architecture of Cloud GNSS Based on Hadoop

Based on cloud computing technology, the cloud GNSS is defined as: based on cloud computing and the relative theory, method and technology in the domain of GNSS, cloud GNSS aims to solve distributed storage, computing, sharing, interoperability and other key problem of large-scale GNSS data. And it offers service in the way of web applications; provides high-performance and high-throughput computing capacity in the form of software; supplies convenient and open Application Program Interface (API) to geophysicist in the form of platform.

62.3.1 Design Objective

Different cloud platform has different architecture and function, the design objectives proposed by the paper are listed as follows.

1. Resource integration and technology fusion. Hardware resources such as Internet, storage, computation and software resources such as database, adjustment, web services can be absorbed into the cloud platform.
2. The extensibility and configurability. For the same and different request, the extensibility includes GNSS data storage, computing and service. Meanwhile, the cloud platform has to be configured to increase or reduce service capacity.
3. Transparent GNSS data access interface. The platform provides user with standard data service, user gets access to data through transparent interface.
4. Standard GNSS service. GNSS data storage and computing should be offered as standard service through Internet.

62.3.2 The Architecture of Cloud GNSS

Shown in Fig. 62.2a, the cloud platform is established in the paper.

The first layer is infrastructure layer. Virtualization and interoperability are used to integrate heterogeneous resources, as a result, an integral resource pool is built, and underlying differences are shielded in the way of service. Finally, a unified virtualization and interoperability interface is provided to the upper layer.

The second layer is the cloud GNSS data management layer, composed by Hadoop modules. HDFS and HBase consist of a flexible distributed storage system and provide fast reading, writing and access; MapReduce and Hama are used to parallelly process GNSS data; Pig and Hive provide a higher level of GNSS data abstraction operation, structured data is queried and processed through a SQL-like syntax HiveQL, data processing ability over MapReduce is offered by Pig Latin; finally Zookeeper is responsible for the cluster management. GNSS data includes

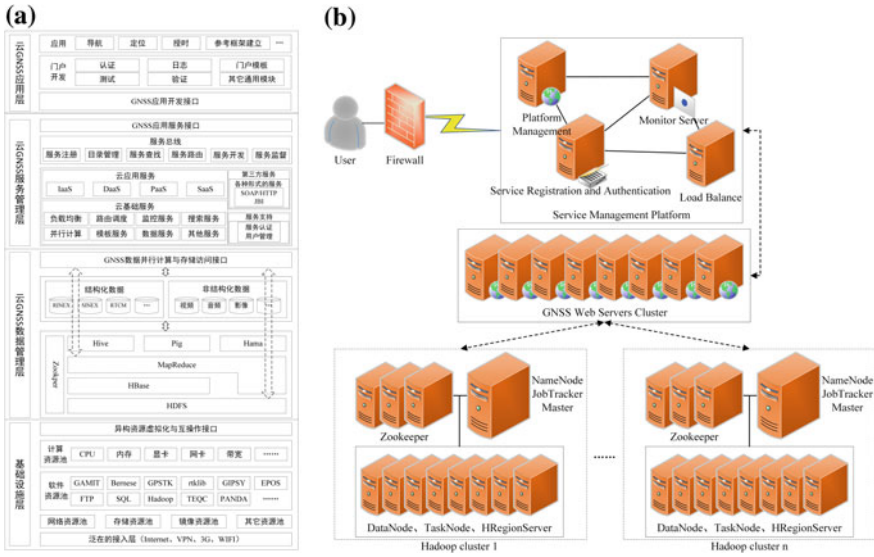


Fig. 62.2 The architecture and deployment model of cloud GNSS Based on hadoop. a the architecture b the deployment model

multi-system, multi-frequency observation data, products, metadata, video, audio and image data. Large files are stored in many blocks; Coordinates, polar motion and covariance matrix are stored in HBase. Archive tools or CombineFileInput-Format class or SequenceFile format is used to management small files, and it will reduce the storage capacity overhead and memory of the master. Meanwhile, these nodes install the processing software, invoke Bernese and GAMIT and large-scale network adjustment, etc. Eventually, this layer provides transparent GNSS data parallel computing and storage interface to the service management layer.

The third layer is the service management layer. It provides users with multi-type services, including cloud application and infrastructure service. The former includes general Infrastructure as a service (IaaS), specific Software as a service (SaaS), and the Platform as a service (PaaS). Besides, GNSS Data as a service is put forward. The latter includes load balancing, parallel computing and search services, etc. Service is registered through the service bus [14], self-help registration service and development are realized according to the actual need. GNSS application service interface is the services encapsulation which achieves the complete decoupling of service management and application layer, as long as the interface remains the same, two layers can evolve independently.

The fourth layer is cloud GNSS application layer. Service is provided through the portal web and various applications, various GNSS Services are encapsulated through the portal, and they are adapted into the service-oriented Architecture (SOA), specific application includes navigation, positioning and timing, etc.

62.3.3 The Deployment Model of Cloud GNSS

The deployment model of cloud GNSS is shown in Fig. 62.2b, it is divided into service management platform, Web server cluster and multiple Hadoop clusters.

Service management platform. Service registration and certification provide the service description and certification outwardly, inwardly it offers the detailed description of the user's requests. The load balancing server is mainly responsible for the management and distribution of users' request, and it monitors the running state of GNSS Web servers, allocates the service to the node who owns lighter load and the corresponding request service content.

GNSS Web server cluster is composed of multiple GNSS Web servers. Upwardly it receives and parses the URL, downwardly it connects different cluster by distributed interface, determines the required storage and computing resources.

The cloud GNSS platform can deploy many Hadoop clusters, each cluster contains three types of nodes. One is the master node, it integrates HDFS NameNode, MapReduce JobTracker and HBase Master into a body. Another is the computing and storage nodes which integrate HDFS DataNode, MapReduce TaskNode and HBase HRegionServer into a body, and they are the main carriers of the services. The last is the service coordinator nodes, made up of an odd number of ZooKeeper, they are responsible for cluster configuration and coordination, and prevent single point failure, load the cluster's balance.

62.3.4 The Characteristics of the Cloud GNSS

The architecture of cloud GNSS proposed in this paper combines Hadoop with massive GNSS data distributed management, computing and service together. The data storage, calculation model and transparent service, application development interface are designed, main characteristics are listed as follows.

1. Strong extensibility. Including the expansion of Web server cluster, Hadoop cluster and nodes in the cluster. The extensibility can effectively improve the service ability, storage and computing capacity.
2. High reliability and availability. In addition to inherit from Hadoop, if one Web node failures or a cluster collapses, it will not lead the platform to collapse.
3. Each layer decouples. Bottom layer provides transparent interface to the upper layer, each layer evolves independently.
4. High concurrent access. Through the queue mechanism, and reasonable storage strategy, the cloud GNSS supports high concurrent access.
5. High-performance computing. The parallel computing interface is designed and realized by MapReduce and Hama, the whole cluster's capacity can be made full use.

62.4 Experiment Analysis

62.4.1 The Establishment and Test of Experiment Platform

Two clusters, 13 Ubuntu 14.04 virtual machines are built by two servers. One works as the main node, the remaining 5 work as slave nodes in each cluster. All are deployed with Hadoop 2.5.1, and equipped with the Intel dual-core 2.0 GHz processor; the master node owns 2 GB memory, while 1 GB for the other. The cluster network topology and roles are shown in Fig. 62.3.

The platform's expansibility and disaster tolerance are tested. First, new nodes are deployed rapidly with Ambari. After new node joins in the cluster, load balancing is applied, and storage loads are distributed averagely. Then, after closing a slave node in cluster 1, failure is monitored through heartbeat, data is backup. There is no effect on the use of the platform during the whole process.

62.4.2 The Data Storage and Retrieval Based on HDFS and HBase

62.4.2.1 The Architecture of Cloud GNSS Based on Hadoop

Through the storage and management of HDFS, GNSS data are directly uploaded, data block is set as 32 MB. As shown in Fig. 62.4, the storage tree directory is established, the corresponding data blocks are located through the file path.

Based on HDFS, shown in Table 62.1, HBase structure is designed. Four characters of the station name is set as the primary key. As for observation data, column families are built according to the value type, such as C1 and P1, satellite number is used as column. Observation data is stored in the storage unit cell. In order to keep the continuity, null is tagged in the epochs of signal loss lock. There is only one column family for the results, including the coordinates, velocity and

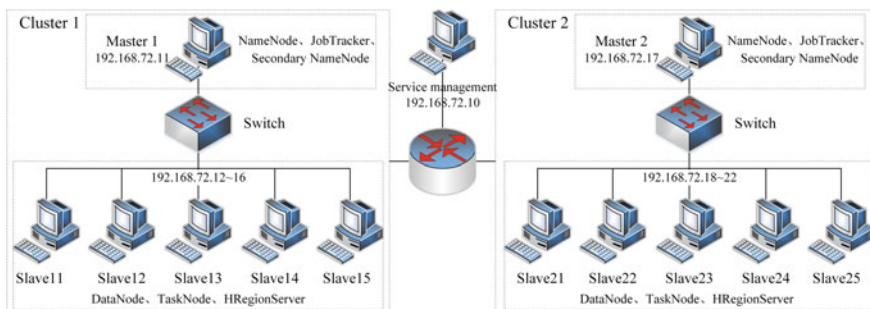


Fig. 62.3 The topological structure of cluster network

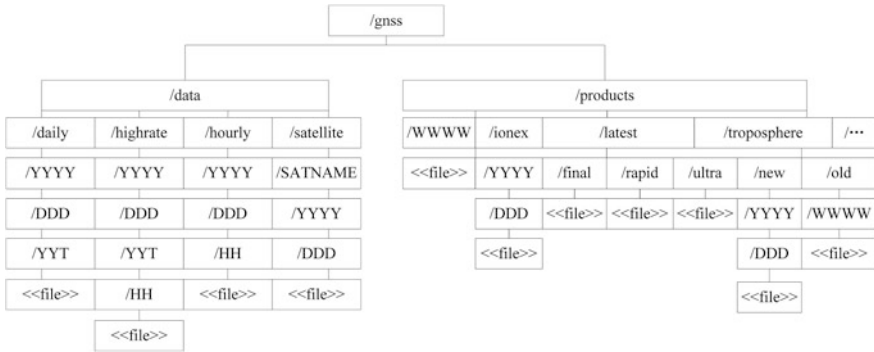


Fig. 62.4 The tree-structured directory

Table 62.1 The HBase architecture for GNSS data

	C1		P1	L1	L2	...	Results
	G1	G2	Coordinate velocity ... Covariance matrix
Station 1							
Station 2							
...							
Station n							

covariance matrix column. In order to distinguish different results in different sessions, version control is realized by the index of timestamp. Results are recorded by storage unit cell determined by rows keys, columns keys and timestamp.

62.4.2.2 The Parallel Retrieval of GNSS Data

In cluster 1, data is retrieved. The retrieval time includes index calculation and read. Curl client is configured to retrieve 50, 100, 200 observation files on March 18, 2014 from 2, 4, 6 nodes, the time is shown in Fig. 62.5. With the increase of the node number, the retrieval efficiency continues to improve.

62.4.3 The Large-Scale Network Solution Based on MapReduce

Based on MapReduce, GAMIT 10.5 is installed and configured in each node, as for the 10 subnets on March 18, 2014, map and reduce functions are achieved, distributed computation is realized. The detailed process is shown in Fig. 62.6.

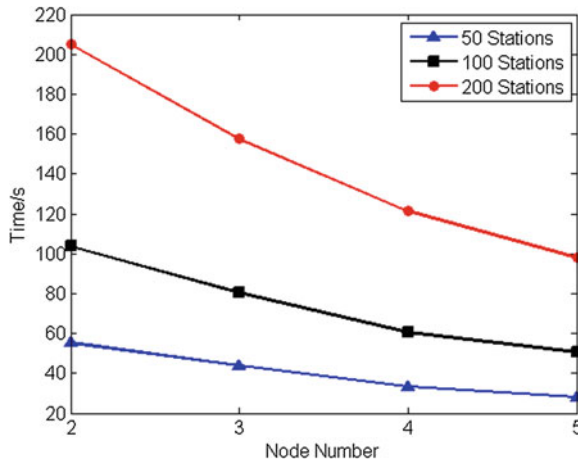


Fig. 62.5 The retrieval time of GNSS data

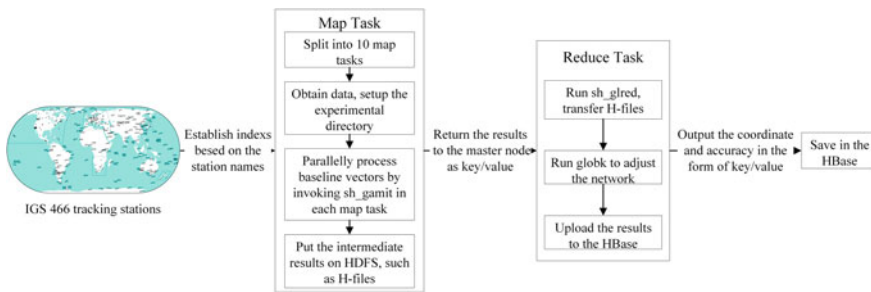


Fig. 62.6 The distributed computing flow of large-scale GNSS network based on MapReduce

1. First, the index file is established, station name is tagged with affiliated subnet.
2. Second, the index file is traversed, 10 map tasks are assigned to five nodes.
3. Third, each map task translates the indexes into the key/value pairs, the key represents the station name; the observation data is retrieved from HBase according to the value, the GAMIT project files are set up, sh_gamit command is invoked to automatically process the subnet. 10 H-files derived from map phase are stored in HDFS, and these ten files return to the master node.
4. Finally, reduce is completed at the master node, GLOBK is used to handle 10 H-file, station coordinates are obtained, the results are stored in HBase.

It costs about 590 min on a single node, while 120 min this method, which is about 3.9 times faster. So this method significantly take advantage of the computation capacity of the cluster, greatly improves the calculation efficiency.

62.4.4 The Data Service Publication

62.4.4.1 The Generation of Specified Interval File

The generating flow of observations with 30 s interval is shown in Fig. 62.7. An observation file is recognized as a Splitting task, and key/value pairs are generated. According to the interval, header is changed and corresponding epochs are extracted, 30 s interval file is generated after merging.

62.4.4.2 Data Visualization

Data access interface is developed into a plug-in by the 50,070 port and HBase Representational State Transfer (REST), visualization display of GNSS data is realized. Data from 466 IGS stations on March 18, 2014 is used, Fig. 62.8 shows the sub-networks division strategy of Scripps Orbit and Permanent Array Center (SOPAC), the left is the subnet 1, while the right is subnet 9.

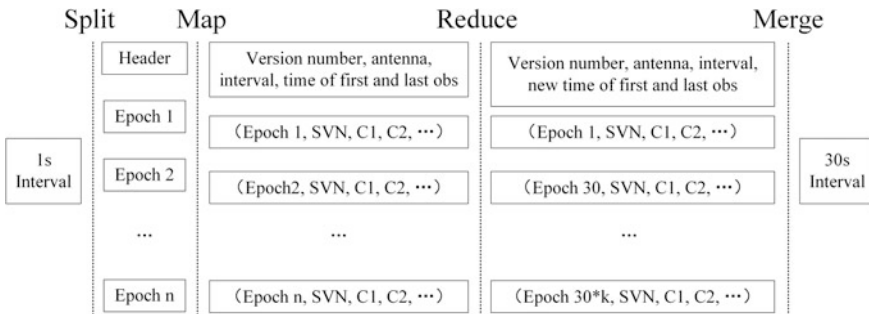


Fig. 62.7 The generating flow of observation with 30 s interval

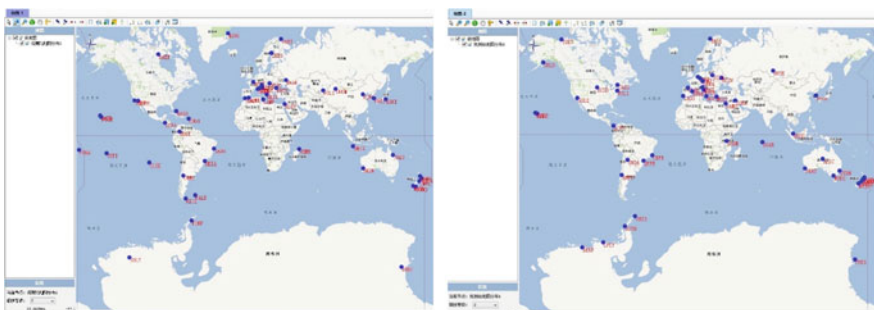


Fig. 62.8 The distribution of IGS sub-networks

62.5 Conclusion

With the development of multi-system and multi-frequency, construction of CORS and continuous observation, GNSS data shows a trend of blowout. Massive, multi-source and heterogeneous data has brought enormous challenges to the management and calculation. Work in this paper is carried out based on Hadoop.

1. The design goals of cloud GNSS are analyzed, the architecture is put forward, and it has the characteristics of high expansibility, reliability and so on.
2. Storage model is designed, GNSS data distributed storage and parallel retrieval are realized, the retrieval efficiency improves significantly.
3. IGS sub-network partition distributed computation is implemented based on MapReduce, the computational efficiency improves significantly.
4. Based on key/value pairs, specified interval file is generated. The release of GNSS data is achieved based on HDFS 50,070 and HBase REST.

Therefore, the architecture proposed in the paper can be used for the data storage, management, processing and publication of large-scale GNSS network.

Acknowledgments The work is funded by the National Natural Science Foundation of China, No. 41274015; the National 863 Program of China, No. 2013AA122501, State Key Laboratory of Geo-information Engineering, NO. SKLGIE2014-M-1-5 and NO. SKLGIE2014-M-1-6.

References

1. UNAVCO (2014) GPS/GNSS Data management overview. <http://www.unavco.org/data/gps-gnss/data-management/data-management.html>
2. Li L, Lu Z, Chen Z (2014) Cloud storage model of massive continuous operating reference station system data. *J Navig Positioning* 2(3):64–70
3. Lian X, Zhengquan XU, Tao W (2014) On the store strategy of small spatio-temporal data files in cloud environment. *Geomatics Inf Sci Wuhan Univ* 39(10): 1252–1256
4. Boomkamp H (2010) Global GPS, reference frame solutions of unlimited size. *Adv Space Res* 46:136–143
5. Serpelloni E, Casula G, Galvani A, Anzidei M, Baldi P (2006) Data analysis of permanent GPS networks in Italy and surrounding regions: application of a distributed processing approach. *Ann Geophys* 49(4/5):897–928
6. Schrock GM (2006) On-grid goal: seeking support for high-precision networks. *GPS World* 17(10):34–40
7. Li J, Li J (2009) A new computing theory model on large-scale CORS. *Geomatics Inf Sci Wuhan Univ* 34(8):992–995
8. Lu Z, Chen Z, Cui Y (2013) Distributed processing of large-scale CORS network's baseline vectors. *J Geomatics Sci Technol* 30(4):433–438
9. Ge M, Gendt G, Rothacher M (2008) Resolution of GPS carrier-phase ambiguities in precise point positioning (PPP) with daily observations. *J Geodesy* 82(7):389–399
10. Shi C (2002) Theory and application of the adjustment of large scale high precision GPS network. Surveying and Mapping Press, Beijing
11. Ge M, Gendt G (2006) Dick G. A new data processing strategy for huge GNSS global networks 80(4):199–203

12. Blewitt G (2008) Fixed point theorems of GPS carrier phase ambiguity resolution and their application to massive network processing: ambizap. *J Geophys Res* 113:B12410
13. Hua C, Jiang W, Ge M (2014) An enhanced strategy for GNSS data processing of massive networks. *J Geodesy* 88(9):857–867
14. Li D, Yao Y, Shao Z (2014) Big data in small city. *Geomatics Inf Sci Wuhan Univ* 39(6):631–640

Chapter 63

Wi-Fi Fingerprint Positioning Updated by Pedestrian Dead Reckoning for Mobile Phone Indoor Localization

Qiang Chang, Samuel Van de Velde, Weiping Wang, Qun Li,
Hongtao Hou and Steendam Heidi

Abstract The widespread deployment of Wi-Fi communication makes it easy to find Wi-Fi access points in the indoor environment, which enables us to use them for Wi-Fi fingerprint positioning. Although much research is devoted to this topic in the literature, the practical implementation of Wi-Fi based localization is hampered by the variations of the received signal strength (RSS) due to e.g. impediments in the channel, decreasing the positioning accuracy. In order to improve this accuracy, we integrate Pedestrian Dead Reckoning (PDR) with Wi-Fi fingerprinting: the movement distance and walking direction, obtained with the PDR algorithm, are combined with the K-Weighted Nearest Node (KWNN) algorithm to assist in selecting reference points (RPs) closer to the actual position. To illustrate and evaluate our algorithm, we collected the RSS values from 8 Wi-Fi access points inside a building to create a fingerprint database. Simulation results showed that, compared to the conventional KWNN algorithm, the positioning algorithm is improved with 17 %, corresponding to an average positioning error of 1.58 m for the proposed algorithm, while an accuracy of 1.91 m was obtained with the KWNN algorithm. The advantage of the proposed algorithm is that not only the existing Wi-Fi infrastructure and fingerprint database can be used without modification, but also that a standard mobile phone is sufficient to implement our algorithm.

Keywords Indoor localization · Wi-Fi fingerprint · K-Weighted nearest node algorithm · Pedestrian dead reckoning algorithm

Q. Chang (✉) · W. Wang · Q. Li · H. Hou
Collage of Information System and Management, National University of Defense
Technology, Changsha, Hunan, China
e-mail: qchang@telin.ugent.be

Q. Chang · S. Van de Velde · S. Heidi
TELIN Department, Ghent University, Ghent, Belgium

63.1 Introduction

Acquiring accurate location information is essential for many applications. Therefore, researchers developed several algorithms to estimate a user's position. Among the solutions, the most popular system is the Global Navigation Satellite System (GNSS), which includes GPS, Galileo and Beidou. Because of the accuracy of GNSS, i.e., normally the average positioning error is 3–10 m, GNSS is widely used for outdoor navigation. On the other hand, with the arrival of the era of mobile Internet, Location Based Service (LBS) developed dramatically: the need of indoor positioning has increased rapidly. However, the indoor environment severely degrades the accuracy of GNSS positioning or makes it totally impossible. As a result, several alternative positioning techniques for indoor positioning were proposed. Some of them are based on the GNSS, such as AGNSS (Assisted GNSS) or DGNSS (Difference GNSS), but most of them rely on other approaches, such as Wireless Sensor Networks (WSNs), cameras, Wi-Fi radio fingerprinting or inertial measurement units (IMUs). In general, all above mentioned algorithms have their strengths and weaknesses, when comparing them with respect to accuracy, complexity and deployment costs. As a result, there still no well-performing positioning technique for indoor localization exists.

Among the solutions for indoor positioning, the Wi-Fi fingerprinting technique has received much attention because Wi-Fi access points are already widely available, implying the deployment costs are negligible compared to other solutions, and some commercial products are already developed, such as google maps, Wi-FiSlam or Rtmmap. Because of the weak relationship between the RSS and the position of the user, a Wi-Fi fingerprinting positioning algorithm consists of two phases: training and localization. First, during the training phase, Received Signal Strength (RSS) samples from the Access Points (APs) are collected and stored in a database together with their location coordinates. Next, in the localization phase, a user's current position is estimated based on the comparison of the measured RSS and those stored in the database. The requirement of an accurate database is the weak point of this technique: because of the Wi-Fi variance problem [1], which is caused by differences in the used device type, the user's direction, measurement time and environmental changes between the two phases, the estimation error is 10 m or even worse, such that the database must be updated regularly. The large estimation error in an outdated database is mainly caused by the selection of irrelevant reference points (RP) that are far from the actual position of the user.

A second widely used indoor positioning technique is Pedestrian Dead Reckoning (PDR) [2], based on information obtained from IMUs. In this technique, raw data from e.g. an accelerometer, a compass and a gyroscope is fused to estimate a user's trajectory. A major advantage of this technique is that no infrastructure is needed to estimate the relative trajectory of a user, although additional fixed anchors are required to find the absolute position of the user.

Both PDR and Wi-Fi fingerprint positioning have their strengths and weaknesses. The PDR algorithm has the advantage of high availability, and immunity to

external environment changes, but the downside of this technique is that it suffers from a drift error that increases with time: e.g., [2] reports a position offset after a 1 km walk of about 10 m, but short term results are accurate. In contrast to the PDR system, the positioning accuracy of the Wi-Fi fingerprinting positioning technique is reasonably low, even on the long term, but susceptible to external disturbances which lead to erratic, but bounded localization errors. Due to the complementary error behaviour, the combination of these two algorithms is expected to have better performance than the two single algorithms. In this paper, we integrate Pedestrian Dead Reckoning (PDR) with Wi-Fi fingerprinting to provide an accurate positioning algorithm. The short term moving distance and walking direction from PDR are applied to assist the KWNN algorithm to select reference points closer to the actual position, so that the positioning accuracy is improved. Hence, the proposed algorithm offers a solution to the RSS variance problem and the aging of the database, as the outdated database still can be used. Therefore, our algorithm reduces the maintenance cost of the system as the database should be updated less regularly.

63.2 Related Works

There is a vast literature on hybrid positioning techniques, combining two or more approaches to estimate a user's position. By combining measurements from different sources, researchers attempt to improve the accuracy of a single approach. Hence, the combination of Wi-Fi fingerprinting and IMU has been considered earlier.

For example, Xiao [3] developed a stochastic system model based on a finite state machine that utilizes the Wi-Fi fingerprint position estimates as its measurements, and the inertial sensing data as control inputs to track the target's position. Although this algorithm improves the positioning accuracy, it comes at the cost of a high computational complexity. A similar approach was used by Korbinian [4] to fuse data from the IMU and the Wi-Fi fingerprint algorithm. However, Korbinian considered shoe mounted IMU devices, such that the practical use for daily life is limited. Both approaches [3] and [4] considered Kalman filters for combining the results, but other types of filters, such as a particle filter [5] are also being considered: HiMLoc [6] combines location tracking and activity recognition using inertial sensors and Wi-Fi fingerprinting via a particle filter. However, HiMLoc requires the knowledge of a basic map including locations of stairs, elevators, corners and entrances. The IMU and Wi-Fi fingerprint based algorithm 'Zee' [7] also needs a map showing the pathways and barriers. Berkovich [8] develops a navigation engine that combines the measurements from a 3D accelerometer, a gyroscope, a magnetometer, Wi-Fi and BLE modules, together with a floor map. The real-time indoor positioning accuracy of the engine is about 1–2 m, but this algorithm is high energy consuming. Herrera [9] creates an indoor positioning algorithm using a particle filter to combine PDR, beacon-based Weighted Centroid position estimates, map information from OpenStreetMap and a users path density

map. This high-energy-consuming algorithm obtains an average accuracy of 2.48 m. Chai [10] presents a PDR/Wi-Fi/barometer integrated system, where an adaptive Kalman filter is employed for sensor fusion. As a barometer is not always available to the users, the practical use of the algorithm is limited. In [11], Jin presents a nearest-neighbor selection algorithm for real-time Wi-Fi fingerprint positioning with the assist of inertial measurement unit (IMU) measurements. The algorithm first selects several RPs according to the conventional KWNN algorithm. Then, filtering out irrelevant reference points based on the position prediction with IMU measurements.

Comparing with the current literature, we combine the results on a much lower level, i.e., we incorporate the movement distance and walking direction directly in the KWNN algorithm. Further, the sensors that are used in our algorithm are a gyroscope and an accelerometer, which are readily available in standard mobile phones. Our algorithm has a very low complexity, especially compared to a particle filter, and requires no assumptions about a noise model. This is in contrast with both the Kalman and particle filters, which both need the knowledge of the noise parameters.

63.3 System Description

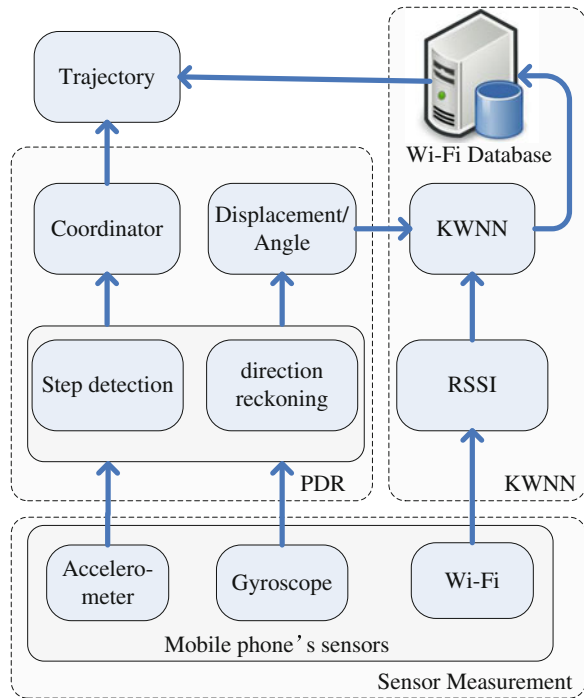
63.3.1 Algorithm Framework

The architecture of the proposed algorithm is given in Fig. 63.1. The algorithm contains three stages: sensor measurement, PDR calculation, and KWNN localization.

In the measurement phase, the algorithm records the internal sensor readings of the mobile phone, which is equipped with an accelerometer, a gyroscope, and a Wi-Fi card. We use the Wi-Fi card to obtain the APs' Radio Signal Strength, the gyroscope to measure rotational forces along the device's three axes, and the accelerometer to measure the acceleration of the device. The digital compass, which is also available in the device, is not used in this algorithm, as it is easily affected by external magnetic fields and operating electronic devices, resulting in non-reliable measurements. In the PDR calculation phase, the readings of the accelerometer and gyroscope are fused to detect the number of steps and walking direction. Based on the number of steps and direction, an estimate of the user's coordinates can be determined. In the last phase, the KWNN algorithm is performed after every m steps, $m = 1, 2, \dots$. By reducing the number of KWNN executions, i.e., by increasing m , the location server's computation load is reduced and the mobile phone's battery life is prolonged.

A detailed description of the PDR and KWNN algorithm is given in the following sections.

Fig. 63.1 Architecture of the proposed algorithm

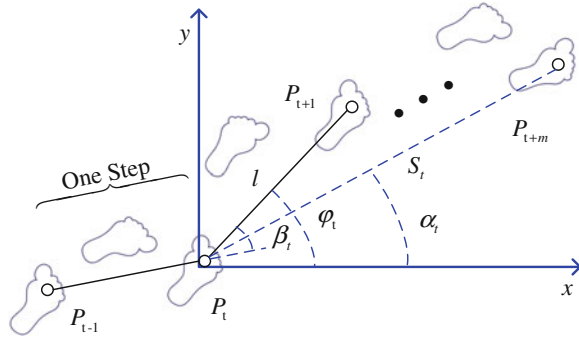


63.3.2 PDR Algorithm

In the PDR algorithm, the readings of the accelerometer and gyroscope are fused to detect the user’s number of steps and walking direction. The process of the step counting is out of the scope of this paper and will not be discussed. The walking direction is estimated by integrating the previous location, and the readings of the gyroscope and the accelerometer. Based on the number of steps and direction, the coordinates (x, y) , moving distance S_t and turning angle α_t can be determined (see Fig. 63.2). We define a step in our algorithm as two actual footsteps, one of both feet, i.e., from a step from the right (or left) foot to the next step from the right (or left) foot. In this paper, P_t refers to the true position of the user at time t , $P_t = \{x_t, y_t\}$, and $P_{t,PDR}$ and $P_{t,WiFi}$ are the positioning results from the PDR and the Wi-Fi fingerprint algorithm at time t , respectively.

Assuming that we have the position estimate from the Wi-Fi fingerprint algorithm at time t , i.e. $P_{t,WiFi}$, we can estimate the locations in the next m steps with the PDR algorithm:

Fig. 63.2 PDR algorithm



$$\begin{cases} x_{t+k,\text{PDR}} = x_{t,\text{WiFi}} + \sum_{i=1}^k l \cos \varphi_{t+k-1} \\ y_{t+k,\text{PDR}} = y_{t,\text{WiFi}} + \sum_{i=1}^k l \sin \varphi_{t+k-1} \end{cases}, k = 1, 2, \dots, m \quad (63.1)$$

In Eq. 63.1, l is the step length and φ is the walking direction as illustrated in Fig. 63.2. The angle φ_t can be calculated as follows:

$$\varphi_t = \begin{cases} \varphi_{t-1} + \beta_t, & t > 0 \\ 0, & t = 0 \end{cases} \quad (63.2)$$

After m steps, the moving distance S_t and turning angle α_t yield:

$$\begin{cases} S_t = \|P_{t+m,\text{PDR}} - P_{t,\text{WiFi}}\| \\ \alpha_t = \arctan \frac{y_{t+m,\text{PDR}} - y_{t,\text{PDR}}}{x_{t+m,\text{PDR}} - x_{t,\text{PDR}}} \end{cases} \quad (63.3)$$

where $\|\bullet\|$ is the Euclidean distance.

In Eq. 63.1, we have introduced a fixed step length l to compute the coordinates with the PDR algorithm. Note that, in reality, the step length l not only varies from person to person, but also for a single person the step length is not the same all the time. Nevertheless, it is observed that on short term, in general, the step length of a person will not change significantly. Therefore, we can update the step length based on the historic walking data.

$$l_t = \frac{\|P_{t-m,\text{WiFi}} - P_{t,\text{WiFi}}\|}{\sqrt{(\sum_{i=1}^m \cos \varphi_{k-m+i})^2 + (\sum_{i=1}^m \sin \varphi_{k-m+i})^2}} \quad (63.4)$$

In the following, for notational convenience, we drop the dependency of the step length on the time index.

63.3.3 KWNN Based Wi-Fi Fingerprint Algorithm

Many algorithms can be used to estimate the user's position based on the RSS measurements and the available database. In this paper, we adopt the K Weighted Nearest Neighbor (KWNN) Wi-Fi fingerprint positioning algorithm. In this algorithm, to estimate the position, the K nearest neighbor reference points (RPs) are selected based on the RSS signal distance, i.e., the difference between the measured RSS and the RSS values available in the database. With this algorithm, the coordinates are calculated as:

$$P_{t+m, \text{WiFi}} = \left\{ \sum_{i=1}^k x_{DB,i} w_i, \sum_{i=1}^k y_{DB,i} w_i \right\} \quad (63.5)$$

where $(x_{DB,i}, y_{DB,i})$ are the coordinates of RP i , and the weight w_i is defined as:

$$w_i = \frac{1/\varepsilon_{DIS,i}^p}{\sum_{j=1}^k 1/\varepsilon_{DIS,i}^p} \quad (63.6)$$

In Eq. 63.6, $\varepsilon_{DIS,i}$ is the RSS signal distance and p is a parameter that can be changed to optimize the positioning accuracy.

In the standard KWNN algorithm, $\varepsilon_{DIS,i}$ is determined by the difference between the RSS values available in the database and the measured RSS value between the user. Because of signal blocking, the RSS values of the nearest reference points can differ significantly from the measured RSS value. This can be illustrated by Fig. 63.3. The nearest 4 reference points, selected by the conventional KWNN algorithm, are not the best ones due to the RSS variance problem. In order to improve the estimation accuracy, we include information obtained from the PDR algorithm in the expression for the error distance $\varepsilon_{DIS,i}$. Note that these two parameters are both acquired on short term only, to avoid the error accumulation in the PDR step. Using this distance definition, the KWNN algorithm is able to select the most relevant RPs, as illustrated in Fig. 63.3.

With the PDR algorithm, we are able to calculate the moving distance and turning angle between two Wi-Fi fingerprint estimations. We include these in the error distance $\varepsilon_{DIS,i}$ as follows:

$$\varepsilon_{DIS,i} = D_{RSS,i} + \lambda D_{LOC,i} + \gamma D_{AGL,i} \quad (63.7)$$

where $D_{RSS,i}$ is the signal distance from the standard KWNN algorithm:

$$D_{RSS,i} = \left(\sum_{j=1}^n |RSSI_{DBi,j} - RSSI_{MR,j}|^q \right)^{1/q} \quad (63.8)$$

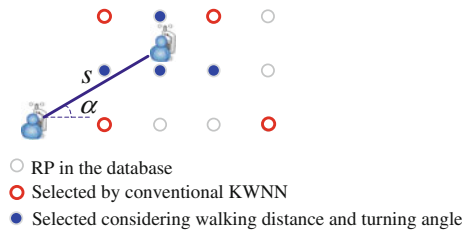


Fig. 63.3 PDR assisted RP selection

n is the number of access points (APs), $RSSI_{DBi,j}$ is the RSS value from the database for AP j , measured in RP i , $RSSI_{MRj}$ is the RSS value measured by the user, and the parameter q can be changed in order to optimize the accuracy.

The location distance $D_{LOC,i}$ is the location distance:

$$D_{LOC,i} = \left| \|P_{DB,i} - P_{t,WiFi}\| - S \right|^q \tag{63.9}$$

where $P_{DB,i}$ is the coordinate of RP i , and $P_{t,WiFi}$ is the coordinate of the previous Wi-Fi fingerprint position estimate. Finally, the angular distance $D_{AGL,i}$ is given by:

$$D_{AGL,i} = \left| \left(\arctan \frac{y_{DB,i} - y_{t,WiFi}}{x_{DB,i} - x_{t,WiFi}} - \alpha_t \right) \% 180 \right|^q \tag{63.10}$$

The location distance $D_{LOC,i}$ and the angular distance $D_{AGL,i}$ in Eq. 63.7 are added to the error distance with the weighting factors λ and γ , which can be selected to optimize the performance.

63.4 Performance Analyses

To evaluate the proposed algorithm, we have created a 3D model of an office environment covering a total area of over 900 m². Eight APs are present in this environment. The radio map for each AP is computed by means of 3D ray tracing.¹ The floor plan of the office area and the coordinates of the APs are shown in Fig. 63.4.

Figure 63.5 shows the RSS radio map for AP 1.

In our simulation setup, we created from the training data a database with 300 RPs. The obtained radio map, originating from the ray tracing program, is considered as the ground truth, and we use it to generate RSS measurements by adding zero mean Gaussian noise with a standard deviation of 5 dBm. For the step length estimation, we define a step length of 1 m and add 0 mean Gaussian noise with a

¹We use the WinProp program from AWE Communications for the 3D ray tracing.

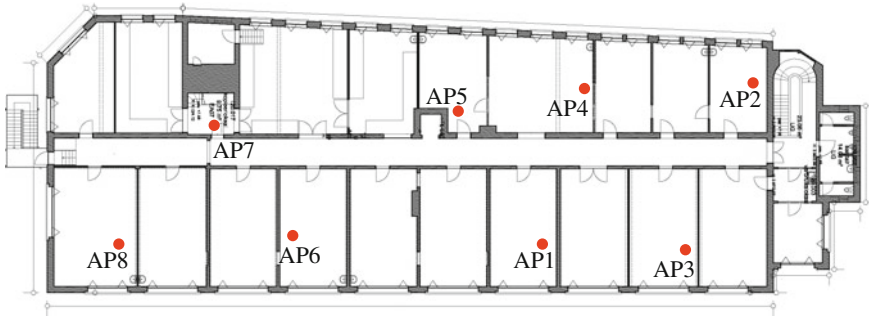


Fig. 63.4 Floor Plan of the indoor environment and the distribution of the APs

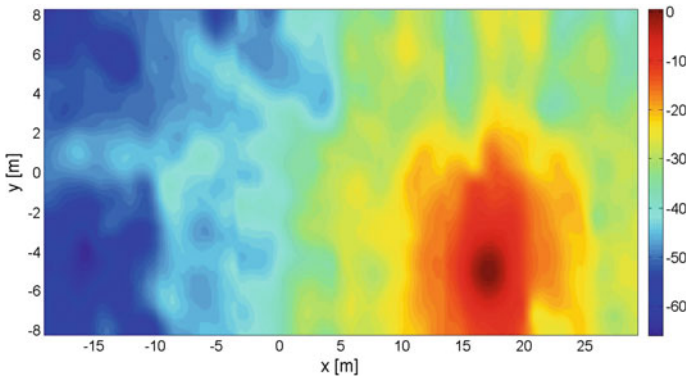


Fig. 63.5 Radio map of AP 1

standard deviation of 0.1 m. The walking direction equals the true direction distorted by zero mean Gaussian noise with a standard deviation of 1 degree. Simulations are performed for 2D localization and the conventional KWNN algorithm is compared with the proposed algorithm. During the simulation, the parameters are selected as follows: $k = 4$, $p = q = 2$, $\lambda = \gamma = 1$, $m = 1$.

The CDF (Cumulative Distribution Function) of the estimation error is shown in Fig. 63.6.

We observe that the new algorithm results in a higher accuracy than the KWNN algorithm: for the proposed algorithm, 80 % of the positioning errors is smaller than 2.05 m, whereas 2.55 m for the KWNN algorithm. Hence, the accuracy is improved with 20 %. Further, the probability of obtaining an error below 2 m is 78 % for the proposed algorithm, as compared to a probability of 59 % for the KWNN algorithm. The average error for the new algorithm is 1.58 m, while 1.91 m for KWNN. Hence, the performance is improved with 17.11 %.

Figure 63.7 shows the RMSE (Root-Mean Square Error) when the number m of steps between two Wi-Fi fingerprinting estimates varies.

Fig. 63.6 CDF curve comparing different algorithms

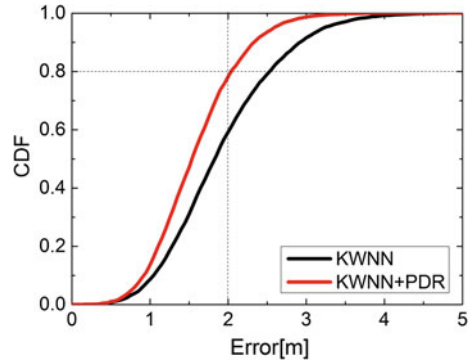
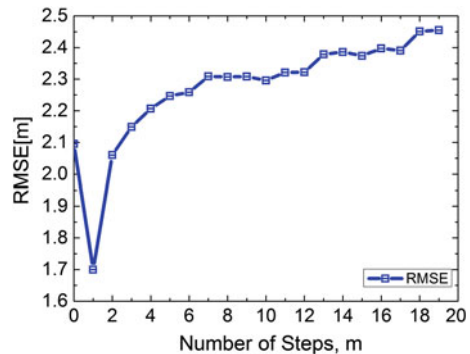


Fig. 63.7 RMSE curve of different number of steps



From Fig. 63.7, it follows that the minimum RMSE is achieved when the number m of steps equals 1, and when m increases, the RMSE increases. This can be explained as the accuracy of the PDR algorithm degrades when the number of steps grows. Hence, although the battery life of the mobile phone benefits from reducing the number of KWNN executions, the location error increases as a result. Nevertheless, our simulation results show that the resulting RMSE is only slightly larger than the RMSE of the conventional KWNN algorithm, even if the number of steps is increased. Hence, to reduce the power consumption of the positioning algorithm, the proposed algorithm offers a solid solution.

63.5 Conclusions and Future Work

In this paper, we propose a novel indoor positioning algorithm based on Wi-Fi fingerprint and PDR. The moving distance and walking direction from the PDR are used to assist the KWNN algorithm to select the most relevant RPs. Extensive simulations demonstrate that the proposed algorithm provides more accurate position estimates than the KWNN-based Wi-Fi fingerprinting positioning

algorithm, if in every step the RSS is measured. Further, we can reduce the complexity of the algorithm by considering RSS measurements only after m steps, by using the information of the PDR algorithm, without significantly degrading the performance. Our algorithm has very low complexity, especially as compared to a particle filter used as in [5], and offers the same accuracy. Moreover, no knowledge about the noise model is required. The existing Wi-Fi infrastructure and fingerprint database can be used without modification, and a standard mobile phone is sufficient to implement our algorithm.

In the future, the parameters of the algorithm (p , q , λ and γ) should be optimized such that the average positioning error is minimized. Further, real-life measurements should be conducted to test the algorithm.

Acknowledgments This work is supported by the Belgian National Fund for Scientific Research (FWO Flanders). Further, the first author gratefully acknowledges the China Scholarship Council (CSC) for their financial support.

References

1. Kim Y, Shin H, Chon Y, Cha H (2013) Smartphone-based Wi-Fi tracking system exploiting the RSS peak to overcome the RSS variance problem. *Elsevier Pervasive Mobile Comput* 9(3)
2. Beauregard, Stephane, and Harald Haas (2006) Pedestrian dead reckoning: a basis for personal positioning. In: *Proceedings of the 3rd workshop on positioning, navigation and communication*
3. Xiao W, Ni W, Toh YK (2011) Integrated Wi-Fi fingerprinting and inertial sensing for indoor positioning. In: *2011 International conference on IEEE indoor positioning and indoor navigation (IPIN)*, pp 1–6
4. Frank K, Krach B, Catterall N et al (2009) Development and evaluation of a combined wlan and inertial indoor pedestrian positioning system. In: *ION GNSS*
5. Atia MM, Korenberg MJ, Noureldin A (2012) Particle-filter-based WiFi-Aided reduced inertial sensors navigation system for indoor and GPS—denied environments. *Int J Navig Obs*
6. Radu V, Marina MK (2013) HiMLoc: Indoor smartphone localization via activity aware pedestrian dead reckoning with selective crowd sourced WiFi fingerprinting. In: *2013 International conference on IEEE indoor positioning and indoor navigation (IPIN)*, pp 1–10
7. Rai A, Chintalapudi KK, Padmanabhan VN et al (2012) Zee: zero-effort crowdsourcing for indoor localization. In: *Proceedings of the 18th annual international conference on mobile computing and networking. ACM*, pp 293–304
8. Berkovich G (2014) Accurate and reliable real-time indoor positioning on commercial smartphones. In: *International conference on indoor positioning and indoor navigation*, pp 27–30
9. Herrera JCA, Plöger PG, Hinkenjann A et al (2011) Pedestrian indoor positioning using smartphone multi-sensing, radio beacons, user positions probability map and indoor floor plan representation. In: *2011 International conference on IEEE indoor positioning and indoor navigation (IPIN)*, pp 1–6
10. Chai W, Chen C, Edwan E et al (2012) 2D/3D indoor navigation based on multi-sensor assisted pedestrian navigation in Wi-Fi environments. In: *2012 IEEE ubiquitous positioning, indoor navigation, and location based service (UPINLBS)*, pp 1–7
11. Jin M, Koo B, Lee S et al (2014) IMU-Assisted nearest neighbor selection for real-time WiFi fingerprinting positioning. In: *2014 International conference on IEEE indoor positioning and indoor navigation (IPIN)*, pp 1–6

Chapter 64

Sequence-Based Motion Recognition Assisted Pedestrian Dead Reckoning Using a Smartphone

Chengxuan Liu, Ling Pei, Jiuchao Qian, Lin Wang, Peilin Liu and Wenxian Yu

Abstract PDR (Pedestrian Dead Reckoning) is a relative positioning method using step length and heading. However, the step length is varying while a pedestrian performs different motions. With estimated motion state, the adaptive step length model can be applied. Considering that human motion is a continuous procedure, in this paper, we propose a sequence-based motion recognition method which estimates the motion states from a sequence of data. In contrast with traditional classifiers, this paper deploys a HMM (Hidden Markov Model) to infer the state labels of sequence motion. And with the assist of motion recognition, a 3D indoor pedestrian localization is presented. Experimental results show that the classification accuracy of sequence-based motion recognition is improved comparing to that of using Naïve Bayes classifier on standalone motion states. Furthermore, the positioning accuracy of a pedestrian in indoor environments is promoted using proposed method in this paper. The mean positioning error is reduced from 0.78 to 0.30 m. The 50th and 95th percentile errors are also cut down in the test within a typical office building.

Keywords Motion recognition · Sequence-based · PDR · 3D localization

64.1 Introduction

The outdoor positioning systems such as GPS, GLONASS, and BDS are well known and have been used successfully in our daily lives [1]. However, GNSS signals are degraded or denied in indoor environments. Therefore, the indoor positioning has been a hot topic recent years. Indoor localization technologies can

C. Liu · L. Pei (✉) · J. Qian · L. Wang · P. Liu · W. Yu
Shanghai Key Laboratory of Navigation and Location-Based Services,
School of Electronic Information and Electrical Engineering,
Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: ling.pei@sjtu.edu.cn

be applied for diverse purposes such as personal navigation, asset tracking, robot localization, ambient living assist, and so on.

With the fast growing popularity of smartphones and giant buildings, there is an explosive demand in indoor location based services using smartphones. PDR is an algorithm which could utilize sensors such as accelerometer, gyroscope and magnetometer to provide location services [2]. Since these sensors have been embedded into smartphones, smartphone-based PDR system becomes an alternative of indoor positioning methods. PDR system detects the gait of a pedestrian, and then estimates the heading and step length to infer the next position of the pedestrian given the current position. Since PDR is a relative positioning method, given irregular step lengths, the positioning error will be accumulated over time.

When a pedestrian moves, however, the motion state such as still, walking normally, turning, climbing stairs can provide extra information. This can help us to reduce the localization error. For example, when the pedestrian is turning, the step length is usually smaller than that of the normal walking [3]. Therefore, adaptive step length model can be applied according to the detected motion state. What's more, there is a height variation when the pedestrian using stairs, and since pure PDR system can't know the altitude information, it only works in the two-dimensional situations. By detecting the staircases, the height information is added. Therefore, the 3D PDR can be implemented.

This paper is organized as follows. The sequence-base motion recognition using HMM is given in Sect. 64.2; the motion recognition assisted PDR is presented in Sect. 64.3; the evaluation is illustrated in Sect. 64.4; and finally the conclusion is made in Sect. 64.5.

64.2 Sequence-Based Motion Recognition

Since the motion of a pedestrian is a continuous process, the current motion is influenced by the previous motion a pedestrian has performed. Taking the relation between adjacent motions into account, we use sequence-based motion recognition to promote the classification accuracy.

64.2.1 Feature Extraction

In this paper, a three-axis accelerometer and a three-axis gyroscope in a smartphone are used to collect the raw motion and orientation data of a pedestrian. To eliminate the noise, we use a three-stage moving average filter. This is simple, but it can reduce jitters effectively.

Features in time and frequency domain are used to constitute the feature vector. The sensor sampling rate is 32 Hz. One second is taken as a sample slot and FFT algorithm is applied to get frequency domain features. To get the features of

motions effectively, a 50 % overlapping moving window is applied. This has been demonstrated useful in previous work [4].

The target device in this paper is a smartphone. We hold it in our hands with the phone screen facing upper. The smartphone is kept relatively stationary which is the same as the previous work [5, 6]. Although the information about the orientation is lost, we can use the information of gravity. In the beginning, we keep still for about 5 s. Then, three-axis accelerations are calculated for the average. Thus, the gravity is got as follow:

$$\vec{\mathbf{g}} = (g_x, g_y, g_z) \quad (64.1)$$

where g_x , g_y and g_z are the projections of gravity on the three axes of the sensor coordinate system. And vector $\vec{\mathbf{g}}$ points to the center of the earth.

Instead of the total acceleration, the linear acceleration actually infers the dynamics of a pedestrian due to the gravity is a constant at a given location. Therefore, features from the linear acceleration are adopted for motion recognition. Suppose the acceleration vector is $\vec{\mathbf{a}} = (a_x, a_y, a_z)$, then the linear acceleration is:

$$\vec{\mathbf{a}}_l = \vec{\mathbf{a}} - \vec{\mathbf{g}} = (a_x - g_x, a_y - g_y, a_z - g_z) \quad (64.2)$$

With the direction information of gravity, we can get the vertical component of linear acceleration by projecting it to the gravity vector. And the horizontal component is inferred later since a vector can be decomposed into vertical and horizontal components.

In order to distinguish the turning motion from other motion states, the rotated angle feature extracted from the gyroscope reading is also applied. All features we use in a feature vector are listed in Table 64.1.

Table 64.1 Features definition

Features	Definition
linaccM_mean	Mean of module of linear acceleration
linaccM_var	Variance of module of linear acceleration
linaccV_mean	Mean of module of vertical linear acceleration
linaccV_var	Variance of module of vertical linear acceleration
linaccH_mean	Mean of module of horizontal linear acceleration
linaccH_var	Variance of module of horizontal linear acceleration
angle_rotation	Angle rotated around the vertical direction
firstfreq	The first dominant frequency
firstpeak	Amplitude of the first dominant frequency
secondfreq	The second dominant frequency
secondpeak	Amplitude of the second dominant frequency
energy	The mean energy of linear acceleration

64.2.2 Hidden Markov Model

In most cases, the current motion a pedestrian is performing will influence the upcoming motion. For example, if a pedestrian is walking, the most probable motion he will perform is still walking. This knowledge can provide us extra information for motion recognition. In this paper, we use the Hidden Markov Model [7] to build the sequence-based motion recognition model.

The Hidden Markov Model is derived from the Markov Chains which is a double stochastic process with an underlying stochastic process that is not observable and can only be inferred from the observation sequences. A typical HMM can be represented as $\lambda = (N, M, A, B, \pi)$:

1. N is the number of hidden states. In HMM, states connect with each other, and every state can be reached from another state. We use $S = \{s_1, s_2, \dots, s_N\}$ to denote the set of all states.
2. M is the number of observation symbols. We use $V = \{v_1, v_2, \dots, v_M\}$ to denote the set of all observation symbols.
3. A is the state transition probability distribution matrix. $A = \{a_{ij}\}$, and $a_{ij} = p(q_{t+1} = s_j | q_t = s_i)$, $1 \leq i, j \leq N$, where q_t denotes the state at time t .
4. B is the observation probability distribution in each state. $B = \{b_j(k)\}$, and $b_j(k) = p(v_k | q_t = s_j)$, $1 \leq j \leq N, 1 \leq k \leq M$, where v_k denotes the observation symbol at time t .
5. π is the initial state probability distribution. $\pi = \{\pi_j\}$, and $\pi_j = p(q_1 = s_j)$, $1 \leq j \leq N$.

According to the distribution of B , HMMs are categorized into Discrete Hidden Markov Model (discrete observation probability distribution) and Continuous Hidden Markov Model (continuous probability distribution).

In this paper, since the motion features are continuous variables, we use the Continuous Hidden Markov Model. And it is presented as follows:

1. Hidden states: The hidden states in our model are the motion states including still, walking, turning, going upstairs, going downstairs, and so on. They are listed in Table 64.2.
2. Observation variables: The observation variables are the features of a vector, which are extracted from a time slot. We use $O = (f_1, f_2, \dots, f_M)$ to represent the

Table 64.2 Motion state definitions

State	Definition
s1	Going downstairs
s2	Going upstairs
s3	Still
s4	Turning
s5	Walking

Table 64.3 Motion states transition probabilities

	Upstairs	Downstairs	Still	Turning	Walking
Upstairs	0.78	0.01	0.01	0.1	0.1
Downstairs	0.01	0.78	0.01	0.1	0.1
Still	0.05	0.05	0.69	0.01	0.2
Turning	0.1	0.1	0.01	0.3	0.49
Walking	0.05	0.05	0.1	0.1	0.7

feature vector, where M is the length of the feature vectors and f_i is the i -th feature.

3. Transition probabilities: The transition probability between two motion states. Every two different motion states correspond to a different transition probability. We get the transition probability matrix from our experience of observing the experiment procedures. They are shown in Table 64.3.
4. Observation probabilities: The observation probability describes the probabilities distribution of the observation variables at each hidden state. The multi-dimensional Gaussian distribution is used to describe the observation probability distribution, and it is defined as:

$$p(O|s_i) = \frac{1}{(2\pi)^{M/2} |\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(O - \mu_i)^T \Sigma_i^{-1} (O - \mu_i)\right), 1 \leq i \leq N \quad (64.3)$$

where $\mu_i = E[O]$ is the mean of feature vectors and $\Sigma_i = E[(O - \mu_i)(O - \mu_i)^T]$ is the covariance matrix of feature vectors when the hidden state is s_i .

5. Initial state distribution: The initial state distribution describes the initial probability distribution of the hidden motion states. Since the motion states mainly include still, walking, turning, going upstairs and going downstairs, we get the initial state probability from our experience.

64.3 Motion Recognition Assisted Pedestrian Dead Reckoning

By observing the characteristic of human walking behaviors, researchers have proposed a new indoor positioning method called pedestrian dead reckoning [2]. It uses the inertial sensors to estimate the position of the pedestrian and has the ability to locate the pedestrian independently. However, there are some problems when it is applied, such as step length error caused by special motion and the infeasibility of 3D localization. With the assist of motion recognition, these problems may be solved.

64.3.1 The Principle of PDR

The main flow of PDR algorithm is:

1. Use accelerometer to measure the acceleration when a pedestrian walks, and check if the pedestrian has stridden a step by comparing the acceleration and the known threshold.
2. If the pedestrian has taken a step, then the acceleration during this step will be used to calculate the step length of this step.
3. Gyroscope and magnetometer will be used to estimate the walking direction of the pedestrian.
4. Since the walking distance and direction are known, the next position of the pedestrian can be calculated.

Therefore, the iteration formula calculating the pedestrian position is as:

$$\begin{cases} x_{k+1} = x_k + l_k * \cos \theta_k \\ y_{k+1} = y_k + l_k * \sin \theta_k \end{cases} \quad (64.4)$$

where (x_k, y_k) is the position of the pedestrian at time k , l_k and θ_k are the step length and walking direction or called heading respectively, (x_{k+1}, y_{k+1}) is the position at time $k + 1$ after the pedestrian takes this step.

64.3.2 Motion Recognition Assist

The PDR method mainly includes three key technologies: step detection, step length estimation and heading estimation [8, 9]. We use motion recognition to assist pedestrian dead reckoning in the following aspects.

64.3.2.1 Step Length Estimation Model

There are several models which estimate the step length [5, 10, 11]. After the model is trained well, it keeps invariable. The step length estimation model we use is:

$$l = a \cdot freq + b \cdot var + c \quad (64.5)$$

where $freq$ is the step frequency which is the reciprocal of the step duration, var is the variance of the module of acceleration during this step, and a , b , c are the parameters.

In our study, we notice that the step length is smaller than the normal walking when a pedestrian is turning. Therefore, the accumulated distance will be larger than the real one if using the trained model with fix parameters. It is obviously that

the indoor positioning error can be decreased if we adapt the model parameters when a pedestrian is turning. Thus, we propose a novel step length model as:

$$l = a(s) \cdot freq + b(s) \cdot var + c(s) \quad (64.6)$$

where s is the motion state of the pedestrian, which means that the step length is adaptive according to the motion state a pedestrian is performing.

64.3.2.2 3-D Pedestrian Localization

The PDR we discussed above is a 2-D localization system, and it doesn't pay attention to the altitude change of a pedestrian. Therefore, when it comes to the stairs situation, the above PDR system will fail. However, by recognizing the motion state of a pedestrian, we propose a new dead reckoning formula to infer the pedestrian's position:

$$\begin{cases} x_{k+1} = x_k + l_{hk} * \cos \theta_{hk} \\ y_{k+1} = y_k + l_{hk} * \sin \theta_{hk} \\ z_{k+1} = z_k + l_{vk} \end{cases} \quad (64.7)$$

where (x_k, y_k, z_k) is the position of the pedestrian at time k , l_{hk} is the horizontal displacement, l_{vk} is the vertical displacement, θ_{hk} is the heading along the horizontal direction, and $(x_{k+1}, y_{k+1}, z_{k+1})$ is the position at time $k + 1$ after the pedestrian takes this step.

64.4 Evaluation

This section first compares the classification accuracy using classic Bayes classifier and sequence-based classifier using HMM. Then a 3D pedestrian trajectory is shown to demonstrate that PDR algorithm with motion recognition assistance works well in the 3D situation.

The experiments were conducted in a typical office building. The sensors we used were imbedded in the smartphone and the sampling rate was 32 Hz. All sensor data were stored and used for offline processing.

64.4.1 Sequence-Based Motion Recognition

First, a training dataset was used to model the Bayes classifier and the sequence-based classifier using HMM. Then typical test dataset was used to evaluate the performance of them. The test dataset was acquired by walking across the stairs downward, the corridor and the office room in order.

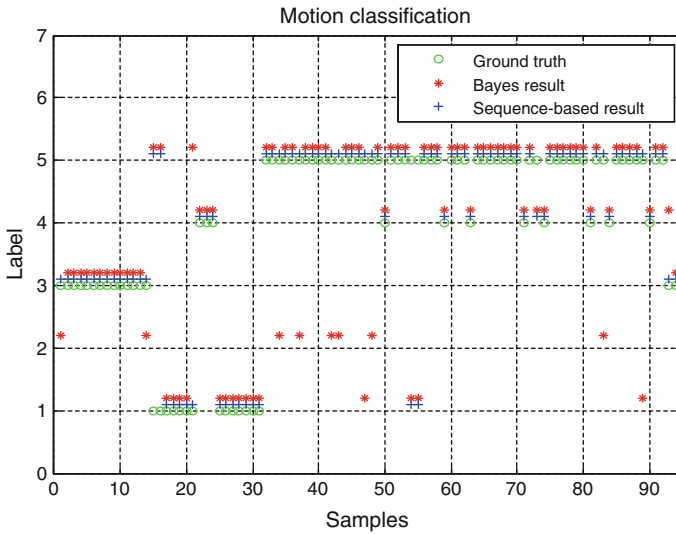


Fig. 64.1 The motion classification result of the classifiers (*label 1 downstairs, 2 upstairs, 3 still, 4 turning, 5 walking*)

Compared to the classic Bayes classifier, since sequence-based motion recognition classifier uses the HMM to construct the relation between samples, it can remove some false states embedded in the sample sequence. As shown in Fig. 64.1, most of the upstairs states are corrected to the walking states by applying our proposed method. The accuracy using Bayes classifier is 82.11 %, and the accuracy of sequence-based classifier using HMM is 94.74 %.

64.4.2 Indoor Pedestrian Localization

In order to evaluate the localization performance of our proposed solution indoors, we executed a comprehensive experiment. In this experiment, a pedestrian first walked around in an office with large open space, then, stepped out of the office by walking through a corridor. After that, he went upstairs and stopped on the upper floor.

Figure 64.2 shows the trajectory estimated from our proposed PDR solution which is in accordance with the trace the pedestrian really passed. The motion state is precisely detected when the pedestrian is going upstairs. Therefore, the vertical dimension of the trajectory is extended and the 3D localization for a pedestrian is performed. The detailed trajectory in staircase is shown in Fig. 64.3. The displacement in vertical dimension is 2.89 m using motion recognition assisted PDR. Comparing ground truth of 3.95 m, the accumulated error is 1.06 m. Although it's not very accurate, it is enough to distinguish the floor level from the height deviation.

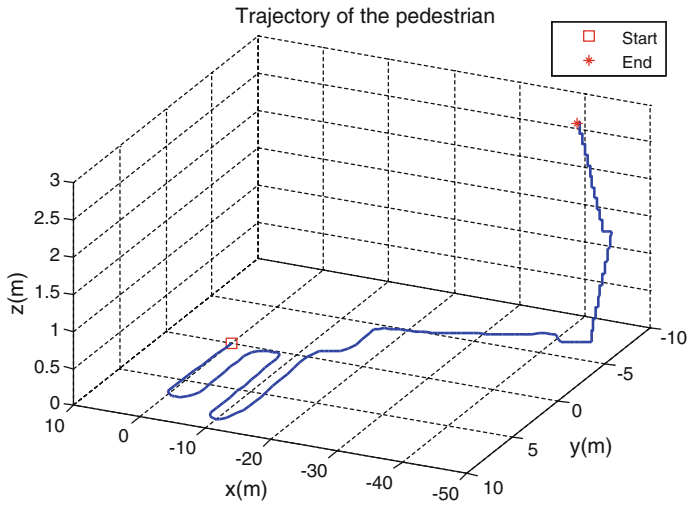


Fig. 64.2 Three-dimensional indoor pedestrian localization

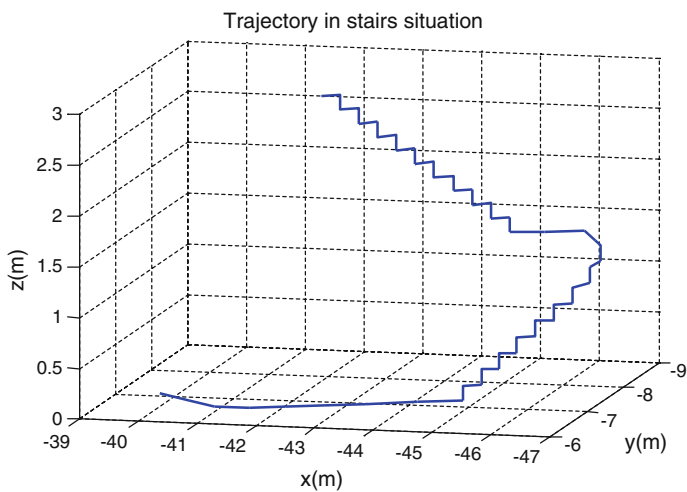


Fig. 64.3 The trajectory of the pedestrian in the staircase

Figure 64.4 shows the horizontal trajectory of the pedestrian in the office. The black square line is the ground truth. Red dot line indicates the trajectory with conventional PDR. The solid green line presents the result of our proposed method (MR-PDR). The horizontal error of trajectory in the office is listed in Table 64.4. It shows that errors are reduced significantly using the motion recognition assisted pedestrian dead reckoning.

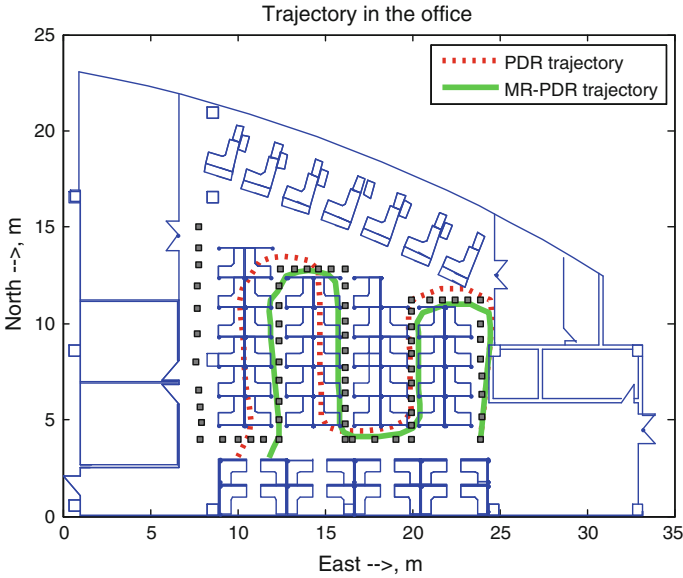


Fig. 64.4 The trajectory of the pedestrian in the office

Table 64.4 The horizontal localization error in an office

Situation	Mean error (m)	50th percentile error (m)	95th percentile error (m)
PDR	0.78	0.60	2.08
MR-PDR	0.30	0.34	0.53

64.5 Conclusion

This paper first describes the sequence-based motion recognition using HMM, then proposes the pedestrian dead reckoning method based on motion recognition. Experiment results demonstrate that the sequence-based motion classifier is more accurate than a classic classifier. And the proposed PDR based on motion recognition performs better than the conventional PDR. What’s more, with the knowledge of motion states, PDR also works in the 3D environments.

Acknowledgments This work is jointly funded by the China Satellite Navigation Office and the Science and Technology Commission of Shanghai Municipality. The funding project number is BDZX005. This work is also supported by funds from Shanghai Science and Technology Commission of Shanghai Municipality under Grant 12511501002, 13511501302, 14511100300, and partly sponsored by Shanghai Pujiang Program (No. 14PJ1405000) and the research fund of ZTE.

References

1. Sadoun B, Al-Bayari O (2007) LBS and GIS technology combination and applications. In: IEEE/ACS international conference on computer systems and applications, AICCSA'07, IEEE, pp 578–583
2. Beauregard S, Haas H (2006) Pedestrian dead reckoning: a basis for personal positioning. In: Proceedings of the 3rd workshop on positioning, navigation and communication, pp 27–35
3. Qian J, Pei L, Ying R, Chen X, Zou D, Liu P, Yu W (2014). Continuous motion recognition for natural pedestrian dead reckoning using smartphone sensors. In: Proceedings of the 27th international technical meeting of the satellite division of the institute of navigation (ION GNSS + 2014), pp 1796–1801
4. Bao L, Intille SS (2004) Activity recognition from user-annotated acceleration data. In: Pervasive computing, Springer Berlin Heidelberg, pp 1–17
5. Chen R, Pei L, Chen Y (2011) A smart phone based PDR solution for indoor navigation. In: Proceedings of the 24th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2011), pp 1404–1408
6. Pei L, Liu J, Guinness R, Chen Y, Kuusniemi H, Chen R (2012) Using LS-SVM based motion recognition for smartphone indoor wireless positioning. *Sensors* 12(5):6155–6175
7. Rabiner L, Juang BH (1986) An introduction to hidden Markov models. *ASSP Mag IEEE* 3(1):4–16
8. Qian J, Ma J, Ying R, Liu P (2013) RPNOS: Reliable pedestrian navigation on a smartphone. In: *Geo-informatics in resource management and sustainable ecosystem*, Springer Berlin Heidelberg, pp 188–199
9. Qian J, Ma J, Ying R, Liu P, Pei L (2013) An improved indoor localization method using smartphone inertial sensors. In: 2013 international conference on indoor positioning and indoor navigation (IPIN), IEEE, pp 1–7
10. Alvarez D, González RC, López A, Alvarez JC (2006) Comparison of step length estimators from wearable accelerometer devices. In: 28th annual international conference of the IEEE engineering in medicine and biology society, EMBS'06, IEEE, pp 5964–5967
11. Shin SH, Park CG, Kim JW, Hong HS, Lee JM (2007) Adaptive step length estimation algorithm using low-cost MEMS inertial sensors. In: IEEE sensors applications symposium, SAS'07, IEEE, pp 1–5

Chapter 65

An Unconventional Full Tightly-Coupled Multi-Sensor Integration for Kinematic Positioning and Navigation

Jian-Guo Wang, Kun Qian and Baoxin Hu

Abstract Conventionally, all of the sensors, except the IMUs, function as aiding sensors in the multisensor integrated kinematic positioning and navigation. In this way, the IMU measurements are only used in free inertial navigation calculation, not through measurement update in Kalman filter (KF) between two adjacent aiding measurement epochs. This paper strives for a novel structure of IMU/GNSS integration KF, which deploys a kinematic trajectory model as the core of the KF system model and utilizes all of the measurements, inclusive of the ones from IMUs, through measurement updates. This novel multisensor integration strategy takes advantage of modern computing technology and well advances the realization of Kaman filter for a better utilization of the IMU measurements, especially either with low-cost IMUs or in poor GNSS and/or GNSS denied environment. Moreover, one no longer needs to distinguish between the core sensors and the aiding sensors. The conceptual comparison with the conventional error-state and error measurement based inertial navigation integration shows its advantages. The results from real road tests along with discussions are also presented.

Keywords Unconventional · Kinematics · Multisensor · IMU GNSS · Tightly-coupled integration · Kalman filter · IMU measurement update

J.-G. Wang (✉) · K. Qian · B. Hu
Department of ESSE, Lassonde School of Engineering, York University,
Toronto, Canada
e-mail: jgwang@yorku.ca

K. Qian
e-mail: qiankun@yorku.ca

B. Hu
e-mail: baoxin@yorku.ca

65.1 Introduction

In multisensor-aided inertial integrated kinematic positioning and navigation, there are four different integration architectures: uncoupled, loosely-coupled, tightly-coupled and deeply-coupled ([4], etc.). With either integration architecture, the key lies on how to realize the Kalman filter (KF) for a better measurement fusion. There are so-called direct and indirect methods to realize the KF [1, 4, 6, 7, 17], etc.). In the realization of their KF, the former uses the whole-value states (navigation parameters) while the latter estimates the error states.

An indirect KF is usually constructed in multisensor-aided inertial navigation. It utilizes the inertial navigation mechanization to estimate the error states and sensor systematic errors through error measurements on the basis of the aiding sensors ([10, 17], etc.). All of the sensors, except the IMUs, function as aiding sensors. At the same time, the direct KF has also been applied once in a while [6, 7, 16, 17]. Particularly, Wagner and Wieneke [11] presented an innovative compact integration mechanism with the pleasing solution quality by directly estimating the optimal navigation parameters where the inertial navigation mechanization was merged into the equivalent system equations.

Besides, with the efforts to integrate low-cost IMUs in variety of kinematic positioning and navigation systems, the focus has obviously been on the adaption of the existing conventional integration strategy and further reinforcement, e.g., by applying extra parametric constraints for specific motions such as the quasi-zero lateral and downward velocities [3], etc.) and/or the self-contained virtual measurements based on *a priori* error characteristics of IMUs and the vehicle's dynamics (Wang and Gao [12], etc.) and so on.

However, the system models used in KF from all abovementioned work have directly been based on the *a priori* inertial error model. In other words, there is principally no difference in any case as the IMU measurements are applied only in the free inertial navigation calculation and thus no measurement update is performed in KF between two adjacent aiding measurement epochs.

What can be done with the same multisensor integration task to take advantages of modern computing technology and how can the KF be realized for a better utilization of IMU measurements, especially with low-cost IMUs and/or in poor GNSS and/or GNSS denied environment? Thus, a significant breakthrough with the multisensor integration strategy is highly demanded for low-cost IMU-based kinematic positioning and navigation systems.

Accordingly, a number of research trials have been made to construct standalone system model so that all of the sensor measurements, inclusive of IMU raw outputs, can directly participate in KF measurement update. Wang [15] developed the system models based on the uniform linear motion, uniform circular motion to construct a direct KF. Variant kinematic models were also applied in KF for kinematic positioning and navigation [1, 18]. Especially, Wang and Sternberg [14] tentatively utilized the 3D kinematic model after the uniform circular motion as the system model in the KF with the six whole-value states and the nine measurements

for the GPS-aided IMU integrated navigation. Along the same path, the authors have been working on constructing novel IMU/GPS navigation KF (IMU/GPS KF) for multisensor integrated navigation systems [9, 13].

This paper continually strives for a novel IMU/GNSS Kalman filter, which utilizes a 3D kinematic trajectory model as the core of the KF system model and allows the measurements from all of the sensors, inclusive of the IMUs, directly participating in KF measurement updates instead of using the error measurements. The proposed IMU/GNSS KF estimates the whole-value states instead of the error states. Since IMU's raw outputs directly participate in KF measurement update as the raw observables from all other kinds of sensors (e.g. the GNSS pseudo-range, carrier phase and Doppler measurements etc.) do, there is no longer need to distinguish between the core sensors and the aiding sensors. This novel KF structure makes the full tightly-coupled multisensor integration for kinematic positioning and navigation come true. This introduction will be followed by the description of the novel KF algorithm for multisensor integrated navigation as Sect. 65.2. In Sect. 65.3, the results and analysis of land vehicle tests are presented. Conclusions and remarks as Sect. 65.4 end the paper.

65.2 Kalman Filter for Unconventional Full Tightly-Coupled Multisensor Integration

65.2.1 The System Model

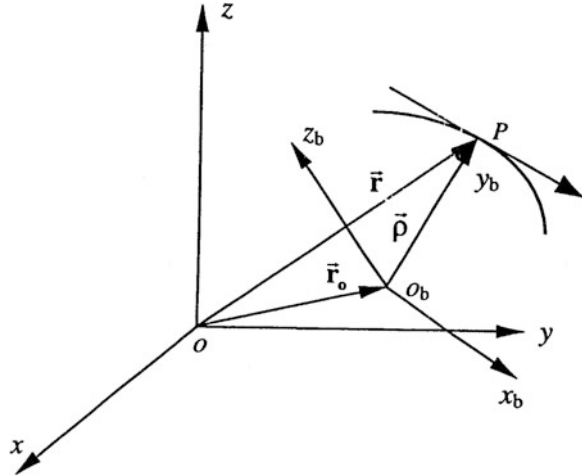
The proposed novel KF system model consists of three parts: (1) the 3D kinematic trajectory, i.e., the position, velocity, acceleration and even the jerk vectors; (2) the attitudes and (3) the angular velocity.

The instantaneous movement of an object, without considering the forces for the reasons for different types of motions, can be described using kinematics. The form for mathematic representation of position, velocity and acceleration of a mechanic system involves coordinate systems in general. Two reference systems, which are moved relatively to each other, are needed (Fig. 65.1). One is called the space-fixed system $S(o-xyz)$ while the other is referred to as the moving system $S_b(o_b-x_b y_b z_b)$. The position vector \vec{r} of a point P using the unit vectors: \vec{i}_x , \vec{i}_y and \vec{i}_z in the directions: x , y and z is given by

$$\vec{r} = x\vec{i}_x + y\vec{i}_y + z\vec{i}_z \quad (65.1)$$

and in components for the instant position of the moving point at time t with respect to a reference instant t_0

Fig. 65.1 Coordinate systems



$$\begin{aligned}
 x(t) &= x(t_0) + \dot{x}(t_0)(t - t_0) + \frac{1}{2} \ddot{x}(t_0)(t - t_0)^2 + \frac{1}{6} \overset{\dots\dots\dots}{\dddot{x}}(t_0)(t - t_0)^3 + \dots \\
 y(t) &= y(t_0) + \dot{y}(t_0)(t - t_0) + \frac{1}{2} \ddot{y}(t_0)(t - t_0)^2 + \frac{1}{6} \overset{\dots\dots\dots}{\dddot{y}}(t_0)(t - t_0)^3 + \dots \\
 z(t) &= z(t_0) + \dot{z}(t_0)(t - t_0) + \frac{1}{2} \ddot{z}(t_0)(t - t_0)^2 + \frac{1}{6} \overset{\dots\dots\dots}{\dddot{z}}(t_0)(t - t_0)^3 + \dots \quad (65.2)
 \end{aligned}$$

wherein $\dot{x}, \dot{y}, \dot{z}, \dots, \overset{\dots\dots\dots}{\dddot{x}}, \overset{\dots\dots\dots}{\dddot{y}}, \overset{\dots\dots\dots}{\dddot{z}}$ are the 1st, 2nd and 3rd derivatives of x, y, z with respect to t . In order to be able to take as less terms as possible, the time interval $t - t_0$ must be small enough. Let \vec{r}_o be the position vector of the origin O_b and \vec{p} and \vec{p}_b be the relative position vector of point P from O_b and the same vector in S_b , then

$$\vec{r} = \vec{r}_o + \vec{p} = \vec{r}_o + \mathbf{D}^T(t) \vec{p}_b \quad (65.3)$$

where in $\mathbf{D}(t)$ is the instantaneous rotation matrix from S to S_b . From (65.3), the velocity, acceleration and jerk vectors can be derived in general. The angular motion is coupled with $\mathbf{D}(t)$. For more about the relevant derivation in details refer to ([2, 14]; etc.).

With a strapdown inertial navigation system, the basic navigation parameters of interest are the position vector r_{nb}^n of the IMU center and three Euler angles (roll γ , pitch p , and heading α) for the rotations between the IMU body frame b and the local navigation frame n . Without further unnecessary statements on kinematics, the velocity vector v_{nb}^n , transformed from its counterpart velocity vector v_{nb}^b in the IMU body frame, the acceleration vector a_{nb}^n and the jerk vector j_{nb}^n are directly given below [9]

$$\dot{\mathbf{j}}_{nb}^n = \mathbf{v}_{nb}^n = \mathbf{C}_b^n \mathbf{v}_{nb}^b \quad (65.4)$$

$$\dot{\mathbf{v}}_{nb}^n = \mathbf{a}_{nb}^n = \mathbf{C}_b^n [\boldsymbol{\omega}_{nb}^b \times] \mathbf{v}_{nb}^b + \mathbf{C}_b^n \begin{pmatrix} \dot{v}_{nbx}^b & \dot{v}_{nby}^b & \dot{v}_{nbz}^b \end{pmatrix}^T = \mathbf{C}_b^n \mathbf{a}_{nb}^b \quad (65.5)$$

$$\dot{\mathbf{a}}_{nb}^n = \mathbf{j}_{nb}^n = \mathbf{C}_b^n [\boldsymbol{\omega}_{nb}^b \times] \mathbf{a}_{nb}^b + \mathbf{C}_b^n \begin{pmatrix} \dot{a}_{nbx}^b & \dot{a}_{nby}^b & \dot{a}_{nbz}^b \end{pmatrix}^T = \mathbf{C}_b^n \mathbf{j}_{nb}^b \quad (65.6)$$

wherein \mathbf{v}_{nb}^b , \mathbf{a}_{nb}^b , \mathbf{j}_{nb}^b , $\boldsymbol{\omega}_{nb}^b$ are the velocity, acceleration, jerk and angular velocity vector in the IMU body frame, respectively, and \mathbf{C}_b^n is the direction cosine matrix (DCM) from the body frame to the navigation frame.

The attitude angles (roll γ , pitch p and heading α) are the functions of the angular velocities which vary with the particular realization of the Euler angles. Under its realization of Kardan angles used in strapdown inertial navigation, they are connected to each other as follows [5]

$$\begin{pmatrix} \dot{\gamma} \\ \dot{p} \\ \dot{\alpha} \end{pmatrix} = \begin{pmatrix} \sin \gamma \tan p & 1 & -\cos \gamma \tan p \\ \cos \gamma & 0 & \sin \gamma \\ \sin \gamma \sec p & 0 & -\cos \gamma \sec p \end{pmatrix} \begin{pmatrix} \omega_{nbx}^b \\ \omega_{nby}^b \\ \omega_{nbz}^b \end{pmatrix} = \mathbf{C}_{3 \times 3} \boldsymbol{\omega}_{nb}^b \quad (65.7)$$

Under the normal driving condition with a smooth steering over a small time interval, three components of the angular velocity of a land vehicle, for instance, can reasonably be treated as independent [9]. So, ω_{nbx}^n and ω_{nby}^n are modeled as zero-mean processes and ω_{nbz}^n is modeled as a non-zero mean random process with random disturbance. The former is usually interpreted by the first order Markov model. Here, the zero-mean Singer motion model is used to express the dynamic variation of two quasi-zero angular rate components ω_{nbx}^n and ω_{nby}^n in the system model, while the modified Singer model is adapted to describe the dynamics of the angular rate ω_{nbz}^n [19].

In practice, the sensor systematic errors, e.g. the IMU biases and scale factor errors and so forth may also be included in the KF system model according to the need and possibility. In summary, the state vector here can specifically consists of 27 components as follows

$$\mathbf{x} = \left(\mathbf{r}^T \quad (\mathbf{v}_{nb}^b)^T \quad (\mathbf{a}_{nb}^b)^T \quad \boldsymbol{\theta}^T \quad (\boldsymbol{\omega}_{nb}^b)^T \quad \mathbf{b}_g^T \quad \mathbf{b}_a^T \quad \mathbf{s}_g^T \quad \mathbf{s}_a^T \right)^T \quad (65.8)$$

with $\mathbf{r} = (X \ Y \ Z)^T$, $\mathbf{v}_{nb}^b = (v_{nbx}^b \ v_{nby}^b \ v_{nbz}^b)^T$, $\mathbf{a}_{nb}^b = (a_{nbx}^b \ a_{nby}^b \ a_{nbz}^b)^T$, $\boldsymbol{\theta} = (\gamma \ p \ \alpha)^T$, $\boldsymbol{\omega}_{nb}^b = (\omega_{nbx}^b \ \omega_{nby}^b \ \omega_{nbz}^b)^T$, $\mathbf{b}_g = (b_{gx} \ b_{gy} \ b_{gz})^T$, $\mathbf{s}_g = (s_{gx} \ s_{gy} \ s_{gz})^T$, $\mathbf{b}_a = (b_{ax} \ b_{ay} \ b_{az})^T$, $\mathbf{s}_a = (s_{ax} \ s_{ay} \ s_{az})^T$, which are the position, body velocity, body acceleration, the attitude, body angular velocity, gyro bias and scale factor error, accelerometer bias and scale factor error vectors, respectively.

The proposed system model in discrete time is summarized below

$$\mathbf{r}_{k+1} = \mathbf{r}_k + \Delta t \mathbf{C}_{b(k)}^n \mathbf{v}_{nb(k)}^b \mathbf{v}_{nb(k)}^b + \frac{\Delta t^2}{2} \mathbf{C}_{b(k)}^n \mathbf{a}_{nb(k)}^b + \frac{\Delta t^3}{6} \mathbf{C}_{b(k)}^n \mathbf{j}_{nb(k)}^b \quad (65.9)$$

$$\begin{aligned} \mathbf{v}_{nb(k+1)}^b = & \left[\mathbf{I}_{3 \times 3} - \Delta t \left[\boldsymbol{\omega}_{nb(k)}^b \times \right] + \frac{\Delta t^2}{2} \left[\boldsymbol{\omega}_{nb(k)}^b \times \right]^2 \right] \mathbf{v}_{nb(k)}^b + \left[\Delta t \mathbf{I}_{3 \times 3} \right. \\ & \left. - \Delta t^2 \left[\boldsymbol{\omega}_{nb(k)}^b \times \right] \right] \mathbf{a}_{nb(k)}^b + \frac{\Delta t^2}{2} \left[\mathbf{v}_{nb(k)}^b \times \right] \dot{\boldsymbol{\omega}}_{nb(k)}^b + \frac{\Delta t^2}{2} \mathbf{j}_{nb(k)}^b \end{aligned} \quad (65.10)$$

$$\begin{aligned} \mathbf{a}_{nb(k+1)}^b = & \left[\mathbf{I}_{3 \times 3} - \Delta t \left[\boldsymbol{\omega}_{nb(k)}^b \times \right] + \frac{\Delta t^2}{2} \left[\boldsymbol{\omega}_{nb(k)}^b \times \right]^2 \right] \mathbf{a}_{nb(k)}^b \\ & + \frac{\Delta t^2}{2} \left[\mathbf{a}_{nb(k)}^b \times \right] \dot{\boldsymbol{\omega}}_{nb}^b + \Delta t \mathbf{j}_{nb}^b \end{aligned} \quad (65.11)$$

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \Delta t \mathbf{C}_{3 \times 3} \boldsymbol{\omega}_{nb(k)}^b + \frac{\Delta t^2}{2} \mathbf{C}_{3 \times 3} \dot{\boldsymbol{\omega}}_{nb}^b \quad (65.12)$$

$$\omega_{nbx(k+1)}^b = e^{-\Delta t/T_x} \omega_{nbx(k)}^b + w_{\omega_x} \quad (65.13)$$

$$\omega_{nby(k+1)}^b = e^{-\Delta t/T_y} \omega_{nby(k)}^b + w_{\omega_y} \quad (65.14)$$

$$\omega_{nbz(k+1)}^b = e^{-\Delta t/T_z} \omega_{nbz(k)}^b + (1 - e^{-\Delta t/T_z}) \omega_{nbz(k-1)}^b + w_{\omega_z} \quad (65.15)$$

$$\mathbf{b}_{g(k+1)} = \mathbf{b}_{g(k)} + \mathbf{w}_{b_g} \quad (65.16)$$

$$\mathbf{b}_{a(k+1)} = \mathbf{b}_{a(k)} + \mathbf{w}_{b_a} \quad (65.17)$$

$$\mathbf{s}_{g(k+1)} = \mathbf{s}_{g(k)} + \mathbf{w}_{s_g} \quad (65.18)$$

$$\mathbf{s}_{a(k+1)} = \mathbf{s}_{a(k)} + \mathbf{w}_{s_a} \quad (65.19)$$

where $\Delta t = t_{k+1} - t_k$, \mathbf{C}_b^n is the same as in (65.4)–(65.6), T_x , T_y , T_z are the time correlation coefficients, w_x , w_y , w_z are the white noises for the angular velocities, \mathbf{w}_{b_g} , \mathbf{w}_{s_g} , \mathbf{w}_{b_a} and \mathbf{w}_{s_a} are the white noise vectors for the biases and scale factor errors of gyros and accelerometers, \mathbf{j}_{nb}^b is the body jerk vector as the process noise for the position, velocity and acceleration vectors, and $\dot{\boldsymbol{\omega}}_{nb}^b$ is the change of the angular velocity vector.

Multiple alternate models can be derived from certain simplification of (65.9)–(65.12) while one may extend the model for sensor systematic errors as in (65.13)–(65.19), especially with the IMU sensors.

65.2.2 The Measurement Model

Only the IMU and GPS measurements are here considered.

In general, an IMU sensor consists of three gyroscopes to measure the angular rate vector ω_{ib}^b and three accelerometers to measure the specific force vector f_{ib}^b . In practice, the measurements from inertial sensors may be contaminated by the following errors: constant start-off biases, drifting bias residuals, errors of linear scale factors, misalignments of sensor axes and Gaussian white noises and so on, which depend on the quality of a specific IMU. As an experimental decision here in this paper, the IMU measurements for the being constructed IMU/GNSS KF are modeled as follows:

$$\omega_{ib-imu}^b = (\mathbf{I} + \mathbf{S}_g)(\omega_{nb}^b + \mathbf{C}_n^b(\omega_{i.e.}^n + \omega_{en}^n)) + \mathbf{b}_g + \Delta_g \quad (65.20)$$

$$f_{ib-imu}^b = (\mathbf{I} + \mathbf{S}_a)(\mathbf{a}_{nb}^b + \mathbf{C}_n^b(2\omega_{ie}^n + \omega_{en}^n) \times \mathbf{v}_{nb}^n - \mathbf{C}_n^b \mathbf{g}^n) + \mathbf{b}_a + \Delta_a \quad (65.21)$$

where in $\omega_{i.e.}^n$, ω_{en}^n and \mathbf{g}^n are the Earth's rotation rate vector and transport rate vector and the local gravity vector in the local ENU frame, ω_{nb}^n , \mathbf{a}_{nb}^n are the rotation rate vector the acceleration vector of IMU body frame with respect to the local ENU frame, \mathbf{S}_g , \mathbf{S}_a are the 3×3 scalar and misalignment error matrices for three gyroscopes and three accelerometers, Δ_g , Δ_a are the Gaussian white noise vectors for gyroscope and accelerometer measurements, respectively. With the low-cost IMUs, depending on specific needs, (65.20) and (65.21) may be further simplified to:

$$\omega_{ib-imu}^b = (\mathbf{I} + \mathbf{S}_g)\omega_{nb}^b + \mathbf{b}_g + \Delta_g \quad (65.22)$$

$$f_{ib-imu}^b = (\mathbf{I} + \mathbf{S}_a)(\mathbf{a}_{nb}^b - \mathbf{C}_n^b \mathbf{g}^n) + \mathbf{b}_a + \Delta_a \quad (65.23)$$

With the GNSS, here specifically the GPS, the double differenced pseudoranges and carrier phases are used for the accuracy level of centimeters. The generic observation equations for the pseudorange PR_i^j and the carrier phases ϕ_i^j (only L1 phases considered for demonstration purpose) from a receiver i to a satellite j are

$$PR_i^j = \rho_i^j + c(\delta t_i - \delta t^j) + d_{i-trop}^j + d_{i-ion}^j + \varepsilon_{PR_i^j} \quad (65.24)$$

$$\phi_i^j = \rho_i^j + c(\delta t_i - \delta t^j) + d_{i-trop}^j - d_{i-ion}^j + \lambda_1 N_i^j + \varepsilon_{\phi_i^j} \quad (65.25)$$

wherein $i = A, B$ (a base station and a rover station) and $j = 1, 2, \dots, n$, ρ_i^j is the geometric range from the receiver i to the satellite j , c is the speed of light, $\delta t_i, \delta t^j$ are the receiver and satellite clock errors, respectively, $d_{i-trop}^j, d_{i-ion}^j$ are the tropospheric and ionospheric delays, λ_1, N_i^j are the L1 wave length and the ambiguity parameter, and $\varepsilon_{PR_i^j}, \varepsilon_{\phi_i^j}$ are the random noises. As usual, the single differencing

performs between two GPS receivers with respect to the same satellite and further the double differencing is introduced to two single differenced measurements between two satellites. Without going into the unnecessary details, the double differenced GPS measurements, denoted by the symbol $\Delta\nabla$, are

$$\nabla\Delta PR_{AB}^{jk}(t) = \nabla\Delta\rho_{AB}^{jk}(t) + \nabla\Delta d_{AB-Iono}^{jk}(t) + \nabla\Delta d_{AB-Trop}^{jk}(t) + \varepsilon_{\nabla\Delta P_{AB}}^{jk}(t) \quad (65.26)$$

With a baseline away from the GPS base station short enough, e.g., < 10 km, the double differenced tropospheric and ionospheric delays become small enough to be ignored so that (65.26) can be simplified to

$$\nabla\Delta PR_{AB}^{jk}(t) = \nabla\Delta\rho_{AB}^{jk}(t) + \varepsilon_{\nabla\Delta P_{AB}}^{jk}(t) \quad (65.27)$$

with $\nabla\Delta\rho_{AB}^{jk} = [\rho_B^k(t) - \rho_A^k(t)] - [\rho_B^j(t) - \rho_A^j(t)]$. For the carrier phase observations, the final double differenced observation equation is directly given below

$$\nabla\Delta\varphi_{AB}^{jk}(t) = \nabla\Delta\rho_{AB}^{jk}(t) + \lambda_1\nabla\Delta N_{AB}^{jk} + \varepsilon_{AB}^{jk}(t) \quad (65.28)$$

Additionally, the DGPS heading using two GPS receivers on the vehicle is also available

$$\gamma_{DGPS}(t) = \gamma(t) + \varepsilon_{DGPS}^{\gamma}(t) \quad (65.29)$$

In particular, the state vector in (65.8) must be extended to include the ambiguity parameters $\nabla\Delta N_{AB}^{jk}$ before they can be fixed as integer numbers.

Similarly, measurement equations for other additional sensors can be practically developed. So, the measurements from all of the integrated sensors can directly participate in KF measurement updates under the proposed integration strategy.

65.2.3 The Kalman Filter Under the Novel Integration Strategy

Straightforward, a KF algorithm is further constructed on the basis of the proposed system model given in (65.9)–(65.19) for time updates starting from the corresponding initial state vector, and the measurement model given in (65.20), (65.21) (or (65.22) and (65.23)), (65.27) and (65.29) for measurement updates. Because most of the above mentioned equations are nonlinear, the extended KF (EKF) algorithm is applied. However, no further detail about it will be provided here due to the space limit and also because the EKF development is not the focus of this manuscript.

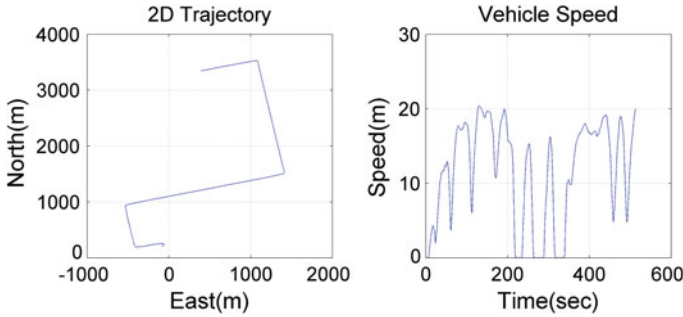


Fig. 65.2 Trajectory's top view and the velocity profile

65.3 Road Tests and Results

The proposed system and measurement models were successfully implemented using the EKF in C/C++ under Microsoft Visual Studio. A number of the road tests were performed by our in-house developed land vehicle navigation system with two NovAtel OEM GPS receivers, and one Crossbow IMU440CA (angular rate: bias stability $<10.0^\circ/h$, angle random walk $<4.5^\circ/\sqrt{hr}$; acceleration: bias Stability $<1.0\text{ mg}$, velocity random walk $<1.0\text{ m/s}/\sqrt{hr}$) with the third GPS receiver as the base station [8]. This section gives the results from one of our road tests. The lever arm vectors of the GPS receivers and cameras with respect to the IMU unit were measured beforehand at the accuracy of 0.5 cm. The data rates were set to 1.0 Hz, 4 and 100 Hz for the GPS base station and rover receiver and the Crossbow IMU, respectively. The used dataset was collected in Vaughan, Ontario, of which the first part for 518 s is used here for demonstration purpose. Figure 65.2 gives the top view of the trajectory and velocity profile of the vehicle.

Due to the space limit, only the selected plots are given here. The 1σ accuracy plots for position (Fig. 65.3), velocity (Fig. 65.4) and the attitude solution with 3σ envelopes and the 1-sigma attitude accuracies (Fig. 65.5) are presented. Figure 65.6

Fig. 65.3 3D positional accuracy

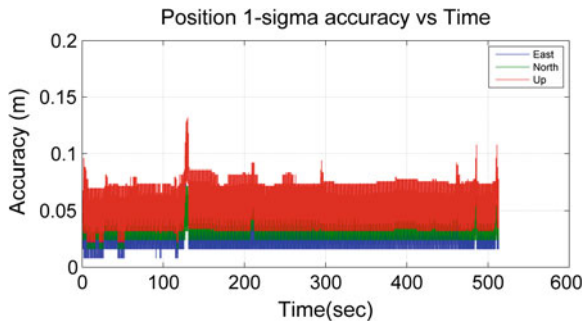


Fig. 65.4 3D velocity accuracy

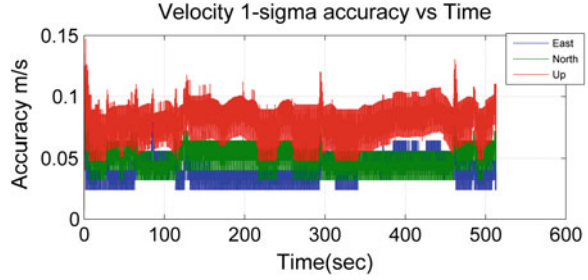
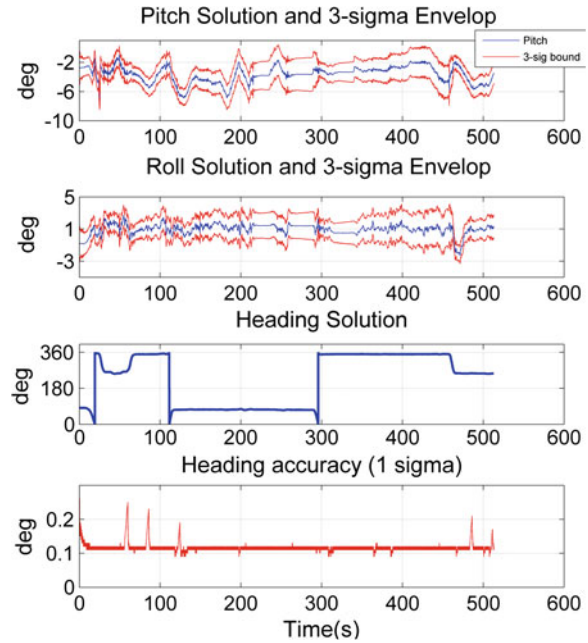


Fig. 65.5 Attitude solution and accuracy



is the number of the available satellites followed by the innovation and residual series of the double differenced carrier phases from PRN15, 18 and 22 in Fig. 65.7.

Furthermore, the estimates of the bias states with their $\pm 3\sigma$ bounds from IMU/GNSS KF are shown in Fig. 65.8. The accuracies of accelerometer biases in right and forward directions are noticeably improved while the vehicle was turning because the observabilities of bias drifts are enhanced in that case.

In order to assess the performance improvement with the proposed models through the EKF, GPS outages were simulated. The solutions from the proposed novel integration strategy (solution with IMU measurement updates in KF) and the conventional integration strategy (typical free inertial solution) were compared here. Figure 65.9 presents the solution with one 20-second GNSS outage against the reference solution. At the beginning of the GNSS outage, the navigation parameters

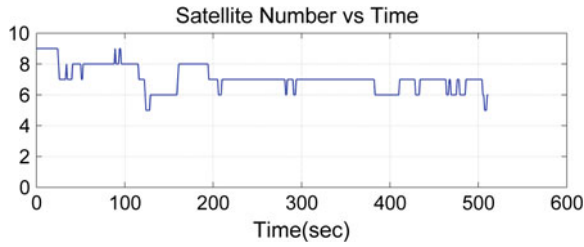


Fig. 65.6 The satellite availability

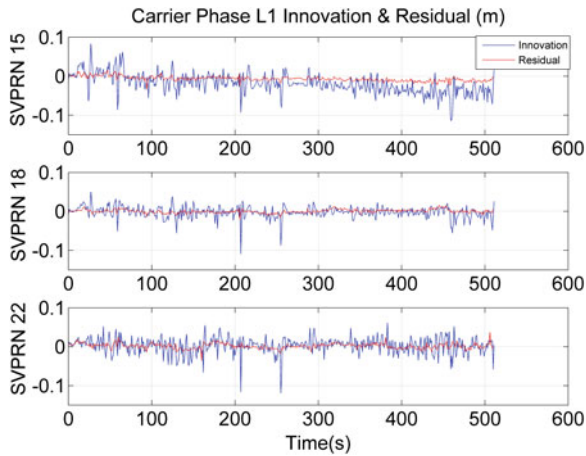


Fig. 65.7 Innovation and residual of GPS satellites (SVPRN 15, 18, 22)

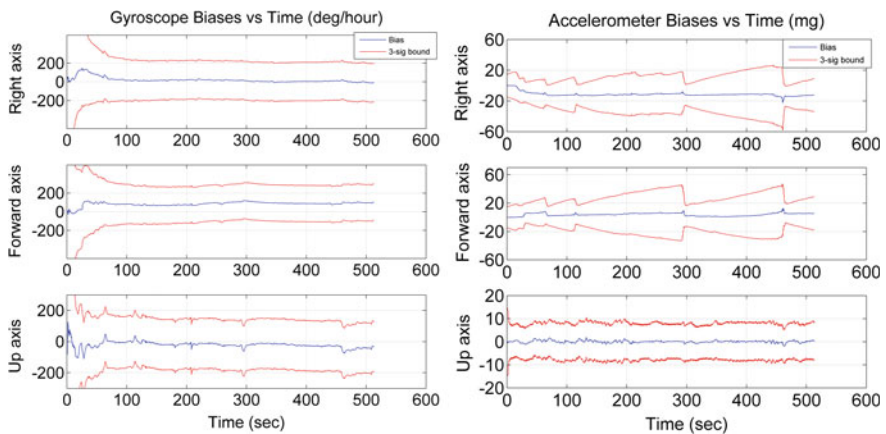
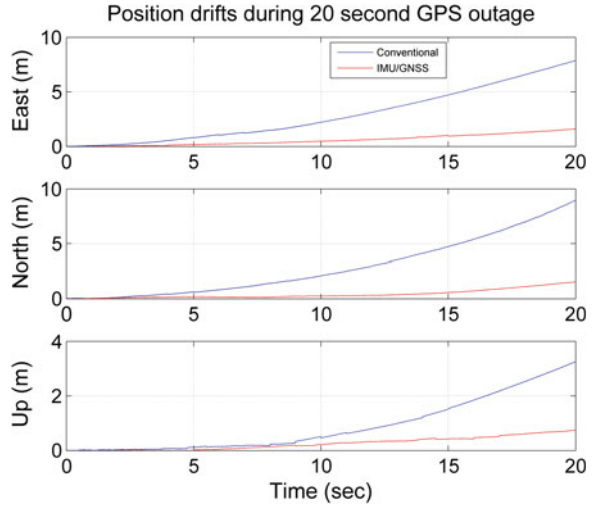


Fig. 65.8 Biases and their $\pm 3\sigma$ bounds from gyroscopes (left) and accelerometers (right)

Fig. 65.9 Position drifts of the novel KF and the conventional free inertial navigation



for both of the integration strategies were initialized to the same integrated navigation solution. Clearly, the positions from the novel multisensor integration KF were drifting much slower than the ones from the conventional free inertial navigation solution.

65.4 Conclusions and Potential Future Research

A novel structure of KF has been proposed for multisensor integrated kinematic positioning and navigation, which applies the kinematics to build up the core of the system model. By using the direct method, all of the measurements from individual sensors can directly participate in the measurement updates of Kalman filter independently so that a novel full multisensor integration strategy comes true as there is no longer need to distinguish between the core sensors and other aiding sensors. The solution from one of our road tests has been taken to have demonstrated the success of the proposed novel multisensor integration strategy with the improved solution performance, and robustness with the low-cost IMU sensors and especially in poor GNSS and/or GNSS denied environment. The next tasks could be on developing the simultaneous estimation of variances for process and measurement noises and systematically studying the effects of different choices of states under the novel multisensor integration strategy to further explore its advantages for low-cost multisensor integrated kinematic positioning and navigation.

Acknowledgments The authors sincerely acknowledge the financial support through research grant provided by Natural Sciences and Engineering Council (NSERC) of Canada.

References

1. Beetz A (2012) Ein modulares Simulationskonzept zur Evaluierung von Positions-sensoren sowie Filter- und Regelalgorithmen am Beispiel des automatisierten Straßenbaus, No. 688, DGK Series C, Munich. ISSN 0065-5325
2. D'Souza AF, Garg VK (1984) *Advanced dynamics: modeling and analysis*. Prentice-Hall Inc., Englewood Cliffs
3. Godha S (2006) Performance evaluation of low cost MEMS-based IMU integrated with GPS for land vehicle navigation application. M.Sc. thesis, UCGE Reports, No. 20239, Department of Geomatics Engineering of University of Calgary
4. Greenspan RL (1996) GPS and inertial integration, chapter 7, *GPS theory and applications*, vol II, edited by Parkinson et al, Progress in Astronautics and Aeronautics, vol 164
5. Magnus Kurt (1971) *Kreisel: theorie und anwendungen*. Springer, Berlin
6. Munguía R (2014) A GPS-aided inertial navigation system in direct configuration. *J Appl Res Technol* 12:803–814
7. Qi H, Moore JB (2002) Direct kalman filtering approach for GPS/INS integration. *IEEE Trans Aerosp Electron Syst* 38(2):687–693
8. Qian K, Wang J, Gopaul N, Hu B (2012) Low cost multisensor kinematic positioning and navigation system with Linux/RTAI. *J Sens Actuator Netw* 1(3):166–182
9. Qian K, Wang J, Hu B (2013) Application of vehicle kinematic model on GPS/ MEMS IMU integration. In: Joint EOGC 2013 and CIG annual conference, Toronto, June 5–7, 2013
10. Titterton D, Weston J (2004) *Strapdown inertial technology*, 2nd edn. Co-published by the Institution of Engineering and Technology, London and The American Institute of Aeronautics, Reston
11. Wagner JF, Wieneke T (2003) Integrating satellite and inertial navigation—conventional and new fusion approaches. *Control Eng Pract* 11(5):543–550
12. Wang JH, Gao Y (2010) Land vehicle dynamics-aided inertial navigation. *IEEE Trans Aerosp Electron Syst* 46(4):1638–1653
13. Wang J, Qian K, Hu B (2014) A novel and unique IMU/GNSS Kalman filter, keynote speaker of session invited presentation (J. Wang). In: China satellite navigation conference. Nanjing, May 21–23, 2014
14. Wang J, Sternberg H (2000) Model development for kinematic surveying of land vehicle trajectories (in German), *Schriftenreihe Studiengang Vermessungswesen UniBw München*, No. 60-1, Germany, pp 317–331. ISSN 0179-1009
15. Wang J (1997) Filtermethoden zur fehlertoleranten kinematischen Positionsbestimmung, *Schriftenreihe Studiengang Vermessungswesen UniBw München*. PhD dissertation, No. 52, Germany. ISSN 0179-1009
16. Wendel J, Schaile C, Trommer GF (2001) Direct kalman filtering of GPS/INS for aerospace applications. In: *International symposium on kinematic systems in geodesy, geomatics and navigation (KIS2001)*, Canada
17. Yi G, Wang H (1987) *Inertial navigation* (in Chinese). Aviation Industry Press, Beijing
18. Yu G, Xiong J, Guo H, Wang J (2014) GNSS/INS/VKM vehicle integrated navigation system. In: Sun J et al (eds). *China satellite navigation conference, proceedings: vol III. Lecture Notes in Electrical Engineering* 305, Heidelberg, pp 585–594
19. Zhou H, Kumar KSP (1984) A “current” statistical model and adaptive algorithm for estimating maneuvering targets. *J Guid Cont Dyn* 7(5):596–602

Chapter 66

A WIFI/INS Indoor Pedestrian Navigation System Augmented by Context Feature

Ling Yang, Yong Li and Chris Rizos

Abstract An Inertial navigation System (INS) is self-contained, immune to jamming/interference and in many other ways is ideally for pedestrian navigation applications especially in indoor environments. However, due to the sensor properties the quality of the navigation solution from a stand-alone INS will degrade rapidly, and must rely on some form of external correction/calibration/aiding to ensure system stability and reliability. In this contribution the authors use GPS when outdoors and WiFi when indoors to aid the INS in order to support seamless pedestrian navigation. To improve the performance of WiFi positioning, an enhanced fingerprinting method is proposed. A new fingerprinting database augmented by two types of map-based information is described. One is the topological relationship linking corridors, and the other is the orientation information of different corridors. Test results confirm that the navigation accuracy and stability is improved.

Keywords Pedestrian navigation · Step detection · Walking status detection

66.1 Introduction

A navigation system that tracks the location of a person is a useful capability for firefighters or other emergency first responders, for location-aware computing, personal navigation assistance, mobile 3D audio, augmented reality, and other

L. Yang (✉) · Y. Li (✉) · C. Rizos (✉)

School of Civil and Environmental Engineering, UNSW, Sydney NSW, 2052, Australia

e-mail: ling.yangsy@gmail.com

Y. Li

e-mail: yong.li@unsw.edu.au

C. Rizos

e-mail: c.rizos@unsw.edu.au

L. Yang

College of Surveying and Geo-Informatics, Tongji University, Shanghai, China

© Springer-Verlag Berlin Heidelberg 2015

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015*

Proceedings: Volume III, Lecture Notes in Electrical Engineering 342,

DOI 10.1007/978-3-662-46632-2_66

applications. Such a system may be referred to as a Pedestrian Navigation System (PNS). Global Navigation Satellite System (GNSS) is a well-known and extremely attractive technology for outdoor PNS applications. However GNSS-alone positioning may not satisfy the navigation integrity and reliability requirements due to its signal vulnerability to signal blocking or Radio Frequency Interference (RFI) [12]. Integration of GNSS with a low-cost MEMS-based Inertial Navigation System (INS) has gained considerable attention for consumer-based navigation applications [9, 11, 5, 6]. However, many pedestrian navigation applications involve walking from outdoors to an indoor destination point, which requires seamless navigation in both outdoor and indoor environments. Therefore, a GNSS/MEMS-INS integrated system is not entirely suitable since the navigation performance will severely degrade during long GNSS signal outages.

Over many years different technological solutions have been investigated. Nowadays the most common alternative solution is based on wireless communications infrastructure such as WiFi. The preferred class of positioning algorithms makes use of the Received Signal Strength (RSS). There are two main types of RSS-based location determination techniques, the trilateration approach and the signal fingerprinting approach. The trilateration approach uses a signal propagation model to convert the RSS into a distance measurement, and several simultaneous distance “measurements” are then used to calculate the position of the WiFi tracking device [1, 3, 9]. This technique is simple and comparatively easy to implement, however it suffers from the requirement for realistic signal propagation models. Moreover the exact locations of the WiFi access points are required. The second approach, known as “fingerprinting”, uses an empirical model to describe the distribution of RSS at various locations across the coverage area of interest [4]. This method does account for the impact of the real environment on WiFi signal propagation. However, the main disadvantage is the need to create and maintain the RSS database across the area of interest [10, 14]. Recent investigations indicate that WiFi stand-alone positioning may not be able to satisfy the stability and reliability requirements for many mission-critical applications in “busy” or dynamic environments (e.g. due to temporary pedestrian signal obstruction).

In this paper, in order to satisfy the accuracy requirement of indoor navigation, WiFi positioning technology is chosen as the most appropriate alternative for GNSS to integrate with INS. The advantages of using INS and WiFi for indoor positioning applications are their availability and relatively low cost. Two kinds of map-based information are introduced into the fingerprinting database to support the WiFi-only and WiFi/INS integrated solutions. One is the topological relationship between connected corridors. The other is the orientation information of different corridors, which is used as an attitude constraint in a WiFi/INS integrated system. With instruments mounted in a backpack, the GNSS/WiFi/INS integrated navigation system has been tested and its performance has been analysed.

66.2 System Design and Algorithm

The Extended Kalman Filter (EKF) estimation algorithm is triggered at every GNSS/WiFi measurement epoch using the difference between the GNSS/WiFi and the INS mechanisation solution (loosely-coupled or tightly-coupled). The EKF algorithm can be found in [8]. The EKF estimates the INS errors using the measurement update equations. Whenever the GNSS/WiFi measurements are missing, the EKF operates in a time prediction mode, thus resulting in errors in position, velocity and attitude estimation that grow with time.

$$\begin{cases} \mathbf{x}_{nav} = [\delta\mathbf{r}^n, \delta\mathbf{v}^n, \delta\boldsymbol{\psi}^n] \\ \mathbf{x}_{ins} = [\nabla_b, \boldsymbol{\varepsilon}_b] \end{cases} \quad (66.1)$$

where \mathbf{x}_{nav} and \mathbf{x}_{ins} are the navigation error vector and the inertial sensor measurement error vector, respectively.

The *psi*-angle error equations of the INS are expressed in the navigation frame (or local east, north and up) for position, velocity and attitude updating [7]:

$$\begin{cases} \delta\dot{\mathbf{r}}^n = -\boldsymbol{\omega}_{en}^n \times \delta\mathbf{r}^n + \delta\mathbf{v}^n \\ \delta\dot{\mathbf{v}}^n = -(\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{in}^n) \times \delta\mathbf{v}^n - \boldsymbol{\psi}^n \times \mathbf{f}^n + \delta\mathbf{g}^n + \mathbf{C}_b^n \nabla^b \\ \dot{\boldsymbol{\psi}}^n = -\boldsymbol{\omega}_{in}^n \times \boldsymbol{\psi}^n - \mathbf{C}_b^n \boldsymbol{\varepsilon}^b \end{cases} \quad (66.2)$$

where the superscript (n) refers to the navigation frame (n -frame), while the subscript (i) and (e) denote the inertial and the earth-centred earth-fixed frame (i -frame and e -frame), respectively; $\delta\mathbf{r}^n$, $\delta\mathbf{v}^n$ and $\boldsymbol{\psi}^n$ are the position, velocity and attitude error vector, respectively; $\boldsymbol{\omega}_{ie}^n$ is the Earth rotation velocity; $\boldsymbol{\omega}_{in}^n$ is the angular rotation velocity of the n -frame with respect to the i -frame; \mathbf{f}^n is the specific force vector; $\delta\mathbf{g}^n$ is the error of the gravity vector in the n -frame; $\boldsymbol{\omega}_{en}^n$ is the rotation vector from the e -frame to the n -frame; \mathbf{C}_b^n is the transition matrix from body frame (b -frame) to the n -frame; and ∇^b and $\boldsymbol{\varepsilon}^b$ are the accelerometer and gyro drift vector expressed in the b -frame, respectively.

The dynamic matrix is obtained by a linearisation of Eq. (66.2). The measurement model is:

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \quad \mathbf{v}_k \sim N(0, \mathbf{R}_k) \quad (66.3)$$

where \mathbf{x} is the error state vector including \mathbf{x}_{nav} and \mathbf{x}_{ins} in Eq. (66.1); \mathbf{z}_k is the measurement vector; \mathbf{R}_k is the measurement variance-covariance matrix; \mathbf{H}_k is the design matrix that relates the measurements to the state vector; and \mathbf{v}_k is the measurement residual vector.

66.3 WiFi Augmentation for Position and Orientation Determination

The WiFi fingerprinting algorithm consists of the following two steps:

- (1) In the training phase, the RSS measurement set from several Access Points (APs) at Reference Points (RPs) are measured and recorded as features of the RPs, to form a database for subsequent positioning.
- (2) In the positioning phase, the RSS measurement set observed at an unknown point is compared with the signal characteristics of the RPs contained within the database. The RP signal data which have the closest match with the RSS measurement set are chosen as the estimated location.

Since the position estimation actually is derived from the fingerprinting database, the quality of the database directly impacts on the final position accuracy. One option for improving accuracy is to build a database with denser set of sampled RPs. However this would require the investment of significant additional survey resources during the training phased, and would increase computation burden in the positioning phase. Another way to improve the quality of the database is to consider additional information that provides more constraints in the searching and matching algorithm. The next subsection describes a new database for WiFi positioning that incorporates map-based information.

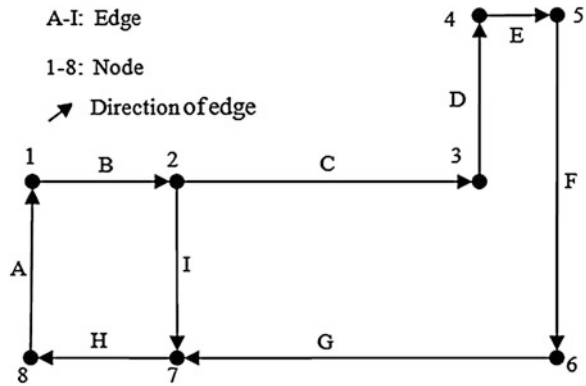
66.3.1 Topology Modelling in the Fingerprinting Database

Two kinds of map-based information are introduced into the database to support the WiFi-only and WiFi/INS integrated positioning solutions. The information is derived from assumptions that pedestrian travel in indoor environment mainly occurs in corridors, for which two conditions apply:

- (1) The person cannot directly pass from one corridor to another disconnected corridor.
- (2) When walking in a corridor, the heading direction always coincides with the direction of the corridor.

Noted that this study is concerned with the walking behaviour from an outdoor origin to an indoor destination, thus those actions such as pausing and facing a wall are not considered. Based on the above two assumptions, the topological relationship between any two corridors can be defined. Topology, which is a very common concept in GIS (Geographic Information System), expresses the spatial relationships between connecting or adjacent vector features (points, polylines and polygons). Corridors inside buildings can be abstracted as polylines, thus the road topology concept in GIS can be used to describe the relationships among corridors in an indoor area. In addition, the trajectory of walking along corridors can be

Fig. 66.1 Topological elements and relationships of a layer inside a building



modelled as a set of discrete points. Figure 66.1 shows the topological elements and relationships among corridors.

Figure 66.1 graphically illustrates the two types of elements and their topological relationships. Labels “1”–“8” are called “nodes”, which are the connection points of adjacent corridors. Labels “A”–“I” are referred to as “edges” in a standard GIS database, and are corridors within a building. The arrows on the edges denote the direction of the edges, which are stored in the database as a feature of each edge. The topological relationships among these nodes and edges can be described as collections of geographic features, which are listed in Table 66.1.

Considering the above topological relationships, the RPs’ features stored in the fingerprinting database will also include this geographic information. Figure 66.2 shows the structure of the new database known as “RP”. The fingerprinting database actually stores features of a set of RPs, which are preselected points with known coordinates along the corridors. The coordinate of each RP is stored in the structure of “LatLon”. During the database building phase, a mobile user device is located at each RP for a short period of time. The MAC address and RSS measurements from each AP are collected and stored in the “AP” structure, named as “MAC_ad” and “RSS”. The new database should also contain a topology table that

Table 66.1 Edge structure

Edge	Direction	From node	To node
A	0	8	1
B	90	1	2
C	90	2	3
D	0	3	4
E	90	4	5
F	180	5	6
G	-90	6	7
H	-90	7	8
I	180	2	7

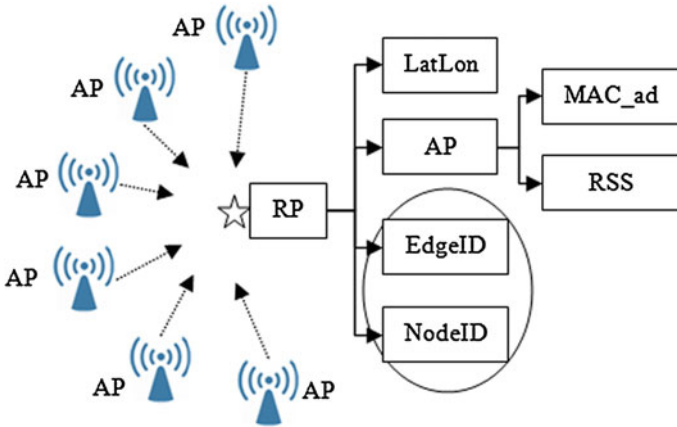


Fig. 66.2 Fingerprinting database structure

describes the geospatial information of each corridor, similar to Table 66.1. Depending on their particular topological properties these RPs are divided into two classes and associated with different features.

- (1) For RPs which are very close to a connection point of several corridors, the corresponding node index will be recorded in the “NodeID” structure, and 0 will be recorded in the “EdgeID”.
- (2) For RPs which are distant from a connection point, the corresponding edge index will be recorded in the “EdgeID” structure, and 0 will be recorded in the “NodeID”.

By searching the values of “EdgeID” or “NodeID” in the topology table, one can determine which corridor the user is located in/at, and can also obtain the orientation information. With this additional information the positioning accuracy can be improved.

Geospatial topology is fundamentally used to ensure data quality as well as to aid data compilation. The use of topology in this study can be grouped into two aspects. The first is to augment the fingerprinting database for WiFi positioning. As described above, the topological relationships among each RP are stored together with the RSS measurements and coordinates, thus the features of RPs are extended. The other use is to provide constraints in the positioning phase, which will be discussed below. Firstly, in WiFi-only position the candidates which are not adjacent or connected with the previous position determination will be excluded. Then the orientation information associated with the selected candidates will be used as pseudo-measurement in the WiFi/INS integrated solution step. In the first step the accuracy of the prior position is important. An erroneous prior estimate will bias the decision concerning the exclusion of candidate solutions. By integrating INS with WiFi, the short-term reliability and stability of the prior position can be guaranteed.

66.3.2 Searching and Matching Algorithms

In fingerprinting positioning phase, an RSS measurement set is collected at each sampling epoch. By comparing the RSS measurement set with those stored in the database, a user's position can be obtained. Some techniques are deterministic, meaning that the applied RSS data are mean values and an RSS measurement set is matched with only one reference point. Transparent location fingerprinting uses a map of RPs [2]. A RP consists of a sequence of pairs (ss_i, c_i) . ss_i is a RSS measurement set and c_i is the corresponding physical coordinates. The matching of a new measurement set ss is done by selecting a number of RPs (k) that are closest to the measurement. c_f , the weighted average of their coordinates, is returned:

$$c_f = \frac{\sum_{j=1}^k \frac{1}{d(ss_j, ss) + \varepsilon} \cdot c_j}{\sum_{j=1}^k \frac{1}{d(ss_j, ss) + \varepsilon}} \quad (66.4)$$

where $d(ss_j, ss)$ is the Euclidean distance between the two RSS measurement sets and ε is a small real constant.

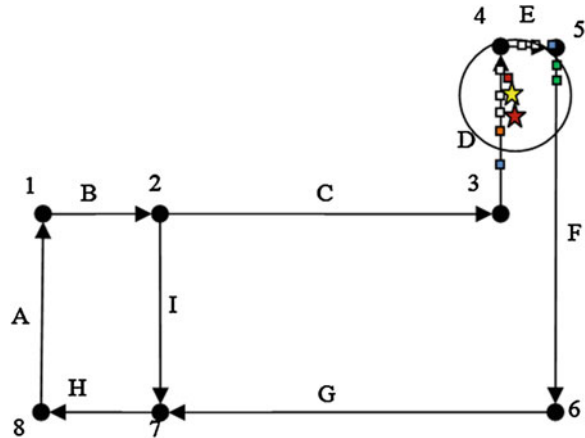
The probabilistic techniques use a RSS probability distribution at each RP rather than a mean value. Roos et al. [13] use two estimators to match a measurement with the database. In the first method, a probability mass is assigned to a "kernel" around each observation in the training data. The second method is closely related to discretization of continuous values.

Without considering the geospatial relationships among these points, some faulty RPs which are physically far away from the mobile terminal may be selected. To exclude these incorrect RPs from the k candidates, the following three criteria based on the topological database are enforced.

- (1) The candidate which is not adjacent with the prior location of the user will be excluded;
- (2) The candidate which is not consistent with the user's walking direction will be excluded;
- (3) The candidate which is not within the estimation error circle will be excluded.

Figure 66.3 is a demonstration of the exclusion decision based on the above three criteria. The yellow star point is the prior estimated position based on the previous coordinates (red star point), and its estimated error is indicated by the circle. The rectangular points are k candidates obtained by the searching algorithm. With the first criterion, the two green rectangular points are excluded since they are actually located at corridor F which is not connective with corridor D, where the current location is supposed to be. Using the second criterion, the orange rectangular point is excluded. The other two blue rectangular points are also excluded by the third criterion. Finally the red rectangular point is obtained as the estimated position.

Fig. 66.3 A demonstration of the exclusion based on topological information



The three criteria may not be always effective. The first criterion will not work for a single corridor. The second criterion is not relevant when the user is turning since the instantaneous walking direction is difficult to determine. The third criterion largely relies on the estimation error thus a miss-detection will occur if the error radius is large or a wrong exclusion will be introduced using an over-strict error radius.

66.4 Experiment and Analysis

66.4.1 Experiment Description

Two experiments were analysed to assess the navigation performance under different indoor environments. In these experiments the distance between two APs is an average of 1 m. The data is sampled at each RP for 1 min. The nAX5, a small MEMS-INS/GPS device, was used to collect inertial and GPS data, with the GPS antenna fixed to the top of the device as shown in Fig. 66.4. For these tests the output rate of the inertial sensors was set to 10 Hz. The WiFi data was collected by a laptop with built-in wireless internet card with a sample rate of 1 Hz.

The first experiment was conducted on the lower ground level of the Electrical Engineering Building at the University of New South Wales (UNSW), Sydney, Australia. The building layout is shown in Fig. 66.5. The second experiment was conducted on level 2 of the Eastgardens Shopping Center, Sydney, Australia, and the building layout is shown in Fig. 66.6. The red points in Figs. 66.5 and 66.6 are the RPs in the corresponding fingerprinting databases.

Fig. 66.4 Experiment setup

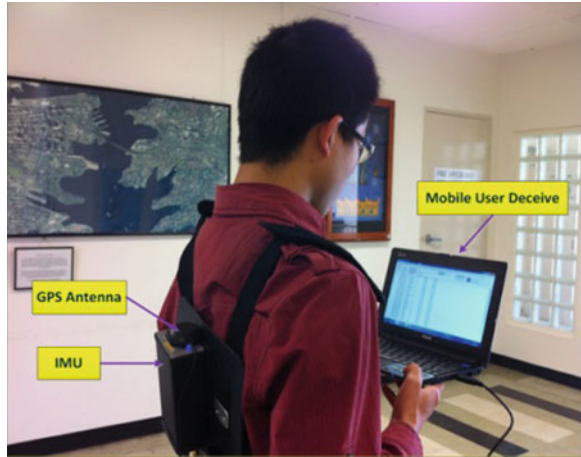


Fig. 66.5 Building layout for experiment 1

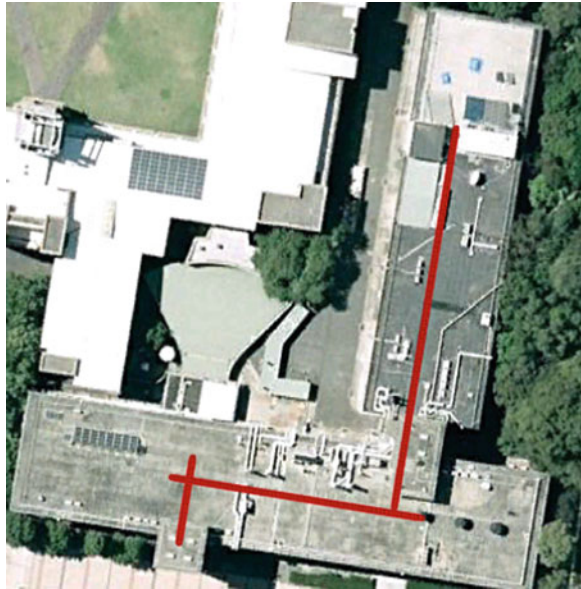


Fig. 66.6 Building layout for experiment 2



66.4.2 Performance Evaluation of WiFi-Only System

Experiments 1 and 2 are typical layouts of offices and shopping centres. In the office building rooms are usually located at two sides of the corridors. A common design of shopping centres locates shops around one side of the circular corridors. Figures 66.7 and 66.8 show the positioning results of these two experiments, where the blue marks are results by using traditional fingerprinting methods, and red marks are results obtained using the proposed method taking into account the topological information associated with corridors.

It can be seen from Fig. 66.7 that the WiFi positioning results with and without topological information are similar. In contrast, Fig. 66.8 shows an obvious improvement by using topological information for Experiment 2, where the red points have a much smoother trajectory than the blue points. The different performances in these two experiments are likely due to the different characteristics of the building corridors. In Experiment 1, there are only two parallel corridors which are more than 30 m apart and connected by a perpendicular corridor. It is nearly impossible to miss-locate one point to the other corridor which is not connected with the current corridor. However, the building layout in Experiment 2 is more complex. There are seven corridors, and some of them are parallel but close to each other. As a result the possibility of locating a point in a wrong corridor is high. By considering the topological relationships among corridors, the probability of miss-matching can be reduced.

Fig. 66.7 WiFi-only positioning results of experiment 1

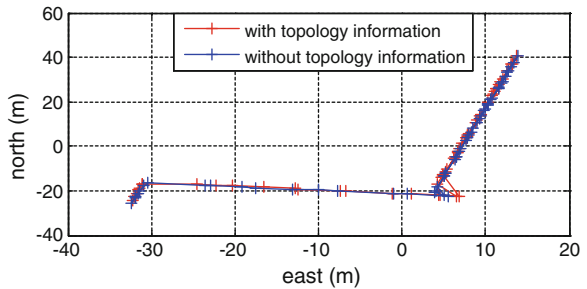
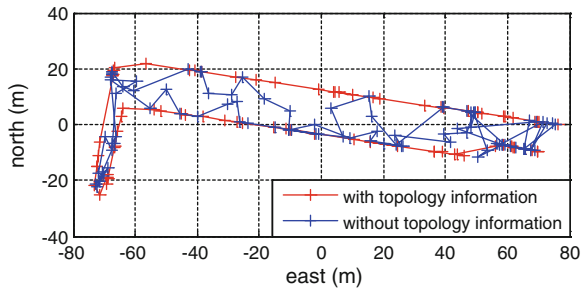


Fig. 66.8 WiFi-only positioning results of experiment 2



66.4.3 Performance Evaluation of the WiFi/INS Integrated System

The position and attitude solutions for Experiment 1 are shown in Figs. 66.9, 66.10 and 66.11, where the red line is the GPS/INS solution, the blue and green lines are results from WiFi/INS integrated system, without and with map-based information. The performance in Experiment 2 is similar and not presented here due to space constraints.

The red line in Fig. 66.9 shows that the INS-only solution diverges quickly in an indoor environment. The blue line indicates that upon adding position corrections

Fig. 66.9 Trajectories of Experiment 1 (red GPS/INS, blue/green GPS/WiFi/INS without/with topology constraints)

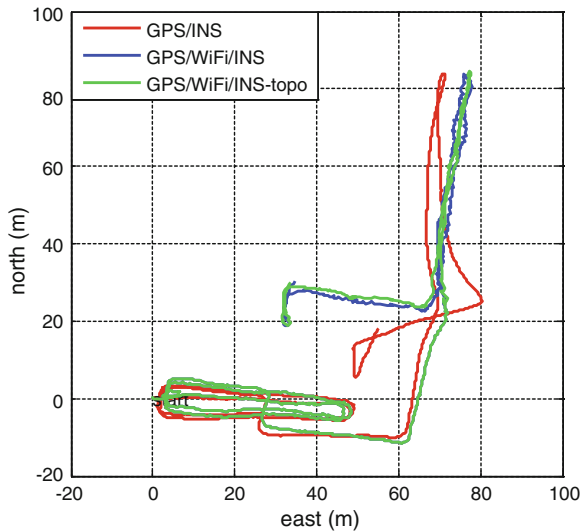
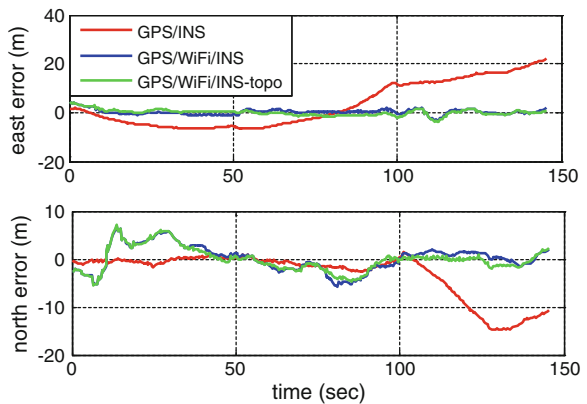


Fig. 66.10 Horizontal position errors of Experiment 1 (red GPS/INS, blue/green GPS/WiFi/INS without/with topology constraints)



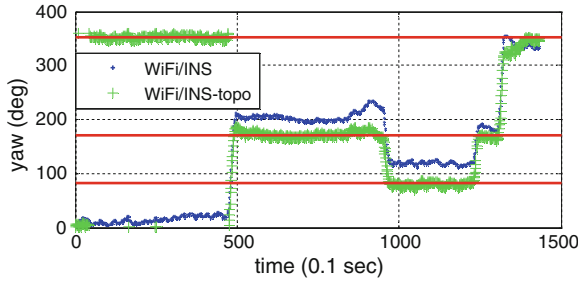


Fig. 66.11 Yaw results of Experiment 1 during indoor period (*red* reference from building orientation information, *blue/green* results of WiFi/INS without/with topological constraints)

Table 66.2 Horizontal positioning errors

	East error (m)		North error (m)	
	Mean	RMS	Mean	RMS
GPS/INS	3.11	9.64	-3.09	5.64
GPS/WiFi/INS	0.21	1.23	0.14	2.67
GPS/WiFi/INS-topo	0.003	1.24	-0.09	2.48

from WiFi the divergence can be largely controlled, such that the trajectory is much more consistent with the geometry of the corridors. The green line indicates the best performance, when considering map-based information in the integrated system. This means that by using building orientation constraints, the position accuracy and solution stability is improved. The horizontal position errors of the trajectories inside the building are plotted in Fig. 66.10. It shows that without the absolute position solutions the performance GPS/INS navigation (red) degrades severely. Aided by WiFi the position errors can be reduced significantly. The mean and RMS (root meat square) values of the horizontal position errors are given in Table 66.2. It shows that with the aid of WiFi the mean values of position errors are around 1 m in east and 3 m on north directions, and the RMS (root-mean-squares) of the errors reduced from 10 to 3 m. The performance of GPS/WiFi/INS system with topological constraints is slightly better than the system without topological constraints.

Figure 66.11 shows the attitude solutions without and with topological constraints. The red lines are the references solution from the building orientation information. It can be seen that the green points are much more consistent with the references, while the yaw angle given by the blue points are not accurate and not stable even when the person is walking along a corridor. This further indicates the accuracy and stability of attitude estimation is improved by considering the building orientation information in the fingerprinting database.

66.5 Concluding Remarks

The extensive use of WiFi-based positioning provides an alternative to realise continuous navigation as a person moves from outdoors to indoor environments, where GPS becomes unavailable. In this paper an enhanced WiFi database is proposed that takes into account the topological relationships among corridors and orientation information of each corridor. By integrating WiFi with INS, the long-term navigation stability for indoor environments is further improved. Experiment results have verified that by adding building/corridor topological information into the fingerprinting database, incorrect candidate solutions can be excluded with a higher success rate, and the searching and matching reliability is improved in WiFi-only positioning. By using the more accurate position measurements as well as the orientation constraint provided by WiFi, the performance of a WiFi/INS integrated navigation system is also improved. The orientation constraint has a significant impact on accuracy of position and attitude estimation.

Acknowledgments The first author was sponsored by the China Scholarship Council (CSC) for her Ph.D. studies at the University of New South Wales, Australia. The authors would like to thank Dr Binghao Li and Jiantong Cheng for their assistance regarding the conventional WiFi positioning algorithm. This work was supported by the Shanghai Huace Navigation Technology Corp.

References

1. Akyildiz IF, Sun Z, Vuran MC (2009) Signal propagation techniques for wireless underground communication networks. *Phys Commun* 2(3):167–183
2. Brunato M, Kiss Kalló C (2002) Transparent location fingerprinting for wireless services. In: *Proceedings of Med-Hoc-Net, mediterranean workshop on Ad-hoc networks*, Baia Chia, Cagliari, 17 September 2002. Report No. DIT-02-071
3. Chen Y, Kobayashi H (2002) Signal strength based indoor geolocation. In: *IEEE international conference on communications*, New York, 28 April–2 May 2002, p 436–439
4. Evennou F, Marx F (2006) Advanced integration of WiFi and inertial navigation systems for indoor mobile positioning. *EURASIP J Appl Sig Process* 1:164–174
5. Godha S, Cannon ME (2007) GPS/MEMS INS integrated system for navigation in urban areas. *GPS Solutions* 11(3):193–203
6. Godha S, Lachapelle G (2008) Foot mounted inertial system for pedestrian navigation. *Meas Sci Technol* 19(7), 075202 (9 pp)
7. Goshen-Meskin D, Bar-Itzhack IY (1992) Unified approach to inertial navigation system error modelling. *J Guid Control Dyn* 15(3):648–653
8. Hwang DH, Oh SH, Lee SJ, Park C, Rizos C (2005) Design of a low-cost attitude determination GPS/INS integrated navigation system. *GPS Solutions* 9(4):294–311
9. Li B, Salter J, Dempster AG, Rizos C (2006) Indoor positioning techniques based on wireless LAN. In: *1st IEEE international conference on wireless broadband and ultra wideband communications*, Sydney, Australia, 3–16 March 2006, paper 113, CD-ROM proc
10. Li B, Wang Y, Lee HK, Dempster AG, Rizos C (2005) Method for yielding a database of location fingerprints in WLAN. *IEE proceedings-communications* 152(5):580–586
11. Li Y, Efatmaneshnik M, Dempster AG (2012) Attitude determination by integration of MEMS inertial sensors and GPS for autonomous agriculture applications. *GPS Solutions* 16(1):41–52

12. Ochieng WY, Sauer K, Walsh D, Brodin G, Griffin S, Denney M (2003) GPS integrity and potential impact on aviation safety. *J Navig* 56(1):51–65
13. Roos T, Myllymäki P, Tirri H, Misikangas P, Sievänen J (2002) A probabilistic approach to WLAN user location estimation. *Int J Wireless Inf Networks* 9(3):155–164
14. Youssef MA, Agrawala A, Udaya Shankar A (2003) WLAN location determination via clustering and probability distributions. In: 1st IEEE international conference on pervasive computing and communications, Fort Worth, USA, 26–26 March 2003, pp 143–150

Chapter 67

A Pedestrian Movement Direction Recognition Method Based on Inertial Sensors

Shunbao Lu, Zhongliang Deng, Chen Xue, Yeqing Fang,
Ruoyu Zheng and Hui Zeng

Abstract Most of the phone has been configured IMU currently, includes an acceleration sensor and an electronic compass. In this paper, using the pedestrian dead reckoning algorithm based accelerometer and electronic compass composition IMU to assist indoor positioning. Detection walking step length and state. Currently detect direction of motion requires the user wear or hand in IMU with a fixed posture. To address this issue, using the terminal gesture recognition algorithm in the third quarter. When pedestrian walking on the same direction, the angle walking forward and walking backward was the same. To address this issue, using pedestrian movement direction detection method based on the differential cross-correlation of the acceleration in the fourth quarter. Based on the above, the article conducted experiments in Section V and the result shows that the proposed method can effectively detect pedestrian movement forward and backward state, the average accuracy of the detection results is 85.67 %.

Keywords IMU · Pedestrian movement direction · Recognize forward and backward state

67.1 Introduction

With the popularity of navigation feature on your phone and development of location based services, people increasingly demands for more accurate and reliable pedestrian navigation. Most of the phone has been configured IMU currently, includes an acceleration sensor and an electronic compass.

S. Lu (✉) · Z. Deng · C. Xue · Y. Fang · R. Zheng · H. Zeng
School of Electronic Engineering, Beijing University of Posts and Telecommunications,
Beijing 100876, China
e-mail: 371368153@qq.com

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015
Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_67

Its basic principle is: wearing the IMU in a fixed posture on a person, in the case of the normal travel of the pedestrian [1]. To do the cadence detection of the pedestrian through the cyclical changes in the acceleration when people walk, to estimate the pedestrian step length based on the pedestrian step length estimation model and using a electronic compass to get the pedestrian movement direction [2]. The cadence detection and step length estimation has been relatively mature, the main difficulty in pedestrian dead reckoning is the detection of the pedestrian movement direction. The pedestrian movement direction detection requires the user wearing the IMU in a fixed posture currently, to keep a certain relationship with the IMU coordinate system while walking. Then transform the coordinate system from the IMU to the pedestrian to gain the direction of the pedestrian. Obviously this method has certain limitation, it brings a lot of inconvenience to pedestrians. More than that, the basic principle to get the pedestrian movement direction is to calculate the arctangent value between two horizontal components of the geomagnetic field on the pedestrian coordinate system. Therefore, the azimuth is the same when a person goes forward and backward in a same route. Obviously it's a mistake, can't recognize the forward and backward state is the primary cause of the problem [3].

To solve these problems, we studied the relationship between the acceleration and direction of pedestrian in the Chap. 2 first. In the Chap. 3, we described the principle of the direction detection by the electronic compass briefly. Based on that, we used a method to detect the pedestrian movement direction which is related to the acceleration difference cross-correlation and studied the recognition pedestrian backward and forward gesture and did some experiments in the Chap. 5 to prove the effectiveness of this method. This article used HMC5883L electronic compass and MMA8451Q accelerometer that can be used in mobile phones and other kinds of terminal [4].

67.2 Correlation Between Accelerations While Pedestrian Walking

The definition of the coordinate system has been shown in Fig. 67.1 which the OX'Y'Z' is the coordinate system of the terminal. OY 'axis points the front of the terminal screen. OX 'axis is pointing to the right of the terminal screen. OZ 'axis is perpendicular to the top of the screen pointing to the terminal screen. The OXYZ is the pedestrian coordinate system. OY axis points pedestrians facing direction. OX-axis points the right of pedestrian. OZ axis is perpendicular to the pedestrian pedestrians pointing upward.

Acceleration in the vertical direction and the horizontal direction contain the complete process of rising-dropping-rising reverse-dropping reverse in each step of the pedestrian. When moving continuously, the acceleration has a periodical change characteristic. To analyze the direction of the relationship between acceleration and pedestrian movement further, we made person start moving from a standstill along the X-axis. The test results have been shown in Fig. 67.2.

Fig. 67.1 Pedestrian coordinate system and platform coordinate system

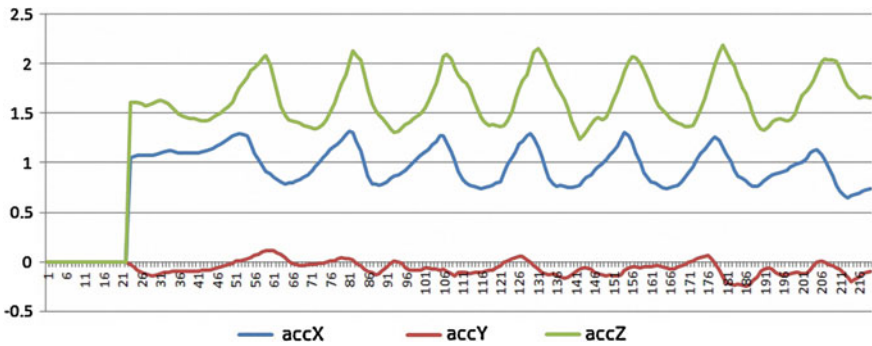
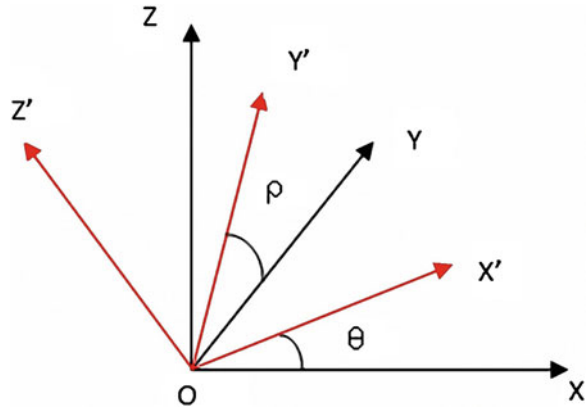


Fig. 67.2 Changes in acceleration along the X-axis during moving on positive direction

From Fig. 67.1 we could found that when a pedestrian is walking along the X-axis of the terminal, the waveform of acceleration appeared in the X-axis and Z-axis are similar and they could reflect the movement of the pedestrian. But on the Y-axis we can hardly find the waveform. When the pedestrian moved along the Y axis, we can also get similar results.

67.3 Electronic Compass Based Direction Detection

67.3.1 Basic Principle of Direction Detection by Electronic Compass

One of the key issues in PDR algorithm is how to get accurate orientation from low-cost sensors. Now the magnetic compass and accelerometer sensors have been widely used in various platforms. In the case of the Earth’s magnetic field is not

interfered by other magnetic field. Through these two elements we could get a more accurate orientation. The principle to get orientation is sensing the acceleration of gravity by the accelerometer. Then calculate the gravitational acceleration component on each axis in the coordinate OX', Y', Z' , and record them as A_X, A_Y, A_Z .

Pedestrians can't guarantee handheld platform parallel to the ground in PDR. In this time, platform coordinate system is no longer coincided to the pedestrian coordinate system. But the three-axis value output by the magnetic sensor remains the three components of the geomagnetic field in the platform coordinate system. We need to transform the value to the pedestrian coordinate system to get orientation according to formula 67.1. We can describe the connection of these two coordinate systems by pitch and roll angle. The pitch angle ρ is the angle that the platform rotated around the OX axis pedestrian coordinate system. The roll angle θ is the angle that the platform rotated around the OY axis pedestrian coordinate system. When the platform rotate around the axis OY and OX , pitch angle ρ and roll angle θ can be calculated by formulas 67.1 and 67.2.

$$\rho = \frac{A_Y}{\sqrt{A_X^2 + A_Z^2}} \tag{67.1}$$

$$\theta = \frac{A_X}{\sqrt{A_Y^2 + A_Z^2}} \tag{67.2}$$

After we get the pitch and roll angle, we can transform the three components of the geomagnetic field on platform coordinate system to the pedestrian coordinate system according to formula 67.3. (M'_x, M'_Y, M'_Z) are three components of the geomagnetic field on pedestrian coordinate system and (M_X, M_Y, M_Z) are three components of the geomagnetic field on platform coordinate system. After we get three components of the geomagnetic field on pedestrian coordinate system, we can calculate the orientation according to formula 67.4.

$$H(M'_x, M'_Y) = \begin{cases} 90.00, & \text{IF } M'_x > 0, M'_Y = 0 \\ 270, & \text{IF } M'_x < 0, M'_Y = 0 \\ 360 + \arctan\left(\frac{M'_x}{M'_Y}\right), & M'_Y > 0 \\ 180 + \arctan\left(\frac{M'_x}{M'_Y}\right), & \text{IF } M'_Y < 0 \end{cases} \tag{67.3}$$

67.3.2 Terminal Gesture Recognition Algorithm

The method said above needs the consistent of the direction between pedestrian and platform. It's difficult to maintain during pedestrian constantly traveling so that we

used another way (formulas 67.4 and 67.5) to calculate the orientation which pitch angle is ρ and roll angle is θ . Ac is the value output by the electronic compass. β is the final orientation.

$$\alpha = \arctan\left(\frac{\theta}{\rho}\right) \quad (67.4)$$

$$\beta = Ac - \alpha \quad (67.5)$$

The above method doesn't require the alignment of the pedestrian heading and reference direction of platform. Pedestrian can handheld the platform at any angle when traveling. But this method has a problem that it still uses the three components of the geomagnetic field. Therefore, the orientation is the same when pedestrian walking along the same route forward and backward. The 180° difference of the orientation does not appear. The orientation is wrong when pedestrian walking backward.

67.4 Forward and Backward Detection of Pedestrian

In Chap. 4, we will introduce a motion direction method based on the cross-correlation characteristic of acceleration. In a complete gait cycle, it is considered that the error caused by the platform attitude measurement errors approximately unchanged. To eliminate these errors and improve the part of the signal, we detect the direction of motion by the differential calculus in a pedometer-cycle and define the m step's acceleration differential cross-correlation value $R_{x,z,m}(\tau'_{x,m})$, $R_{y,z,m}(\tau'_{y,m})$ as formula 67.6:

$$\begin{cases} R_{x,z,m}(\tau'_{x,m}) = \sum_{i=N_{m-1}+1}^{N_m} \{a_x[t_s(i - \tau'_{x,m})] - a_x[t_s(i - \tau'_{x,m} - 1)]\} \{a_z(t_s \cdot i) - a_z[t_s(i - 1)]\} \\ \quad = \text{sgn}(v_x) (S_{x,z,m}(\tau'_{x,m}) + I_{x,z,m}(\tau'_{x,m})) \\ R_{y,z,m}(\tau'_{y,m}) = \sum_{i=N_{m-1}+1}^{N_m} \{a_y[t_s(i - \tau'_{y,m})] - a_y[t_s(i - \tau'_{y,m} - 1)]\} \{a_z(t_s \cdot i) - a_z[t_s(i - 1)]\} \\ \quad = \text{sgn}(v_y) (S_{y,z,m}(\tau'_{y,m}) + I_{y,z,m}(\tau'_{y,m})) \end{cases} \quad (67.6)$$

t_s is Sensor data sampling interval, N_m Indicates the num of the sampling points when the m step ends. $\tau'_{x,m}$ and $\tau'_{y,m}$ are local offset delay for correlation calculation. $S_{x,z,m}(\tau'_{x,m})$ and $S_{y,z,m}(\tau'_{y,m})$ are signal components in the integration result. $I_{x,z,m}(\tau'_{x,m})$ and $I_{y,z,m}(\tau'_{y,m})$ are Noise components in the integration result. τ_m is the

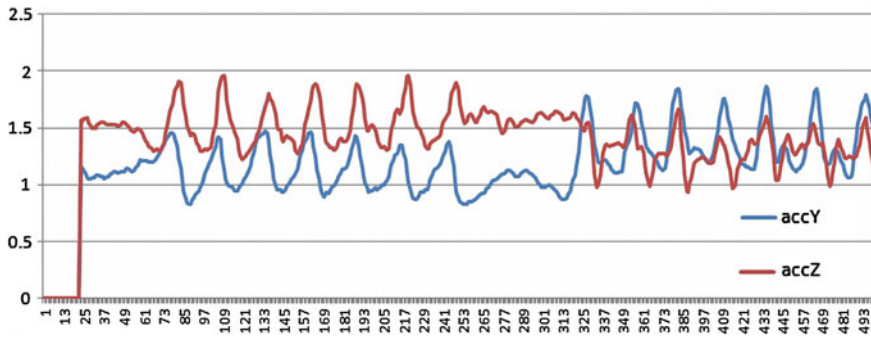


Fig. 67.3 Waveform phase difference

amount of lead that acceleration changing in horizontal than in vertical. When τ'_m equals to τ_m , the signal component in $R_{x,z,m}(\tau'_{x,m})$ and $R_{y,z,m}(\tau'_{y,m})$ are maximizing.

After a lot of research, we found that the presence of the Z-axis motion acceleration waveform is different with the main motion axis direction (shown as Fig. 67.3) while pedestrian walking forward and backward. Therefore, the correlation value of acceleration between Z-axis and main motion axis tends to be stable while pedestrian walking forward and backward. While pedestrian changes the motion state from forward to backward, the correlation value would have a significant trend of change (shown as Fig. 67.4). While this tendency is caught, we could found that the pedestrian motion state has changed.

We need to calculate the correlation value of acceleration between Z-axis and main motion axis each step of pedestrian and pick out the max value from the data series. We found that the correlation value is a stable negative when Y-axis is the main motion while pedestrian moving forward. The correlation value is a stable positive when Y-axis is the main motion while pedestrian moving backward. We can find the changing of pedestrian motion direction. With the orientation we get from formulas 67.4 and 67.5, we can do a 180° processing. Then we can get a more accurate heading value.

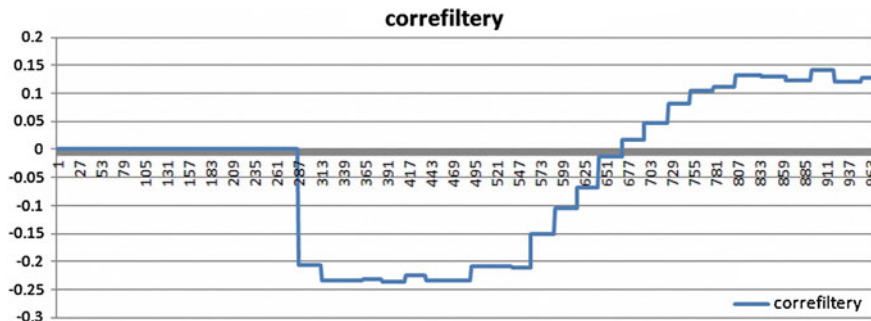


Fig. 67.4 Correlation value change

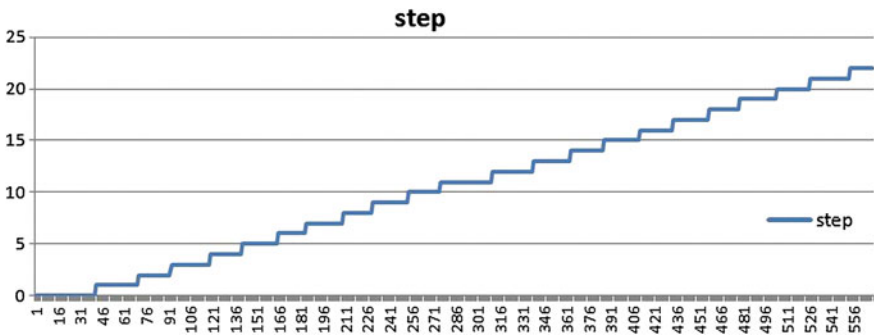


Fig. 67.5 The changing process of step forward and backward

67.5 Experimental Verification

Based on the method described in Sect. 67.4, we did some tests on handheld terminal and took 12 directions. In each direction, we forward 10 steps and backward 10 steps for 50 times. The pedometer effects are shown in Fig. 67.5 and testing results are shown in Table 67.1. The results show that the method described in Sect. 67.4 can detect the pedestrian motion state. The average accuracy rate of the test results is 85.67%. The orientation changing while pedestrian goes forward and then goes backward along North are shown in Fig. 67.6.

The testing results can prove that the method described in this article can be used to detect the pedestrian motion state. Integrating this method with PDR can enhance the effect of wireless indoor location and reduce the error further.

Table 67.1 Handheld terminal test results

Real angle (forward) (°)	Real angle (backward) (°)	Num of success detection	Num of detection failure	Detection accuracy (%)
0	180	42	8	84
30	210	41	9	82
60	240	44	6	88
90	270	43	7	86
120	300	42	8	84
150	330	46	4	92
180	0	43	7	86
210	30	44	6	88
240	60	41	9	82
270	90	43	7	86
300	120	42	8	84
330	150	43	7	86

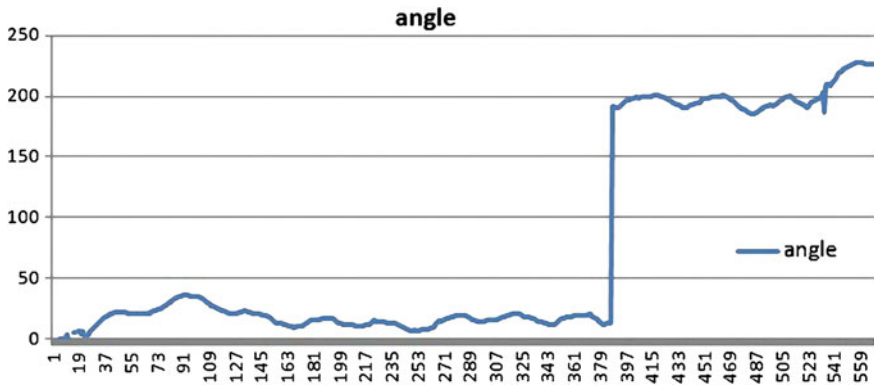


Fig. 67.6 The change of orientation when forward and backward

Acknowledgments Foundation project: The National High Technology Research and Development Program (“863” Program) of China (No. 2012AA120801) and The National Natural Science Foundation of China (No. 61372110).

References

1. Foxlin E (2005) Pedestrian tracking with shoe-mounted inertial sensors. *IEEE Computer Graph Appl* 25(6):38–46
2. Chen W (2009) Research on GPS, Self, contained sensors based seamless outdoor/indoor pedestrian positioning algorithm
3. Susi M (2013) Motion mode recognition and step detection algorithms for mobile phone users. *Sensors* 13:1539–1562
4. Asano S (2012) Pedestrian dead-reckoning unit for navigation system using mobile device. In: *The 1st IEEE global conference on consumer electronics*

Chapter 68

High-Precision Simulator for Strapdown Inertial Navigation Systems Based on Real Dynamics from GNSS and IMU Integration

Gongmin Yan, Jinling Wang and Xinyi Zhou

Abstract Traditional Strapdown Inertial Navigation Systems (SINS) simulator based on pure mathematical model cannot reflect the reality of complex dynamics of a carrier. In this paper, based on the actual flight data of an airborne navigation system, navigation algorithms of high-precision carrier phase differences Global Navigation Satellite System (GNSS) and medium-precision Inertial Measurement Unit (IMU) are used, and combined with the partial feedback principle of Kalman filter and cubic spline interpolation method, to generate smooth trajectory parameters. Then, the simulation of inertial sensors is achieved with the inversed SINS algorithm, and the new algorithm considers the impacts of attitude coning error and velocity sculling error compensation. Numerical verifications of airborne flight test have demonstrated that the inertial sensor simulator has high accuracy and good frequency characteristic, and can meet the high-accuracy SINS requirements for simulated inertial sensor data source and frequency complexity.

Keywords GNSS · SINS simulator · Integrated navigation · Actual flight data

68.1 Introduction

For the simulated research of Strapdown Inertial Navigation Systems (SINS) and its integration with other systems, it is a fundamental step to obtain the carrier's trajectory parameters and the simulation outputs of inertial sensors (including

G. Yan (✉) · X. Zhou

School of Automation, Northwestern Polytechnical University,
Youyi West Rd, Xi'an 710072, Shaanxi, China
e-mail: yangongmin@163.com

J. Wang

School of Civil and Environment Engineering,
UNSW (University of New South Wales), Sydney NSW 2052, Australia
e-mail: jinling.wang@unsw.edu.au

© Springer-Verlag Berlin Heidelberg 2015

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_68

gyroscopes and accelerometers). Especially for the research of high-accuracy SINS algorithms, high precision dynamic simulation for a carrier is critical [1]. In the simulation of airborne SINS, for example, traditional trajectory simulator based on pure mathematical model can only provide the carrier's simple maneuvers, including still, accelerate, climb, cruise, twists and turns, and the integration of these flight maneuvers. The above trajectory form is too simple and thus, cannot reflect the reality of complex dynamics of a carrier. However, if the impact of the pneumatic environment of the carrier aircraft is considered, physical and mathematical models are often too complex to be described. Actually, more or less, there still remains some difference between ideal model and the actual flight situation [2].

In this paper, based on the actual flight data of inertial sensors, the integrated navigation algorithms of high-precision carrier phase differences Global Navigation Satellite System (GNSS) and medium-precision Inertial Measurement Unit (IMU) are used, combined with the partial feedback principle of Kalman filter and cubic spline interpolation method, to generate smooth trajectory parameters, i.e. attitude, velocity and positioning information. The incremental outputs of inertial sensors are inversely deduced from attitude, velocity and positioning information with the use of traditional high-precision SINS updating algorithm [3–5], which can be called as inversed SINS algorithm, and the new algorithm considers the impacts of attitude coning error and velocity sculling error compensation. Then, the simulated outputs of inertial sensors are achieved with the inversed SINS algorithm. Numerical verifications and comparison analysis using airborne flight experiment have demonstrated that the simulated outputs of inertial sensors have high navigation accuracies, and such accuracies are only limited by the numerical calculation error. The power spectrum of simulated outputs and actual outputs are basically the same in the low-frequency band, indicating that simulated outputs can properly reflect the dynamic environment characteristics of the carrier flights.

68.2 Traditional Strapdown Inertial Navigation Algorithms

68.2.1 SINS Differential Equations

Firstly, some coordinate frames are briefly introduced: the i -frame as inertial coordinate frame, the e -frame as earth coordinate frame, the n -frame as 'East-North-Up' SINS navigation frame and the b -system as 'Right-Forward-Up' body frame.

SINS navigation algorithms consist of a set of differential equations, i.e. attitude, velocity and position differential equation, respectively, as

$$\dot{C}_b^n = C_b^n (\omega_{nb}^b \times) \quad (68.1)$$

$$\dot{\mathbf{v}}^n = C_b^n \mathbf{f}_{sf}^{cb} - (2\omega_{ie}^n + \omega_{en}^n) \times \mathbf{v}^n + \mathbf{g}^n \quad (68.2)$$

$$\dot{\mathbf{p}} = \mathbf{M}_{pv} \mathbf{v}^n \quad (68.3)$$

where:

$$\begin{aligned} \boldsymbol{\omega}_{nb}^b &= \boldsymbol{\omega}_{ib}^b - (\mathbf{C}_b^n)^T \boldsymbol{\omega}_{in}^n, & \boldsymbol{\omega}_{in}^n &= \boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n, \\ \boldsymbol{\omega}_{ie}^n &= [0 \quad \omega_{ie} \cos L \quad \omega_{ie} \sin L]^T, & \boldsymbol{\omega}_{en}^n &= \left[-\frac{v_N^n}{R_{Mh}} \quad \frac{v_E^n}{R_{Nh}} \quad \frac{v_U^n}{R_{Nh}} \tan L \right]^T, \\ \mathbf{M}_{pv} &= \text{diag}(\sec L/R_{Nh} \quad 1/R_{Mh} \quad 1), & R_{Mh} &= R_M + h, \quad R_{Nh} = R_N + h, \\ R_M &= \frac{R_N(1 - e^2)}{(1 - e^2 \sin^2 L)}, & R_N &= \frac{R_e}{(1 - e^2 \sin^2 L)^{1/2}}, \quad e = \sqrt{2f - f^2}, \\ \mathbf{g}^n &= [0 \quad 0 \quad -g]^T, & g &= g_0(1 + \beta_1 \sin^2 L + \beta_2 \sin^4 L) - \beta_3 h, \end{aligned}$$

and where:

\mathbf{C}_b^n : transformation DCM (Direct Cosine Matrix) from the b -frame to the n -frame, i.e. SINS attitude matrix;

$\boldsymbol{\omega}_{ib}^b, \mathbf{f}_{sf}^b$: gyro sensed angular rate and accelerometer sensed specific force, respectively;

$\mathbf{v}^n = [v_E^n \quad v_N^n \quad v_U^n]^T$: inertial navigation velocity expressed in the n -frame and v_E^n, v_N^n, v_U^n denoting velocity along east, north and up-vertical direction;

$\mathbf{p} = [\lambda \quad L \quad h]^T$: SINS positioning vector and λ, L, h being longitude, latitude and altitude above the mean sea level;

R_e, f : the Earth's semi-major axis and flattening, which are $R_e = 6,378,137$ m and $f = 1/298.257$;

ω_{ie}, g_0 : the Earth's angular rate and gravity magnitude at the equatorial sea-surface, which are $\omega_{ie} = 7.2921151467 \times 10^{-5}$ (rad/s) and $g_0 = 9.7803267714$ (m/s²);

$\beta_1, \beta_2, \beta_3$: gravity magnitude coefficients related with latitude and altitude above sea level, respectively, as $\beta_1 = 5.27094 \times 10^{-3}$, $\beta_2 = 2.32718 \times 10^{-5}$ and $\beta_3 = 2g_0/R_e = 3.086 \times 10^{-6}$ (s⁻²).

68.2.2 SINS Updating Algorithms

SINS updating algorithms, a type of dead reckoning algorithms, recursively calculate navigation information at the current moment (at time t_m) from given attitude, velocity and positioning at the previous moment (at time t_{m-1}) and the inertial sensors' outputs within the updating interval. The SINS updating interval is denoted as $T_m = t_m - t_{m-1}$. In this paper, it is assumed that the gyro sampling output is angular increment and the accelerometer output is velocity increment, and high-precision inertial navigation system always uses this type of sampling method.

For the high-precision of SINS updating algorithms and the convenience of the derivation of the inversed algorithm, the key points of SINS updating algorithms will be given with the use of 'one sample plus previous sample' error compensation method as follows [3–6].

68.2.2.1 Attitude Updating

Using the chain rule of DCM production, C_b^n at time t_m , i.e. $C_{b_m}^{n_m}$, can be constructed as

$$C_{b_m}^{n_m} = C_{n_{m-1}}^{n_m} C_{b_{m-1}}^{n_{m-1}} C_{b_m}^{b_{m-1}} \tag{68.4}$$

where

$C_{b_{m-1}}^{n_{m-1}}$ is the DCM at time t_{m-1} ;

$C_{n_{m-1}}^{n_m}$ is the transformation matrix from the i -frame to the n -frame within updating interval T_m , which can be determined by the rotation vector $\zeta_m = -\omega_{in,m-1/2}^n T_m$ and the formula is given by Eq. (68.6). Note that the subscript $m - 1/2$ denotes the middle time of $[t_{m-1}, t_m]$, i.e. time $t_m - T_m/2$ and short for $t_{m-1/2}$. The Related parameters at time $t_{m-1/2}$ can be estimated by linear extrapolation algorithm, such as $L_{m-1/2} = L_{m-1} + (L_{m-1} - L_{m-2})/2$;

$C_{b_m}^{b_{m-1}}$ is determined by rotation vector Φ_m , which is also given by Eq. (68.6). Consider the rotation vector coning compensation and use the ‘one- plus-previous sample’ gyro angular increment compensation algorithm, we have

$$\Phi_m = \Delta\theta_m + 1/12 \cdot \Delta\theta_{m-1} \times \Delta\theta_m \tag{68.5}$$

where $\Delta\theta_{m-1}$, $\Delta\theta_m$ are gyro angular increments within time interval $[t_{m-2}, t_{m-1}]$ and $[t_{m-1}, t_m]$, such that $\Delta\theta_m = \int_{t_{m-1}}^{t_m} \omega_{ib}^b dt$.

The relationship between DCM C and rotation vector V is given by

$$C = I_{3 \times 3} + \frac{\sin |\Phi|}{|\Phi|} (\Phi \times) + \frac{1 - \cos^2 |\Phi|}{|\Phi|^2} (\Phi \times)^2 \tag{68.6}$$

68.2.2.2 Velocity Updating

SINS velocity v_m^n at time t_m is recursively calculated from velocity v_{m-1}^n at time t_{m-1} , as

$$v_m^n = v_{m-1}^n + C_{n_{m-1}}^{n_m-1/2} C_{b_{m-1}}^{n_{m-1}} \Delta v_{sf,m}^{b_{m-1}} + \Delta v_{g/cor,m}^n \tag{68.7}$$

$$\Delta v_{sf,m}^{b_{m-1}} = \Delta v_m + \Delta v_{rot,m} + \Delta v_{scull,m} \tag{68.8}$$

$$\Delta v_{g/cor,m}^n = [g_{m-1/2}^n - (2\omega_{ie,m-1/2}^n + \omega_{en,m-1/2}^n) \times v_{m-1/2}^n] T_m \tag{68.9}$$

$C_{n_{m-1}}^{n_m-1/2}$ in Eq. (68.7) is determined by rotation vector $\zeta_m/2$, which is similar to the calculation formula of $C_{n_{m-1}}^{n_m}$ in attitude updating. Consider the velocity sculling

compensation and also use the algorithm of ‘one-plus-previous sample’ gyro angular and accelerometer velocity increment compensation, then in Eq. (68.8) we have

$$\Delta \mathbf{v}_{rot,m} = 1/2 \cdot \Delta \boldsymbol{\theta}_m \times \Delta \mathbf{v}_m \quad (68.10)$$

$$\Delta \mathbf{v}_{scull,m} = 1/12 \cdot (\Delta \boldsymbol{\theta}_{m-1} \times \Delta \mathbf{v}_m + \Delta \mathbf{v}_{m-1} \times \Delta \boldsymbol{\theta}_m) \quad (68.11)$$

where $\Delta \mathbf{v}_{m-1}$, $\Delta \mathbf{v}_m$ are the accelerometer specific increments within the time interval $[t_{m-2}, t_{m-1}]$ and $[t_{m-1}, t_m]$, i.e. $\Delta \mathbf{v}_m = \int_{t_{m-1}}^{t_m} \mathbf{f}_{sf}^b dt$.

68.2.2.3 Position Updating

After navigation velocity \mathbf{v}_m^n updating, position p_m at time t_m can be recursively calculated from position p_{m-1} at previous time t_{m-1} , as

$$p_m = p_{m-1} + M_{pv,m-1/2}(\mathbf{v}_{m-1}^n + \mathbf{v}_m^n)T_m/2 \quad (68.12)$$

In Eqs. (68.9) and (68.10), $\mathcal{S}_{m-1/2}^n$, $\omega_{ie,m-1/2}^n$, $\omega_{en,m-1/2}^n$, $\mathbf{v}_{m-1/2}^n$, $M_{pv,m-1/2}$ denote related calculating values at time $t_{m-1/2}$, which can also be calculated by liner extrapolation method with the navigation parameters at time t_{m-2} and t_{m-1} .

68.3 GNSS/SINS Integrated Navigation

68.3.1 GNSS/SINS Integrated Kalman Filter

The positioning accuracy of GNSS can reach centimeter level, and the carrier may experience severe maneuvers and high speed. Thus, for the navigation systems integrating of high-accuracy GNSS carrier phases and SINS, we must consider the impact of the installation lever arm error $\delta \mathbf{l}$ and time synchronization error δt between GNSS and SINS. A 19-dimension GNSS/SINS integrated Kalman filter is established, and the state vector and filter equations are listed respectively below:

$$\mathbf{x} = [\phi \quad \delta \mathbf{v}^n \quad \delta p \quad \varepsilon^b \quad \nabla^b \quad \delta \mathbf{l} \quad \delta t]^T \quad (68.13)$$

$$\dot{\mathbf{x}} = \mathbf{F}\mathbf{x} + \mathbf{G}\mathbf{w} \quad (68.14)$$

$$\mathbf{z} = \mathbf{p}_{SINS} - \mathbf{p}_{GNSS} = \mathbf{H}\mathbf{x} + \mathbf{v} \quad (68.15)$$

Given the limited space here, please refer to Refs. [6, 7] for the symbols and the system description details.

68.3.2 The Partial Feedback Principle of Kalman Filter

In order to minimize inertial navigation system errors, that is to keep the linear of the navigation error equation and improve the accuracy of filter model, the Kalman filter state estimation should constantly feed back to inertial navigation system and correct the navigation errors. For the traditional feedback method, some states are chosen from Kalman filter and feed back to the navigation system at a time, then navigation parameters are modified and all of the feedback states are reset to zero. However, the traditional all-feedback method may easily lead to navigation outputs being severe serrated, and then affect the results of inversed SINS algorithm, which leads to large intermittent pulse distortion for inertial simulator. In order to avoid the negative affect of the all-feedback method, the partial feedback principle for Kalman filter is put forward. The state partial feedback and longitude error modification, for example, can be written respectively as

$$\lambda_m = \hat{\lambda}_m - (1 - \alpha) \cdot \hat{x}_7 \quad (68.16)$$

$$\hat{x}_7^- = \alpha \cdot \hat{x}_7 \quad (68.17)$$

$$\alpha = \exp(-T_m/\tau) \quad (68.18)$$

where $\hat{\lambda}_m$ is longitude in the SINS updating algorithm, λ_m is longitude output after partial feedback, and the latter always has higher precision than the former; \hat{x}_7 is longitude error estimation of the Kalman filter, \hat{x}_7^- is the residual estimation after partial feedback, and the latter can be used to the subsequent partial feedback or as the filter initial value in the next filtering step; α is the weight coefficient for modification. Equation (68.18) shows that the weight coefficient α can be determined by a delay time parameter τ , and the larger the τ , the smoother the navigation parameter outputs. Obviously, if $\alpha = 0$, then it indicates an all-feedback method.

The measurement updating interval of GNSS/SINS integration is generally much larger than the updating interval T_m in SINS. However, if the feedback interval is consistent with SINS updating interval and the all-feedback method is used, smoother navigation parameter outputs will be achieved, and then a more realistic trajectory will be provided to the inversed navigation simulator. Surely, two filter or smoothing algorithms can also be used to further improve the accuracy of trajectory parameters [8].

68.4 Inversed SINS Algorithms

The inversed SINS algorithm can be seen as a reverse processing problem of traditional SINS updating algorithm. In the inversed algorithm, assuming that the attitude angular and positioning information are given, one for angular motion

information and another for liner motion, then gyro angular increment and accelerometer velocity increment are inversely achieved by the transformation of traditional SINS updating algorithm to realize high-precision simulation of the inertial sensor. The key points of the inversed SINS algorithm are described as follows.

68.4.1 Cubic Spline Fitting for Attitude and Position

Taking the longitude in position vector for example, fit the longitude sequence of integrated navigation output $\dots, \lambda_{m-1}, \lambda_m, \lambda_{m+1}, \dots$ as piece-wise continuous cubic spline. Assume the cubic spline fitting equation at the time interval $t \in [t_{m-1}, t_m]$ to be

$$\lambda(t) = a_{m0} + a_{m1}t + a_{m2}t^2 + a_{m3}t^3 \quad (68.19)$$

where $a_{m0}, a_{m1}, a_{m2}, a_{m3}$ are spline fitting coefficients.

The relationship between longitude and navigation east velocity is

$$\dot{\lambda}(t) = \sec L_{m-1/2} / R_{Nh, m-1/2} \cdot v_E^n(t) \quad (68.20)$$

Generally, the interval $[t_{m-1}, t_m]$ is very small, so $L(t)$ and $R_{Nh}(t)$ can be chosen as the values at time $t_{m-1/2}$ in Eq. (68.20). As the position function is known, $L_{m-1/2}$ can be achieved exactly by spline interpolation. While in the traditional SINS updating algorithm, an estimation of liner extrapolation is always used to get $L_{m-1/2}$.

Rearranging Eq. (68.20) and by the substitution of Eq. (68.19) into Eq. (68.20), the east velocity is achieved as

$$\begin{aligned} v_E^n(t) &= R_{Nh, m-1/2} \cos L_{m-1/2} \cdot \dot{\lambda}(t) \\ &= R_{Nh, m-1/2} \cos L_{m-1/2} \cdot (a_{m1} + 2a_{m2}t + 3a_{m3}t^2) \end{aligned} \quad (68.21)$$

Similarly, the north and up-vertical velocities can be achieved respectively by latitude and altitude calculation, then the position function $\mathbf{p}(t)$ and velocity function $\mathbf{v}^n(t)$ at any continuous time can be achieved. Also, if the Euler angle is manipulated by the same method, the Euler angle piece-wise fitting function can be derived, denoted as $\mathbf{A}(t)$. Note that special treatment is needed when there exists some singular points in attitude Euler angles.

68.4.2 Gyro Angular Increment Simulation

Assuming that the inertial sensor calculation interval in the inversed SINS algorithm is $T_k = t_k - t_{k-1}$, it may be different from T_m , always smaller than T_m , and the smaller T_k , the higher the calculation precision.

Take T_k as equal sampling time interval, then attitude angular sequence \mathbf{A}_k can be achieved by the interpolation of angular function $\mathbf{A}(t)$, where \mathbf{A}_k is equivalent to the attitude transformation matrix $\mathbf{C}_{b_k}^{n_k}$. By changing the time subscript from m to k and rearranging Eq. (68.4), the following relationship is established:

$$\mathbf{C}_{b_k}^{b_{k-1}} = (\mathbf{C}_{n_{k-1}}^{n_k} \mathbf{C}_{b_{k-1}}^{n_{k-1}})^{-1} \mathbf{C}_{b_k}^{n_k} \quad (68.22)$$

where $\mathbf{C}_{n_{k-1}}^{n_k}$ is determined by $\zeta_k = -\omega_{in,k-1/2}^n T_k$. As velocity and position function are known, $\omega_{in,k-1/2}^n$ can be obtained accurately by spline interpolation.

By Eq. (68.6), the rotation vector Φ_k can be inversely derived by the transformation matrix $\mathbf{C}_{b_k}^{b_{k-1}}$, then by Eq. (68.5), gyro angular increment can be obtained, as

$$\Delta\theta_k = (I_{3 \times 3} + 1/12 \cdot \Delta\theta_{k-1} \times)^{-1} \Phi_k \approx \Phi_k - 1/12 \cdot \Delta\theta_{k-1} \times \Phi_k \quad (68.23)$$

In the initialization of the inversed SINS algorithm, it can be assumed that $\Delta\theta_0 = 0$.

68.4.3 Accelerometer Velocity Increment Simulation

Also T_k is taken as sampling time interval, the interpolation of velocity cubic spline $\mathbf{v}^n(t)$ will lead to velocity sequence \mathbf{v}_k^n . By rearranging Eq. (68.7), we have

$$\Delta\mathbf{v}_{sf,k}^{b_{k-1}} = (\mathbf{C}_{n_{k-1}}^{n_{k-1/2}} \mathbf{C}_{b_{k-1}}^{n_{k-1}})^{-1} (\mathbf{v}_k^n - \mathbf{v}_{k-1}^n - \Delta\mathbf{v}_{g/cor,k}^n) \quad (68.24)$$

where the calculation of $\mathbf{C}_{n_{k-1}}^{n_{k-1/2}}$ and $\Delta\mathbf{v}_{g/cor,k}^n$ is not so complex, with no more repeat here.

In addition, substitute Eqs. (68.10) and (68.11) into Eq. (68.8), we get

$$\begin{aligned} \Delta\mathbf{v}_{sf,k}^{b_{k-1}} &= \Delta\mathbf{v}_k + 1/2 \cdot \Delta\theta_k \times \Delta\mathbf{v}_k + 1/12 \cdot (\Delta\theta_{k-1} \times \Delta\mathbf{v}_k + \Delta\mathbf{v}_{k-1} \times \Delta\theta_k) \\ &= 1/12 \cdot \Delta\mathbf{v}_{k-1} \times \Delta\theta_k + [I_{3 \times 3} + (1/2 \cdot \Delta\theta_k + 1/12 \cdot \Delta\theta_{k-1}) \times] \Delta\mathbf{v}_k \end{aligned} \quad (68.25)$$

By rearranging Eq. (68.25), it leads to

$$\begin{aligned} \Delta\mathbf{v}_k &= [I_{3 \times 3} + (1/2 \cdot \Delta\theta_k + 1/12 \cdot \Delta\theta_{k-1}) \times]^{-1} (\Delta\mathbf{v}_{sf,k}^{b_{k-1}} - 1/12 \cdot \Delta\mathbf{v}_{k-1} \times \Delta\theta_k) \\ &\approx \Delta\mathbf{v}_{sf,k}^{b_{k-1}} - 1/2 \cdot \Delta\theta_k \times \Delta\mathbf{v}_{sf,m}^{b_{k-1}} - 1/12 \cdot (\Delta\theta_{k-1} \times \Delta\mathbf{v}_{sf,k}^{b_{k-1}} + \Delta\mathbf{v}_{k-1} \times \Delta\theta_k) \end{aligned} \quad (68.26)$$

Here, it's also assumed that $\Delta\mathbf{v}_0 = 0$. Now, the accelerometer velocity increment can be obtained by Eqs. (68.24) and (68.26), and then the simulation of inertial sensor is fully completed.

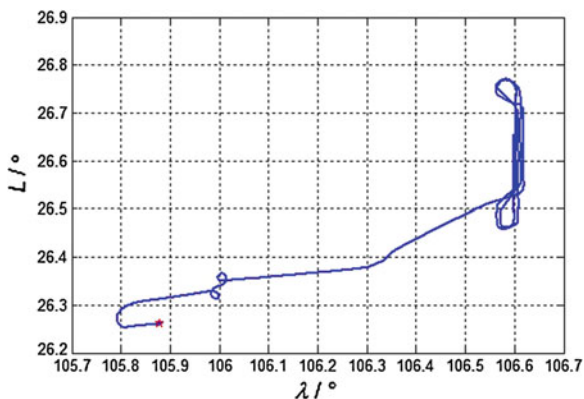
68.5 Testing Data Processing and Analysis

In this section, we provide a set of airborne navigation testing data, with a gyro constant drift of about 0.03 arcdeg/h and accelerometer bias of about 200 μ g for SINS and positioning accuracy of 5 cm for carrier phase based GPS RTK. The SINS raw data output rate is 200 Hz, while the GPS position output rate is 1 Hz. The 1 h flight profile consists of several stages, including preparation in static base, take off and climb, turning, cruise and some specific operations in working area, which are shown in Fig. 68.1 with leveling longitude-latitude projection and ‘☆’ marked as the start point.

The data processing scenario is described as follows.

1. Use two sub-sample attitude coning error and velocity sculling error compensation algorithm to update SINS and execute GPS/SINS integrated navigation Kalman filter. Meanwhile, SINS calculating errors are modified by partial feedback of state estimation from Kalman filter. The above processes can be well implemented by PSINS toolbox [9].
2. For 100 Hz integrated navigation attitude and position outputs, the cubic spline fittings are adopted to get piece-wise continuous trajectory descriptions $\mathbf{A}(t)$, $\mathbf{v}^n(t)$ and $\mathbf{p}(t)$.
3. For continuous $\mathbf{A}(t)$, $\mathbf{v}^n(t)$ and $\mathbf{p}(t)$, 200 Hz equal interval interpolations are made to get discrete serials \mathbf{A}_k , \mathbf{v}_k^n and \mathbf{p}_k .
4. The inverted SINS algorithm are used to simulate the inertial sensor sampling results $\Delta\theta_k$ and $\Delta\mathbf{v}_k$. The power spectrum density (PSD) comparisons of the IMU raw sampling data and the simulated ones are shown in Fig. 68.2. Notice that only x-axis gyro and x-axis accelerometer PSD results within 1000 s-1100 s are displayed, and the PSDs of other axis and time interval are similar then with no more showing.
5. Also based on two sub-sample error compensation algorithms, inertial sensor simulated samples $\Delta\theta_k$ and $\Delta\mathbf{v}_k$ are used to update pure SINS with altitude

Fig. 68.1 Leveling projection for airborne flight profile



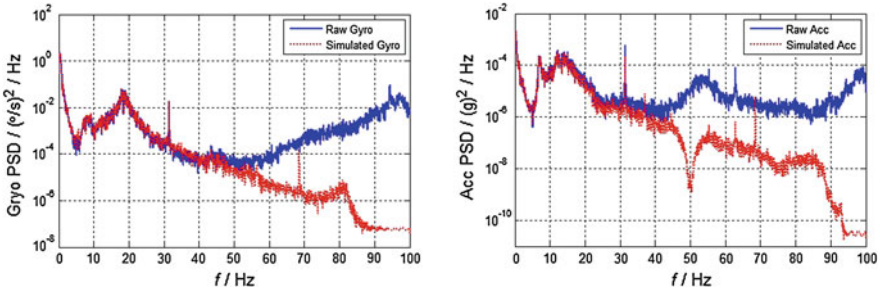


Fig. 68.2 PSD comparison for IMU raw data and simulated sampling data

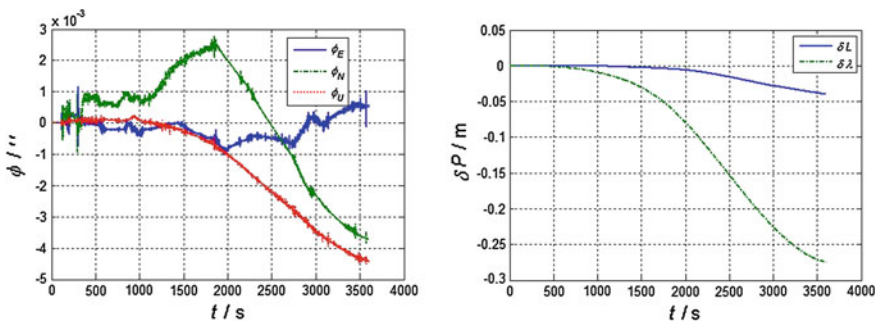


Fig. 68.3 SINS attitude and position errors based on simulated sampling data

damping. The attitude and position comparisons for pure SINS and integrated navigation outputs from step 1 are shown in Fig. 68.3.

Figure 68.2 demonstrates that the PSD of IMU raw sampling date well coincides with that of simulated sampling data in low frequency band, and it reveals that the inertial sensor simulator can simulate the airborne flight characteristic. Figure 68.3 shows the simulated sampling data is of high accuracy, with attitude error of 0.001arcsec and position error of 0.1 m within 1 h pure SINS.

68.6 Conclusions

Based on traditional high-accuracy SINS numerical updating algorithm, whose attitude, velocity and position navigation outputs are updated by inertial sensor sampling data, the inversed SINS algorithms are proposed by using attitude and position information to reversely produce simulated sampling data for inertial sensor simulator. For actual airborne flight navigation testing data, the techniques of GNSS integrated Kalman filtering and SINS error modification by partial feedback

are applied to improve navigation accuracy and get smooth trajectory parameters, which are essential for a high precision inertial sensor simulator. The testing data processing reveals that the simulator can reach a positioning accuracy of about 0.1 m within 1 h standalone SINS, which meets the high-accuracy SINS requirements for inertial sensor simulated data source. Hopefully, further decrease of simulated errors can be made by improving the simulator updating frequency. If higher frequency coincidence between IMU raw data and simulated sampling data is demanded, the increasing of the sampling frequency for raw data or using single-sample SINS updating algorithm can be applicable.

Furthermore, when using high-accuracy simulated sampling data for specific navigation task, the sensor's errors, such as stochastic errors and misalignment errors, can be modeled and added into the ideal data to implement more complex inertial navigation or integrated navigation simulations.

References

1. James H, Robert G (2014) Radio Astro Inertial Doppler Error Simulator. <http://www.industrycortex.com>. Accessed 20 Dec 2014
2. GPSoft LCC. Inertial Navigation System Toolbox 3.0 for Matlab User's Guide. <http://gpssoftnav.com>. Accessed 20 Dec 2014
3. Paul GS (1998) Strapdown inertial navigation integration algorithm design Part 1: attitude algorithms. *J Guid Control Dyn* 21(1):19–28
4. Paul GS (1998) Strapd-n Part 2: velocity and position algorithms. *J Guid Control Dyn* 21(2):208–221
5. Yuanxin W, Jinling W, Dewen H (2014) A new technique for INS/GNSS attitude and parameter estimation using online optimization. *IEEE Trans Signal Process* 62(10):2642–2655
6. Gongmin Y (2006) Research on vehicular position and orientation system. Ph.D thesis, Northwestern Polytechnical University, Xi'an, China
7. Chunlian Z (2014) Research on precise POS navigation algorithm in aerial photogramme try. Master thesis, Northwestern Polytechnical University, Xi'an, China
8. Wanliang H, Fuping S, Jianyong C, Po C (2013) Application of two-filter smoothing to GNSS/INS integrated navigation system. In: The 4th China satellite navigation conference
9. Gongmin Y. Precise Strapdown Inertial Navigation System (PSINS) Toolbox for Matlab. http://blog.sina.com.cn/s/blog_40edfdc90101heg0.html. Accessed 20 Dec 2014

Chapter 69

A Research on All Source Navigation and Positioning and Its Critical Technology

YongBin Zhou, Jun Lai, XiYe Guo and Jun Yang

Abstract Due to the increased complexity of the application scenario, the conventional navigation and positioning system, which uses a single method, can't meet the challenge of accuracy and reliability. Comparing with multi-sensor fusion, All Source Navigation and Positioning (ASPN) provides an all-in-one solution. It uses a uniform architecture to achieve data fusion of numerous sensors and data filtering with the ability of re-configuration. It tries to provide the maximum accuracy and reliability when some navigation and positioning methods, especially the GNSS, are not available. This paper mainly introduces the ASPN's concept and its development status. It also analyses and predicts the critical technology of ASPN.

Keyword ASPN plug-and-play

69.1 Introduction

Navigation and positioning services are playing an irreplaceable role in modern Military and civil areas. Due to the increased complexity of the application scenario, the conventional navigation and positioning system, which uses a single method, can't meet the challenge of accuracy and reliability. Multi-sensor Navigation and Positioning are becoming more popular. ASPN (All Source Positioning and Navigation) is based on Multi-sensor Navigation and Positioning. It aims at providing a uniform method and framework for multiple navigation technology. It tries to provide navigation and positioning service with maximum accuracy and reliability and adapt the change of sensor's performance and of the availability of each sensor. This paper mainly introduces the ASPN's concept and its development status. It also analyses and predicts the critical technology of ASPN.

Y. Zhou · J. Lai (✉) · X. Guo · J. Yang
Department Mechanic Engineering and Automation, NUDT, Changsha, China
e-mail: mfary@139.com

69.2 Background of ASPN

In the Command Automation System, C4ISR, proposed by the United States, navigation and positioning are important means of providing information for command and control. It can be predicted that the reliability problem with navigation system of combat platform will remain one of the eternal themes about military navigation applications.

The variety of navigation and positioning sensors, including global navigation satellite systems (GNSS, global navigation satellite system), inertial navigation system, magnetic fields, gravity, imaging, odometry, mobile phone base stations, radio beacons, ground-based navigation system. But military use commonly relies on GNSS. In the United States, For example, almost all precision-guided munitions have used GPS targeting data. GPS receivers are also mounted on ships, aircrafts, tanks, armored vehicles, and transport vehicles. Corresponding to much convenient, the GPS dependence becomes an essential problem. GPS has its inherent vulnerability. Its signal power reaching to the earth surface is merely -160 dBW. Obstacles can greatly fade the signal, making it nearly impossible being using in under-water, underground or indoor scenario. Meanwhile, an improvised jamming device is able to cause a deficiency of navigation service in a vast area.

It is achievable to give the error position and time information to GPS receiver by using deception jamming. In June 2012, an experiment organized by the U.S. Department of Defense shows the ability of Texas University to disable a UAV's navigation system only using a deception device that costs less than \$1000 [1]. Restricted by the defects of the sensor as well as the external environment, a single navigation method often has a lot limitations in practical.

Therefore, it becomes popular to fusing multi sensors' data for a higher reliability and accuracy. However, the conventional fusion system can't tolerate the dynamic change of sensor availability, so the available scenario is restricted. ASPN is proposed to solve the problem.

69.3 Research Status

69.3.1 Abroad Research Status

At present, several universities and research institutions in the United State have being studying the concepts, techniques, and trends of ASPN and obtained some initial results.

In November 2010, the concept of ASPN was brought up by DARPA (Defense Advanced Research Projects Agency). The ASPN program was announced aimed at developing a low cost navigation sensor fusion technologies, and achieving a Plug-and-Play architecture [2]. In the initial phase of the ASPN program, scientists from Draper Lab and Argon ST concentrated on developing the architectures,

abstraction method, and navigation filtering algorithms necessary for rapid navigation sensors integration and reconfiguration [3].

In 2011, the United States air force doctor Fisher stated United States air force that only by using ASPN approach, GPS precise navigation in blocked environment issues can be resolved [4]. Elsner [5] proposed a software abstraction method to achieving Plug-and-Play.

In June 2012, DARPA launched ASPN phase II project, whose goal is to demonstrate and evaluate adaptive navigation system, optimization performance and real-time operation, showing a plug-and-play navigation platform with these attribute: low-cost, low-power-consumption and small size [6].

MIT Laboratory and the Georgia Institute of technology proposed an ASPN algorithm framework based on Factor Graph. This framework includes Constrained Optimal Sensor Selection, Incremental Smoothing, and Multi-Sensor Data Fusion Based on Factor Graph [7, 8]. They have tested this with 19 different sensor types (total 57 sensors: each sensor type includes multiple kinds of sensors) mounted on dismount, ground, and aerial platforms. Now it can estimate navigation and positioning state in a fix-lag and nonlinear method [9].

69.3.2 Domestic Research Status

Comparing to the abroad research status, the domestic research is mainly in the data fusion of inertial system, satellite navigation system or stereo vision. Research on fusion of more than four sensors is still at theoretical phase. The study on ASPN is also on tentative phase.

69.4 The Critical Technology of ASPN

The section analyses and predicts the critical technology of ASPN, with which the APSN is able to match its purpose.

The comparison between APSN and conventional multi-sensor positioning is shown as Table 69.1. The biggest difference between ASPN and conventional multi-sensor positioning is that the former not only should solve the multi-sensor fusion problem, but also provide a uniformed sensor interface and uniformed navigation filtering algorithm to adapt the dynamic change of available sensors. Besides, it also should be able to reduce the volume and power consuming. Specifically, the critical technology of ASPN includes Uniformed Interface Technology for Navigation Sensors, Modular Platform Technology, Optimal Sensors Selection Algorithm Based on Sensor Data, and Real-time Navigation Data Fusion and Filtering in Dynamic Condition.

Table 69.1 Comparison between ASPN and multi-sensor navigation and positioning

Property	ASPN	Multi-sensor positioning
Number of available sensors	Dynamic	Static
Support for plug-and-play	Yes	No
Interface uniformity	Yes	No
Fusion and filtering algorithm	Uniform	Private
Scenario adaptability	Good	Poor
Reliability	High	Low

69.4.1 Uniformed Interface Technology for Navigation Sensors

ASPN requires Plug-and-Play property, so a uniformed interface for navigation sensors is required, which is responsible for data acquisition and sensor configuration. This interface can be divided into hardware and software abstract interface.

The hardware interface should use as few kinds of interface types, such as LAN, CAN, etc., as possible to guarantee the consistency of ASPN. Various sensors have various data acquisition speed and data size, making more difficult to design the interface with a flexible bandwidth, which should be able to adapt to the various data transition speed, and remain low-power consumption. The fusion problem requires the interface to provide unified standards of space and time and low latency attribute.

For software abstract interface, the sensor description data format, navigation data format, sensor status and control data format and standards of space and time. There should be time flag for each navigation data as the asynchronous feature.

69.4.2 Modular Platform Technology

A typical ASPN platform is shown as Fig. 69.1. As the former section figured out, different navigation scenarios, as well as dynamic access of hardware sensors can cause bandwidth and processing load of sensor interface as well as navigation platform dynamics. In pursuit of the hardware platform consistency, it is a waste that using abundant bandwidth and computation resources to uniform the “easy scenario” and “difficult scenario” platform. To solve the conflict between these scenario requirements, it is a wise idea to modular the resources of ASPN platform.

There are two method being proposed to meet the goal. One of them is inserting a Co-processing module between the sensor interface and sensor that requires high bandwidth or much computation resource, as Fig. 69.2 shows. For example, if using the stereo vision as the sensor, the most of the algorithm about image will be done in the co-processing module, and the sensor interface should merely receive simple data type such as velocity, angular speed, time flag and so on, saving resources greatly.

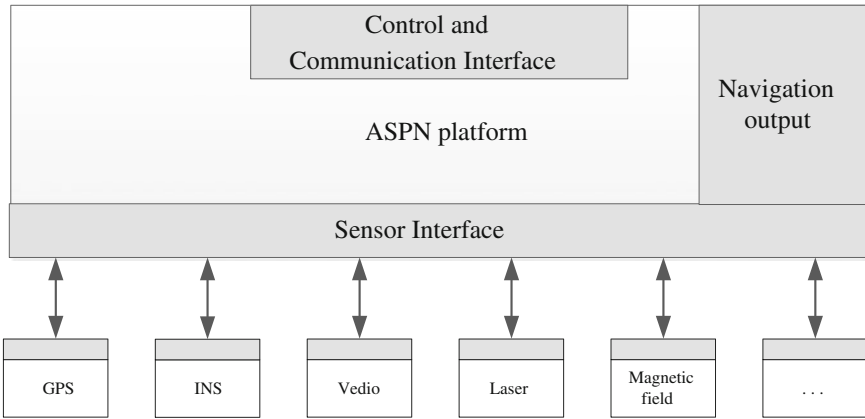
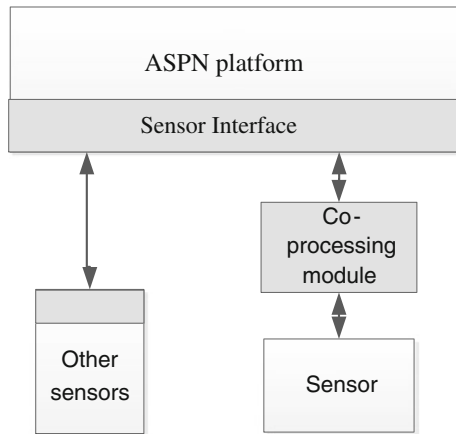


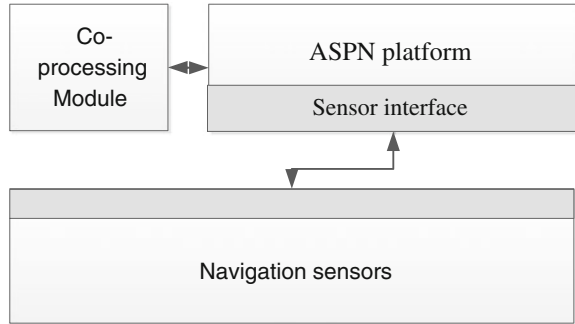
Fig. 69.1 Hardware framework of ASPN

Fig. 69.2 Co-processing module between interface and sensors



Another method is regarding the co-processing module mounted as a resource independent from the ASPN platform as Fig. 69.3 shows. When requiring extra computing resource, sensor data can be transmitted to Co-processing module by the ASPN platform. The advantage of this method is reducing the design difficulty of the Co-processing module, meanwhile improving the modularity. Furthermore, the Co-processing module can be used between several ASPN platforms, thus the ASPN platform can be made to a smaller, lighter embedded platform, which is essential for individual soldier use. But this method has the inherent deficiency that it can't reduce the bandwidth load.

Fig. 69.3 Co-processing module separated from the framework



69.4.3 *Optimal Sensors Selection Algorithm Based on Sensor Data*

Even when the sensors have been connected to the ASPN platform, the data validity is still unknown. For example, when be jammed, the GPS signal can't be used to navigate and position. In addition, in some scenario not requires very high performance, some navigation sensor can be turned off for saving energy. To achieve these demands, an optimal sensor selection algorithm should be constructed.

Groves et al. [10] emphasized the importance of environmental perception for multi-sensor positioning and navigation. Although environmental perception is an optimal solution to abundant scenario positioning. But to ASPN, it means that the sensor selection should be done after finishing environmental perception, which requires analyzing sensor data and priori knowledge. So, the environmental perception is a cumbersome step. Chiu et al. [8] shows a good example, they used sensor data and some priori knowledge to classify the sensors available in the multi-sensor platform, combined with the constrained condition, and finally get the optimal sensors collection. The Sensor data they used is real-time data from each sensor in a short period. But this method can't analyze sensor data in a moving platform, meanwhile, it can't classify some kinds of sensors which provide information like velocity and angular speed.

69.4.4 *Real-Time Navigation Data Fusion and Filtering in Dynamic Condition*

The multi-sensor data fusion navigation and positioning is the basis of the ASPN. There are some challenges in it. First of all, in multi-sensor system, there are a lot of differences between each kind of sensor in data type, rate, and confidence features. Especially, there could be data conflict problem, making the increase of ambiguous, and less confidence of navigation result. Besides, the data fusion approach can be classify into centralized, distributed and hybrid integration types, each of which has

its own advantages and limitations. It is challenge that ASPN wants to have all the advantages and avoid all limitations.

The main substance of multi-sensor data fusion problem is an optimal estimation problem to the joint random variables. In a static system, it is practical to use weighted mean method, LS method. If priori probability is known, then MLE, MMSE can be used. In a dynamic system, Kalman filter can be used in linear estimation. In nonlinear estimation, There are Extended Kalman filter, the UKF, GSF methods [11, 12]. Based on probability distribution, the Bayesian estimation can be used theoretical in any state with a general distribution. To reduce the computational complexity of traditional Bayesian estimation, the Markov Chain Monte Carlo (MCMC) and Particle filter based on random sampling is proposed [13–15].

Plug-and-Play feature not only influences the design of the ASPN platform, but also the design of the fusion algorithms. Traditional fusion algorithm doesn't consider to optimize for the feature. Furthermore, the algorithm's performance is unclassified.

MIT and the Georgia Institute of technology proposed a multi-sensor fusion navigation and positioning algorithm based on Factor Graph in [9]. This algorithm exploits the easy-to-extend and non-linear estimating support features of Factor Graph and an incremental filtering method, and finally expresses the process of data-acquiring and filtering in a natural way. Its navigation accuracy is higher than the EKF [16].

69.5 Conclusions

ASPN creates a new way to provide a high-accuracy positioning and navigation service in a complex and dynamic environment. It guides the future's development of navigation applications. This paper starts with the concept and background of ASPN, then compares the differences between ASPN and multi-sensor fusion poisoning. It also analyses and predicts the critical technology of ASPN. Equipped with these technologies, ASPN can meet requirements in power-consumption, volume, accuracy and reliability. But the development of ASPN has lots of possibilities, such as sensor online calibration, collaboration of multiple ASPN platforms in single carrier. Even using ASPN platform as signal beacon to aid navigation is a critical usage.

References

1. UAVs Vulnerable to Civil GPS Spoofing (2012). <http://www.insidegnss.com/node/3131>. Accessed 11 Dec 2014
2. Liu, B-C (2013) All Source Positioning and Navigation for the US army. *Int Space* 46–49
3. Keller J (2013) Three join DARPA ASPN navigation sensor-fusion program to reduce U.S. need for GPS. <http://www.militaryaerospace.com/articles/2013/02/Three-join-ASPN.html>. Accessed 11 Dec 2014

4. Fisher KA, Raquet JF (2011) Precision position, navigation, and timing without the global positioning system, DTIC Document
5. Elsner DL (2012) Universal plug-n-play sensor integration for advanced navigation, DTIC Document
6. Broad Agency Announcement All Source Positioning and Navigation (ASPN) Phase 2, DARPA, Editor 2012
7. Chiu HP et al (2013) Robust vision-aided navigation using sliding-window factor graphs. In: IEEE international conference on robotics and automation (ICRA), 2013, IEEE
8. Chiu HP et al (2014) Constrained optimal selection for multi-sensor robot navigation using plug-and-play factor graphs. In: Robotics and Automation (ICRA), 2014, IEEE International Conference on (pp. 663–670). IEEE
9. Indelman V et al (2012) Factor graph based incremental smoothing in inertial navigation systems. In: 2012 15th international conference on information fusion (FUSION), 2012, IEEE
10. Groves PD et al (2014) The four key challenges of advanced multisensor navigation and positioning
11. Alspach DL, Sorenson HW (1972) Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Trans Autom Control* 17(4):439–448
12. Han CZ, Zhu HY, Duan ZS (2006) Multi-sensor information fusion. Tsinghua University Press Inc, Tsinghua
13. Doucet A, Logothetis A, Krishnamurthy V (2000) Stochastic sampling algorithms for state estimation of jump Markov linear systems. *IEEE Trans Autom Control* 45(2):188–202
14. Neal RM (1993) Probabilistic inference using markov chain monte carlo methods (Technical Report CRG-TR-93-1). University of Toronto
15. Arulampalam MS et al (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans Sig Process* 50(2):174–188
16. Kaess M et al (2012) ISAM2: Incremental smoothing and mapping using the Bayes tree. *Int J Robot Res* 31(2):216–235

Chapter 70

Mixed Interacting Filter for Tracking with Multiple System and Model

Xiaoguang Zhang, Dongyan Wei, Ying Xu and Hong Yuan

Abstract Due to the complicated and changeable environment for moving target. Tracking is difficult through single system to observe. And single motion model cannot describe the moving state for changeable state. On Bayesian estimation theorem, the mixed interacting multiple system and model filter algorithm for tracking is proposed (IMSM). Its performance is better than single dimension interacting filter. Finally simulation results show the effectiveness of the proposed algorithm.

Keywords Target tracking · Mixed interacting filter · Multiple system · Multiple model

70.1 Introduction

Due to the changeable state of maneuvering target, accurate tracking the maneuvering target with single model may be difficult. Based on the generalized Bayesian theorem, the interacting multiple model algorithm (IMM) has been proposed by

Project interacting multiple system and multiple model positioning data Fusing supported by the innovation program of academy of Opto-Electronics (AOE), Chinese academy of science (CAS).

X. Zhang (✉) · D. Wei · Y. Xu · H. Yuan
Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing 100080, China
e-mail: zxcg@aoe.ac.cn

D. Wei
e-mail: weidiongyan@aoe.ac.cn

Y. Xu
e-mail: nadinexy@aoe.ac.cn

H. Yuan
e-mail: yuanh@aoe.ac.cn

Blom and Bar-Shalom. Parallel filters of multiple model and switching models, described by one-step Markov Chain, have been applied in the IMM algorithm to inaccurate estimation caused by model mismatching [1, 2].

From then on, the study about IMM has been emerged endlessly. In paper [3, 4], particle filter has been introduced into IMM algorithm to avoid error caused by linearization and limitation of the Gaussian noise. To solve motivation model mismatching problem of single module set, multiple module sets switching algorithm has been introduced in paper [5]. A new IMM algorithm with adaptive Markov transition probabilities, which is based on the innovation, has been proposed in paper [6], and the performance has been improved obviously. Besides the improvement to IMM, its theorem has also been used in the study of multiple sensors cooperation. In paper [7, 8] the interacting multiple sensor algorithm has been proposed and used for target tracking under complex environment.

But for target tracking, the performance is influenced by maneuvering state changing and fast environment changing. In this paper, based on IMM theorem, two-dimension mixed interacting filter with multiple models and system has been proposed.

70.2 System Modeling

Assuming the maneuvering target is tracked with n models and m observation system. At time sampling t , the state equation for maneuvering target can be written as:

$$\mathbf{x}(t) = A^p \hat{\mathbf{x}}^p(t-1) + G^p w^p(t) \quad (70.1)$$

where $\hat{\mathbf{x}}^p(t)$ denotes the state vector of target under model p . A^p denote the transforming matrix from one moment to the next time, and $w^p(t)$ denotes Gauss white noise with mean 0 and covariance matrix $Q_p(t)$. The observing equation corresponding to sensor I at time t can be described as:

$$\mathbf{z}_i(t) = H_i \mathbf{x}(t) + v_i(t) \quad (70.2)$$

where $\mathbf{z}_i(t)$ is the observing value. H_i is measuring matrix. Cumulative set for measuring value can be described as:

$$\mathbf{Z}^t = \{\mathbf{Z}(i)\}_{i=1}^t \quad (70.3)$$

70.3 Calculation Steps of the Algorithm

70.3.1 Mixed Probability Calculation of Multiple Models

Assuming model l is a match for target at time $t-1$, and at the next time t it is model p . And the course can be described using one-step markov chain. The switching probability is:

$$P^\mu(l, p) = P\{M^p(t)|M^l(t-1)\} \quad (70.4)$$

The total number of models is m . We have $\sum_{p=1}^m P^\mu(l, p) = 1$, so the mixed probability of multiple model is

$$\begin{aligned} \mu^{p|l}(t-1|t-1) &= P\{M^l(t-1)|M^p(t), \mathbf{Z}^{t-1}\} \\ &= \frac{1}{\bar{c}^p} P^\mu(p, l) \mu^l(t-1) \end{aligned} \quad (70.5)$$

where $\bar{c}^p = \sum_{l=1}^m P^\mu(p, l) \mu^l(t-1)$ is the normalizing factor.

70.3.2 The Initial State After Mixing Operation

$$\hat{\mathbf{x}}^{op}(t-1|t-1) = \sum_{l=1}^m \mu^{p|l}(t-1|t-1) \hat{\mathbf{x}}^l(t-1|t-1) \quad (70.6)$$

$$\begin{aligned} P^{op}(t-1|t-1) &= \sum_{l=1}^m \mu^{p|l}(t-1|t-1) \{P^l(t-1|t-1) \\ &\quad + (\hat{\mathbf{x}}^{op}(t-1|t-1) - \hat{\mathbf{x}}^l(t-1|t-1)) \\ &\quad \times (\hat{\mathbf{x}}^{op}(t-1|t-1) - \hat{\mathbf{x}}^l(t-1|t-1))^T\} \end{aligned} \quad (70.7)$$

70.3.3 Multiple System Interacting Filter

70.3.3.1 Innovation Information Calculation

For different observing system under certain model, $\hat{\mathbf{x}}_i^p(t|t-1)$ and $P^p(t|t-1)$ is the same:

$$\hat{\mathbf{x}}^p(t|t-1) = \mathbf{A}_p \hat{\mathbf{x}}^{op}(t-1|t-1) \quad (70.8)$$

$$P^p(t|t-1) = \mathbf{A}_p P^{op}(t-1|t-1) (\mathbf{A}_p)^T + G_p Q_p (G_p)^T \quad (70.9)$$

The corresponding innovation information and covariance matrix can be described as:

$$v_i^p(t) = \mathbf{z}_i(t) - H_p \hat{\mathbf{x}}^p(t|t-1) \quad (70.10)$$

$$S_i^p(t) = \mathbf{H}_i(t) P^p(t|t-1) (\mathbf{H}_i(t))^T + R_i(t) \quad (70.11)$$

70.3.3.2 The Updated System Probability Calculation

According Bayesian theorem, at time t, system ε_i probability can be described as:

$$\begin{aligned} \beta_i^p(t) &= p\{\varepsilon_i(t) | z_i(t), M^p(t), \mathbf{Z}^{t-1}\} \\ &= \frac{1}{c_p} p\{z_i(t) | \varepsilon_i(t), M^p(t), \mathbf{Z}^{t-1}\} \\ &\quad \times p\{\varepsilon_i^p | M^p(t), \mathbf{Z}^{t-1}\} \end{aligned} \quad (70.12)$$

where $c^p = p\{z_i(t) | M^p(t), \mathbf{Z}^{k-1}\}$ is normalization factor. For simplicity, we assume the observing noise is white Gauss noise. So under model p, measuring value $z_i(t)$ probability, which also be called likelihood value, is

$$\begin{aligned} L_i^p(t) &= p\{z_i(t) | \varepsilon_i(t), M^p(t), \mathbf{Z}^{t-1}\} \\ &= N\{v_i^p(t), S_i^p(t)\} \\ &= |2\pi S_i^p(t)|^{\frac{1}{2}} \\ &\quad \times \exp\left[-\frac{1}{2} (v_i^p(t))^T (S_i^p(t))^{-1} v_i^p(t)\right] \end{aligned} \quad (70.13)$$

$p\{\varepsilon_i(t) | M^p(t), \mathbf{Z}^{t-1}\}$ is the predicting probability of system i. And according to Chapman-Kolmogorov theorem, c_p and $\beta_i^p(t|t-1)$ can be described as:

$$\begin{aligned} c_p &= \sum_{i=1}^n p\{z_i(t) | \varepsilon_i(t), M^p(t), \mathbf{Z}^{t-1}\} \\ &\quad \times p\{\varepsilon_i(t) | M^p(t), \mathbf{Z}^{t-1}\} \end{aligned} \quad (70.14)$$

$$\begin{aligned}
\beta_i^p(t|t-1) &= p\{\varepsilon_i(t)|M^p(t), \mathbf{Z}^{t-1}\} \\
&= \sum_{j=1}^n p\{\varepsilon_i(t)|\varepsilon_j(t-1), M^p(t), \mathbf{Z}^{t-1}\} \\
&\quad \times p\{\varepsilon_j(t-1)|M^p(t), \mathbf{Z}^{t-1}\} \\
&= \sum_{j=1}^n P_{\beta}(i,j)\beta_j(t-1)
\end{aligned} \tag{70.15}$$

where $P_{\beta}(i,j)$ is switching probability from system j at time $t-1$ to system i at time t .

$$\sum_{j=1}^n P_{\beta}(i,j) = 1 \tag{70.16}$$

So under model p , the observing system probability can be updated as:

$$\beta_i^p(t) = \frac{L_i^p(t)\beta_i^p(t|t-1)}{\sum_{i=1}^n L_i^p(t)\beta_i^p(t|t-1)} \tag{70.17}$$

70.3.3.3 Multiple System Filter and State Update

The corresponding gain factor is:

$$K_i^p(t) = P^p(t|t-1)\mathbf{H}_i(t)[S_i^p(t)]^{-1} \tag{70.18}$$

So according certain model p , the state and error covariance estimation can be described as:

$$\begin{aligned}
\hat{\mathbf{x}}^p(t|t) &= E[\mathbf{x}^p(t|t)|\mathbf{Z}^t] \\
&= \sum_{i=1}^n E[\mathbf{x}^p(t|t)|\varepsilon_i(t), M^p(t), \mathbf{Z}^t] \\
&\quad \times p\{\varepsilon_i(t)|M^p(t), \mathbf{Z}^t\} \\
&= \sum_{i=1}^n \{\hat{\mathbf{x}}^p(t|t-1) + K_i^p(t)v_i^p(t)\}\beta_i^p(t) \\
&= \hat{\mathbf{x}}^p(t|t-1) + \sum_{i=1}^n \beta_i^p(t)K_i^p(t)v_i^p(t)
\end{aligned} \tag{70.19}$$

$$\begin{aligned}
 P^p(t|t) &= P^p(t|t-1) \\
 &\quad - \sum_{i=1}^n \beta_i^p(t) K_i^p(t) S_i^p(t) (K_i^p(t))^T \beta_i^p(t)
 \end{aligned} \tag{70.20}$$

70.3.4 Multiple Model Fusing and State Update

$$\begin{aligned}
 \hat{\mathbf{x}}(t|t) &= \sum_{p=1}^m E[\mathbf{x}(t|t)|M^p(t), \mathbf{Z}^t] p\{M^p(t)|\mathbf{Z}^t\} \\
 &= \sum_{p=1}^m \hat{\mathbf{x}}^p(t|t) \mu^p(t)
 \end{aligned} \tag{70.21}$$

$$\begin{aligned}
 P(t|t) &= \sum_{p=1}^m \mu^p(t) \{P^p(t|t) \\
 &\quad + (\hat{\mathbf{x}}^p(t|t) - \mathbf{x}(t|t))(\hat{\mathbf{x}}^p(t|t) - \mathbf{x}(t|t))^T\}
 \end{aligned} \tag{70.22}$$

According to Bayesian theory, $\mu^p(t)$ can be written as:

$$\begin{aligned}
 \mu^p(t) &= p\{M^p(t)|\mathbf{Z}^t\} \\
 &= \frac{1}{c} p\{\mathbf{z}(t)|M^p(t), \mathbf{Z}^{t-1}\} \\
 &\quad \times p\{M^p(t)|\mathbf{Z}^{t-1}\}
 \end{aligned} \tag{70.23}$$

$$p\{M^p(t)|\mathbf{Z}^{t-1}\} = \bar{c}^p(t) \tag{70.24}$$

$$\begin{aligned}
 p\{\mathbf{z}(t)|M^p(t), \mathbf{Z}^{t-1}\} &= \sum_{i=1}^n p\{\mathbf{z}(k)|\varepsilon_i(t), M^p(t), \mathbf{Z}^{t-1}\} \\
 &\quad \times p\{\varepsilon_i(t)|M^p(t), \mathbf{Z}^{t-1}\} \\
 &= \sum_{i=1}^n L_i^p(t) \times \beta_i^p(t|t-1)
 \end{aligned} \tag{70.25}$$

$$\mu^p(t) = \frac{1}{c} \times \bar{c}^p(t) \times \sum_{i=1}^n L_i^p(t) \times \beta_i^p(t|t-1) \tag{70.26}$$

$$c = \sum_{p=1}^m \left(\bar{c}^p(t) \times \sum_{i=1}^n L_i^p(t) \times \beta_i^p(t|t-1) \right) \tag{70.27}$$

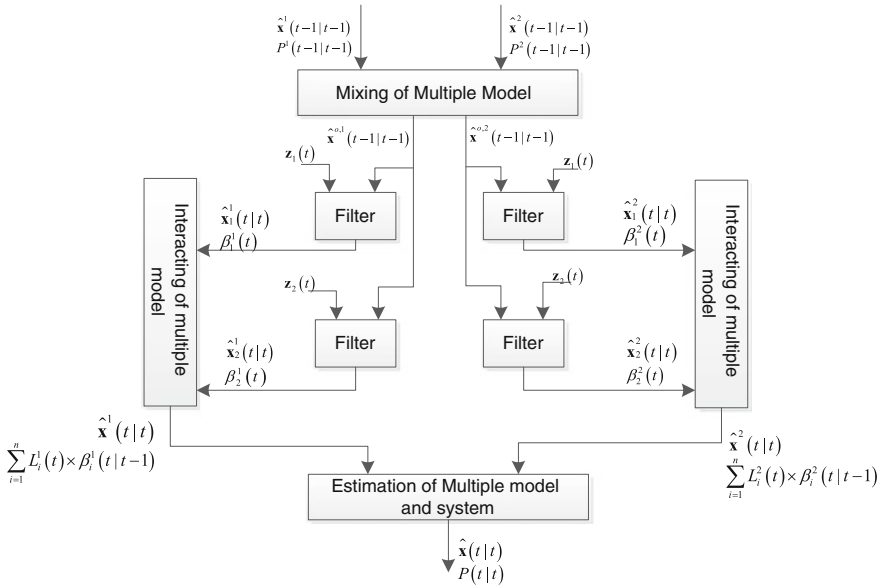
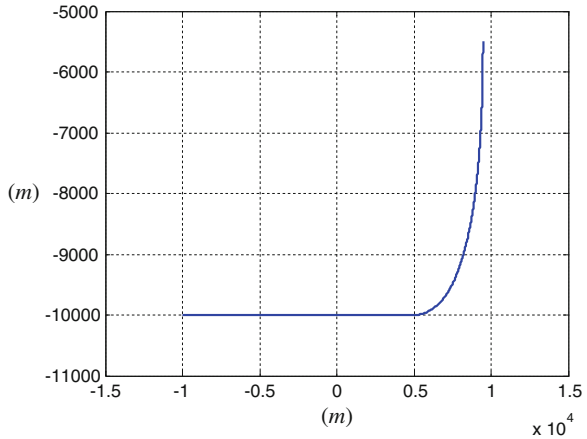


Fig. 70.1 Diagram of mixed tracking filter

70.4 Simulation Analysis

The trajectory of object is diagramed in Fig. 70.1. The two CA models with acceleration noise with mean 0.5, 5 m/s² are taken. And there are three system for observing, in stage 0–200, 200–600, 600–800, with noise mean 1 m, 10 m, and the switching matrix of model and system are consistent (Fig. 70.2).

Fig. 70.2 Track of the target



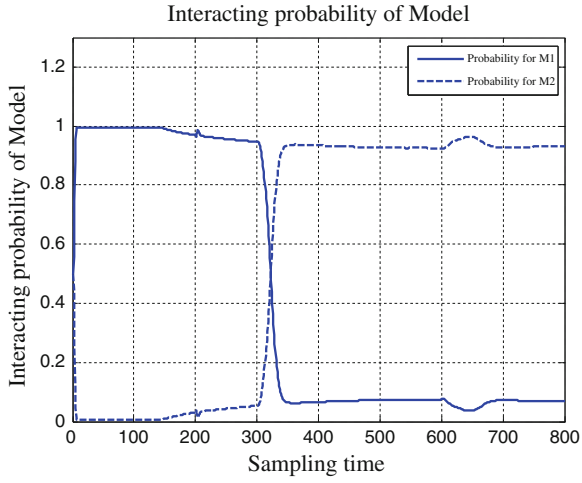


Fig. 70.3 Interacting probability of models

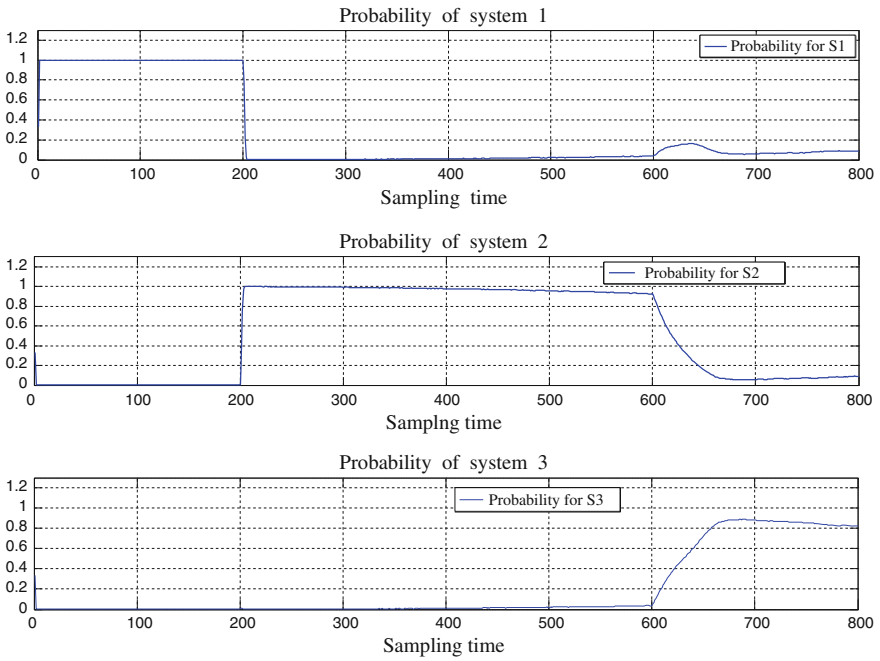


Fig. 70.4 Sensor probability on model I

$$P_u = \begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix}, \quad P_b = \begin{bmatrix} 0.98 & 0.01 & 0.01 \\ 0.01 & 0.98 & 0.01 \\ 0.01 & 0.01 & 0.98 \end{bmatrix}$$

From Figs. 70.3 and 70.4, the algorithm can help the decisive model switch among different model, and also help decisive system switch among different system adaptively.

70.5 Conclusion

On the basis of IMM algorithm, a mixed interacting multiple model and system filter is designed, and the algorithm has been detected. Simulation result show the algorithm can help the decisive model and observing system switch between different models and system adaptively.

References

1. Blom HAP, Bar-shalom Y (1988) The interacting multiple model algorithm for systems with arkovian switching coefficients. *IEEE Trans Autom Control* AC-33(8):780–783
2. Mazor E, Averbuch A, Bar-shalom Y, Dayan J (1988) Interacting multiple model methods in target tracking: a survey. *IEEE Trans Aerosp Electron Syst* 34(1):103–123
3. Boers Y, Driessen JN (2003) Interacting multiple model particle filter. *IEEE Proc Radar Sonar Navig* 150(5):344–349
4. Yang N, Tian W, Jin Z (2006) An interacting multiple model particle filter for maneuvering target location. *Meas Sci Technol* 17(6):1307–1311
5. Can S, Jianping X, Haozhe L, Wei W, Yanbo Z (2012) S-IMM: switched IMM algorithm for maneuvering target tracking. *J Convergence Inf Technol* 14(7):461–468
6. Luo X, Wang H, Li X (2005) Interacting multiple model algorithm with adaptive markov transition probabilities. *J Electron Inf Technol* 27(10)
7. Liu Z, Wang J (2012) Interacting multiple sensor filter for sensor networks. *J Electron Inf Technol* 40(4)
8. Liu Z, Wang J, Xue Y (2012) Interacting multiple sensor filter. *Sig Process* 92:2180–2186

Chapter 71

Research on Multi-Source Fusion Based Seamless Indoor/Outdoor Positioning Technology

Ying Xu, Hong Yuan, Dongyan Wei, Qifeng Lai, Xiaoguang Zhang and Weina Hao

Abstract GNSS has been applied widely. Yet because satellite signals are vulnerable and susceptible to blockage, the operability of GNSS in urban canyons are greatly hampered and GNSS even proves useless for indoor settings. This paper proposes system architecture for the integration of WLAN fingerprinting, visual positioning, baroreceptor-derived altitude estimation and GNSS for seamless indoor/outdoor positioning for vehicles and pedestrians. This architecture augments GNSS through the integration of terminal-side/network-side positioning and position/measurement domain. After temporal and spatial synchronization, data from each sensor is filtered by sub filters and then processed by the main filter. The purpose of these operations is to provide accurate and continuous estimates of positions. Tests conducted in the new technology center of CAS show that the architecture proposed can achieve seamless indoor/outdoor positioning, with a better accuracy performance than any single-source method as the former still maintains accuracy and continuity when the later generates noticeable errors. Calculation shows that multi-source fusion has an accuracy level of better than 1 m (outdoor)/3 m (indoor), hence capable of meeting users' demand for seamless indoor/outdoor positioning.

Keywords Seamless indoor/outdoor · Multi-source fusion · Decentralized kalman filtering

Y. Xu (✉) · H. Yuan · D. Wei · Q. Lai · X. Zhang · W. Hao (✉)
Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, China
e-mail: nadinexy@aoe.ac.cn

W. Hao
e-mail: weinahao@163.com

71.1 Introduction

In recent years, with the development of information and sensor technology, there has been an increased interest in Location Based Service, or LBS, and a greater demand for navigation and positioning services in different contexts. GPS, the most prevalent navigation technology with the widest coverage, has been improved significantly. However, satellite navigation signals are vulnerable by nature, when interfered with or blocked, they will fail to provide continuous navigation services with accuracy especially in urban canyons and enclosed indoor settings which would greatly attenuate GNSS signals even to the point of invalidity. Against that backdrop, an integrated multi-source positioning technology for ubiquitous signals in both indoors and outdoors settings is an inevitable trend for navigation and positioning services as well as a key issue in the construction and development of a PNT system.

Many researchers have done a lot of work on multi-source fusion based navigation worldwide. Reference [1] synthetically analyses developing trend of indoor positioning techniques. Reference [2] adopts random signal to achieve positioning, and implements technique tests, which verify initial feasibility of navigation technique using random signal. Reference [3] synthetically analyses current wireless network based positioning techniques, concluding that due to effects of environment etc. on the signal strength, wireless technique is unable to provide high-precision positioning. Reference [4] combines methods of 2D and 3D to detect pedestrians. Reference [5] realizes fusion of RFID and supersonic methods via triangular positioning. Reference [6] adapts RSSI based indoor positioning algorithm, and derives path loss index estimation algorithm from wireless propagation loss model of indoor environment based on 4 fixed nodes. However, multi-source fusion based outdoor/indoor seamless positioning techniques aimed at multiple heterogeneous positioning methods still waits for further research.

Fusion and filter algorithm weighs heavily in multi-source fusion techniques. Aside from common Kalman filter technique [7, 8], particle filter [9, 10], neuron network [11], Bayesian estimation [12], maximum posteriori probability estimation [13] etc., many novel filter techniques are also being explored. Reference [14, 15] construct recursive model of sensing data to lower computing loads. Reference [16] proposes using Gaussian mixed model to detect pedestrian in laser imaging space. Reference [17] raises a method of fusion of different sensors. Reference [18] brings forward a semantics based fusion method according to Markov logic network. Reference [19] shows a new fusion method of information from multiple sensors, which theoretically realizes fusion of INS/COD/SAR system by means of ARTMAP neuron network and DS. References [20, 21] bring some new theories such as sparse representation and compressive sensing in visual tracking to solve problem of representation of human body characteristics. Throughout the researches on filter algorithms, no single filter algorithm can effectively solve the problem of optimal fusion of different sensors under different environments. In actual

implementation, proper selection of fusion and filter algorithm according to different elements such as sensor's categories, fusion positioning demands, resource consumption is suggested.

71.2 Sensor Classifications and Property Analysis

When ubiquitous signals from both indoor and outdoor settings are utilized for positioning, heterogeneity between sensors is a major constraint for the integration of multiple sources. Such heterogeneity is reflected in the following aspects:

- (1) Heterogeneity in the utilization of frequency spectrum resources as a result of differences in wireless spectrum properties;
- (2) Errors caused by different measuring methods and heterogeneity in environmental adaptiveness;
- (3) Diversity in interface designs and realization of relevant protocols for different networking technology;
- (4) Heterogeneity in time and space references;
- (5) Heterogeneity in the time of sampling and positions of measurement;
- (6) Variety in terminals and services;

Multiple positioning technologies which can be integrated are assorted based on common sensor types: GNSS and its augmented system are the most commonly used for geo-location both indoors and outdoors; while other supplementary methods can be categorized as measurement-based wireless positioning, fingerprint-based wireless positioning, vision-based positioning, temperature-based cognitive positioning and inertial navigation-based active positioning. Characteristics of and representative positioning methods for each positioning technology are presented below:

(1) Measurement-based wireless positioning technology

Measurement-based wireless positioning technology mainly refers to measurement methods based on distances, angles and time of arrival. Typical applications include TOA/TDOA, UWB ranging scheme, and CMMB-based geo-location.

Wireless signals employed by this type of positioning technology are mostly not specifically for navigation and positioning and are significantly stronger than GNSS signals. They feature good penetrability and do not require massive deployment of additional infrastructures. However, affected by factors such as time synchronization, measurement-based wireless positioning technology tend to lack accuracy and are susceptible to multipath and NLOS in urban crayons and indoor environment.

(2) Fingerprint-based wireless positioning technology

Fingerprint-based wireless positioning technology mainly refer to positioning methods that are based on fingerprint information including signal strength and that obtain a location estimation through finding a closest match for the current signal

strength measurement. Typical applications include WLAN and ZIGBEE positioning.

Fingerprint-based wireless positioning technology feature low equipment costs and is free from the effect of multipath and NLOS. Moreover, it does not require temporal simultaneity of transmitters, and is able to effectively deliver continuous positioning. On the other hand, it requires pre-measurement of signal strengths which will later serve as reference data. Additionally, it is vulnerable to changes in the environment.

(3) **Electro-optical measurement-based assistive positioning technology**

Electro-optical measurement-based assisted positioning technology refers to positioning which is assisted by photogrammetry, infrared detection, and laser measurement. Typical applications include laser sensors, infrared indoor positioning, and active/passive visual positioning.

This type of technology is good at environment exploration and delivering positioning estimates with high precision. It is user-friendly and free from the effect of multipath and NLOS both indoors and outdoors. On the other hand, it requires the deployment of measurement equipment. Devices such as laser sensors are costly and susceptible to light.

(4) **Meteorological effect-based cognitive positioning technology**

Meteorological effect-based cognitive positioning technology mainly refers to the determination of user's altitude through the measurement of barometric pressure, wind velocity and humidity. Altitude determination through barometric pressure is a typical application.

Baroreceptors employed by this type of positioning are low in cost, easy to apply and use simple algorithms. On the other hand, they are vulnerable to wind velocity and the environment. Moreover, accuracy is low.

(5) **Inertial navigation-based active positioning technology**

Inertial navigation-based active positioning technology mainly refers to the positioning technology assisted by users' devices such as gyroscopes and accelerometers to measure positions and speed.

Inertial navigation-based active positioning technology is convenient and effective; therefore, it has been adopted widely. However, errors increases as time drifts; therefore, there is a need to correct them real-time, implying that this technology cannot operate independently for long.

(6) **Geographic information-based matching technology**

Geographic information-based matching technology refers to the acquisition of high-precision information such as location and heading by comparing user information with pre-existing maps.

This type of technology is capable of providing location services with high precision, but it requires the establishment of an accurate database which takes time to build and must be updated real-time with changes in geographic locations. Moreover, querying can be time-consuming if a priori information is lacking.

71.3 Evaluation of Multi-Source Fusion Positioning

71.3.1 Definition of Seamless Positioning

Defined by demand, seamless positioning means seamless coverage of navigation and positioning areas; technically, it means seamless handover between different positioning methods; defined by terminals, it means seamless fusion of navigation and positioning information.

(1) Seamless coverage of navigation and positioning areas

Traditional GNSS is inadequate in terms of positioning accuracy and operability in dense urban areas and enclosed indoor environments where signals are susceptible to blockage and multipath effect. Multi-source fusion positioning technology, on the other hand, can meet positioning needs in different environments because of the combined coverage of each positioning method.

Coverage of navigation and positioning signals can be categorized as passive and active coverage. Passive coverage refers to the situation where users directly receive navigation and positioning signals from mobile positioning systems. These signals can be shared by multiple users. Active coverage refers to the situation where a user determines his/her location while he/she move by detecting the surroundings with the sensors and detectors in hand. Signals are separately used by users for their intended purposes and are exclusive to others.

(2) Seamless handover between navigation methods

While a user moves, the environment he/she is in changes, and the primary navigation method will change accordingly. So same is true when a user's location changes. In these situations, multi-source fusion technology should guarantee the smooth handover between different navigation methods so that there is continuity in this user's location information.

When a user enters an indoor environment from outdoors, the navigation and positioning needs cannot no longer be met by GNSS, as signals attenuate indoors. Therefore, the major navigation and positioning method changes from GNSS to indoor positioning. Multi-source fusion technology must ensure that such transition is seamless so that continuous navigation is delivered reliably and without interruptions or sudden and significant errors.

(3) Seamless fusion of navigation and positioning data

In the same setting there are a variety of navigation and positioning methods available, each with its pros and cons. Therefore, these methods can be combined to achieve the best performance. The key to seamless integration of navigation and positioning data lies in choosing the most appropriate method based on its characteristics and the environment, adjusting the weight of individual methods in the integrated approach so that the advantages of each are fully utilized.

Seen from source data, seamless fusion can be further divided into seamless fusion for measurement domain and seamless fusion for position domain. The former refers to fusion of measurements calculated by multiple methods to delivery position information. An example is GNSS/INS tight coupling. Seamless fusion for position domain mainly integrates positioning estimates by different methods to improve positioning performance.

Seen from how fusion is achieved, seamless fusion has two subcategories: network-side and terminal-side fusion. Network-side fusion involves collecting and fusion of positioning data in information processing center before sending accurate position information to users. Terminal-side fusion is fulfilled through user terminals which achieve fusion of information from multiple sources so as to obtain more precise information about their own positions.

71.3.2 The Purpose of Multi-Source Fusion

The primary purpose of multi-source fusion positioning is to augment GNSS performance in areas including accuracy, continuity, integrity and operability.

Accuracy: A match between the location delivered by the system real-time and the user's actual location.

Continuity: The system's ability in continuously delivering services which meet the established performance requirements under normal operating mode.

Integrity: The system's ability to send alarms to users when it is not effective.

Operability: The percentage of time when the system is providing services which meet the established performance requirements under normal operating mode.

Among the metrics above, accuracy, continuity and operability are key to seamless indoor/outdoor positioning. Operability is affected by accuracy and continuity.

Therefore, a key target for seamless indoor/outdoor positioning is seamless handover and improved accuracy and continuity, hence better navigation and positioning services for users.

71.4 Architecture Design of Multisource Fusion System

71.4.1 Rationale for Multi-Source Fusion Positioning

Technical rationale for multi-source fusion positioning is shown in Fig. 71.1.

Information pre-processing is needed before filtering. Information pre-processing involves: evaluation of signal and data quality, space alignment, time synchronization and outlier processing. Then the processed multi-source navigation information is inputted into multi-source fusion filters, which will deliver high-performance navigation and positioning results after information fusion.

In real navigation scenarios, information transmission and reference data can be non-standard and faulty because of external interference, and issues of timing system and interface technology, making navigation information delivered by navigation devices/systems unsuitable for direct use. Therefore, it must be processed before entered into information fusion filters. These operations before fusion are called pre-processing of information.

Major factors affecting the quality of navigation information are shown below:

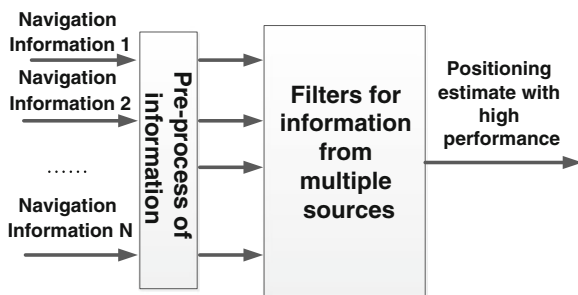
1. Without evaluation and processing, variations in the quality of navigation information sources would hamper the accuracy of multi-source fusion positioning.
2. The coordinate reference system for different navigation devices/systems is not standardized.
3. Navigation and positioning measurements sampled in a certain moment are not for the exact time.
4. Interference of various kinds leads to frequent outliers.

Solutions for the above-mentioned issues are given below.

(1) Evaluation of information quality

In an integrated multi-source system, different measuring approaches can provide location information, operating state and corresponding confidence coefficient; the last two can be used for the evaluation of information quality. When the confidence coefficient fall short of specific requirements or when the operating state

Fig. 71.1 Rationale for multi-source fusion positioning



is abnormal, the corresponding information will be automatically excluded so that the accuracy of multi-source positioning will not be “polluted”. This exclusion operation is called evaluation and processing of information quality.

(2) Spatial alignment

A coordinate system is the chosen benchmark to describe the location and patterns of motion of an object. In a multi-sensor fusion navigation system, measurements for the same physical quantity as described by different coordinate systems must be registered within a single coordinate system so as to shed light on the characteristics and patterns of motion. Moreover, measuring centers for different navigation sources differs and therefore, there is a need to align them to the same measurement point, to achieve consistency of measuring positions.

(3) Information synchronization

An inevitable issue to be addressed during data processing for multi-source fusion is the fact that sampling of time dimensions is out-of-sync due to differences in information sources. It would lead to latency in output seriously affect navigation accuracy if not addressed properly.

Information synchronization of time dimensions can be classified into two levels: the first is consistency of time reference for navigation information from different sources. After the first level has been achieved, the next level is consistency in sampling time.

(4) Outliers handling

Engineering practices show that the combined effect of multiple chance factors result in 1–2 %, or as high as 10–20 % of data points in the measured data set that represent significant variations from the targeted truth value. That is true even for high-precision devices. These variations are referred to as outliers in the field of engineering data processing. Systems that adopt Kalman filters can identify outliers during filtering through predicating the residue. Innovation is defined as: $\varepsilon(k) = Z(k) - H(k)X(k|k-1)$, And identification calculation is done through

$$|\varepsilon(k)| \leq r \sqrt{(H(k)P(k|k-1)H^T(k) - R(k))_{i,i}} \quad (71.1)$$

In this equation, r is a coefficient, usually taken as 3 or 4, $H(k)P(k|k-1)H^T(k)$, represents the variance matrix of one-step-ahead prediction vector, $H(k)$ represents ranging matrix, $X(k)$ represents the state, $P(k|k-1)$ represents predicated covariance matrix, $R(k)$ represents the covariance of the observation noise. If the equation is true, no outliers are detected in the observations; otherwise there exist variations.

If component i of Z_k is detected by the Kalman filter system as an outlier, elements in the same row as Z_i in the gain matrix are set as zero; therefore the effect of the outlier is eliminated. In such situations, only time update is used in the estimation of the state parameters of the system. If outliers occur continuously, error level for state estimation will be increased because of the unavailability of measured

values. In such situations, fitting is needed for some outliers, namely using information from the limited time before the current moment to fit the data. For instance, the smoothed value of the innovation sequence for discrete times 1 to m is used to replace the innovation for time k.

$$\bar{\varepsilon}(k) = \frac{1}{m} \sum_{i=1}^m \varepsilon(k - i) \tag{71.2}$$

71.4.2 Fusion Methods for Information from Multiple Sources

The architecture of fusion for positioning information from multiple sources is shown in Fig. 71.2.

Kalman filters adopted by multi-source fusion navigation systems are either centralized or decentralized (Figures 71.3 and 71.4).

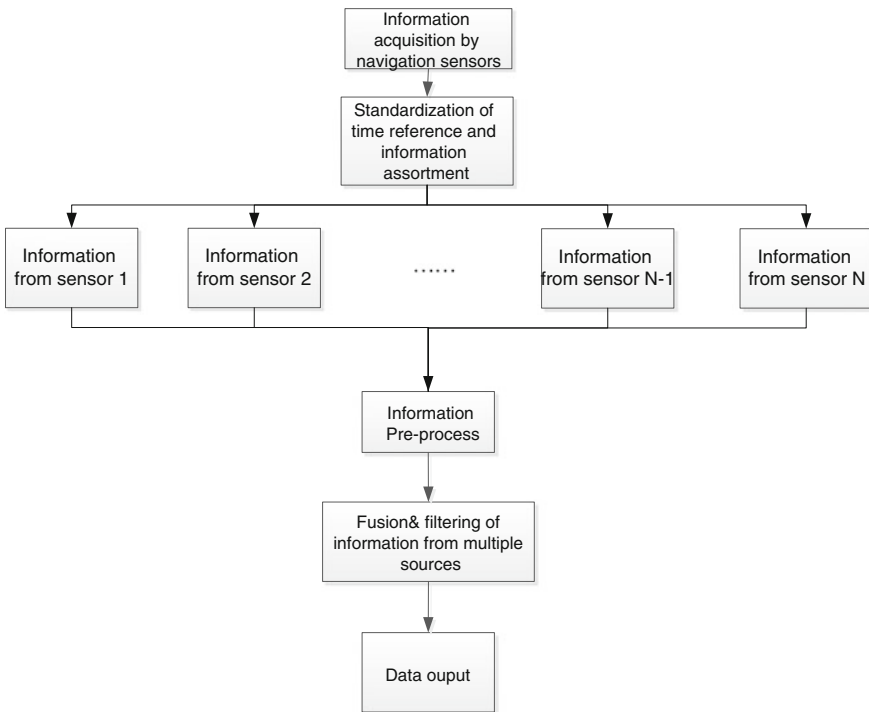


Fig. 71.2 Architecture for multi-source fusion positioning

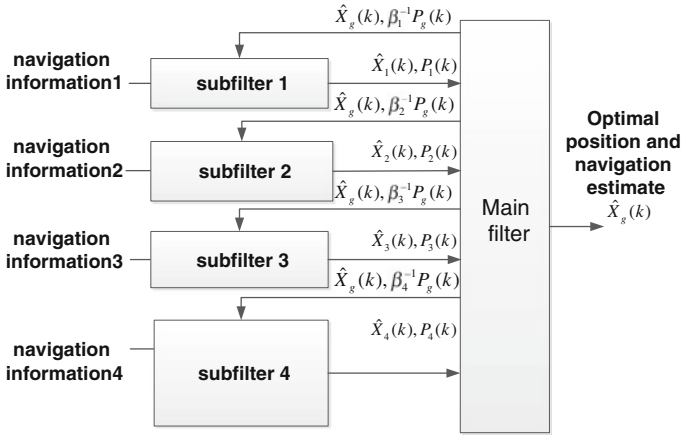


Fig. 71.3 Overall system structure for decentralized filters for multi-sensor fusion

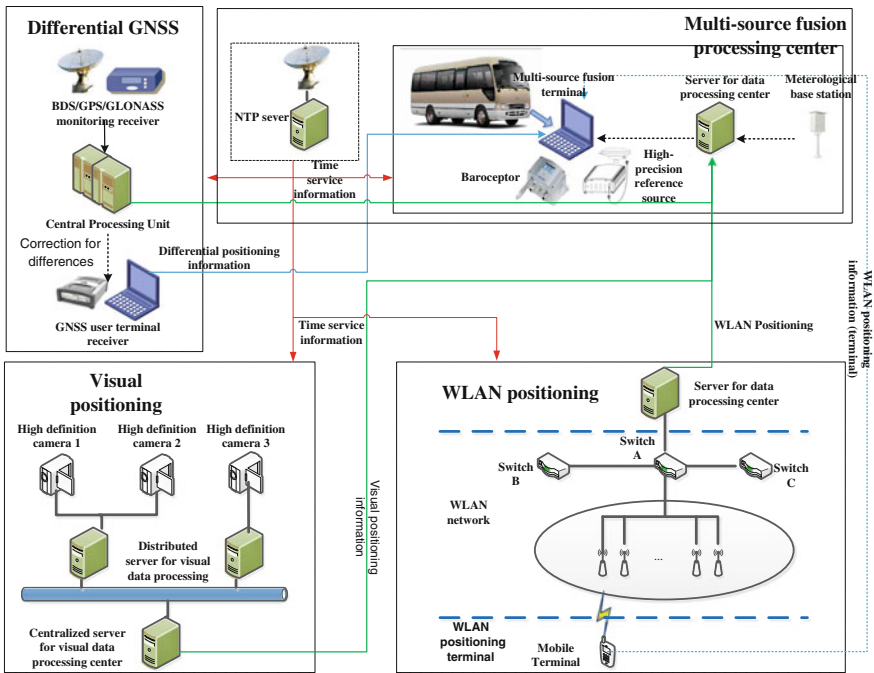


Fig. 71.4 Connection diagram for test equipment

Decentralized Kalman filters for multi-source information fusion are much more robust to errors than centralized ones. Their advantages are listed below:

1. Because their fusion cycle are longer than that of sub filters, there is sufficient time for gradient errors to develop to the level which can be detected by the main filter.
2. Sub filters' states are estimated separately and are not affected by faults in other information sources within the subtitles' cycle. They will only be affected after a long fusion cycle.
3. When faults in a measurement method has been detected and excluded, observations from other sub filters are still there (as long as no resetting has been made) which can lead to a global estimate through simple fusion algorithm. Therefore, decentralized Kalman filters are robust to errors.

For reasons listed above, decentralized Kalman filters are chosen for multi-source information fusion.

The structure for Kalman-based decentralized filters for multi-source fusion is shown above. It is a non-reset one (resetting in decentralized filters would cause cross-contamination of sub filters, resulting in poor robustness to errors.)

It is worth noting that though this filter design is intended for location domain-based information fusion, the overall structure can still be applied to measurement domain. Designs for sub filters and the main filter are discussed in the ensuing part.

(1) Design of sub filters

The main filter uses position domain information from different navigation sensors, while measurements provided by each sub filters differ with different weights, therefore, the state equation and measurement equation for each sub filter should be different. This paper takes differential GPS, passive visual positioning, fingerprinting (WLAN) and positioning based on atmospheric pressure as navigation sources. Sub filter design for each is elaborated in the ensuing passages.

(a) Modeling for sub filters used in GPS

Differential relative positioning can be used for GPS receivers to obtain highly accurate positioning estimates; moreover, estimation of speed can be made with high accuracy via Doppler observations. Therefore, the filtering model can be:

$$\left. \begin{aligned} X_{GNSS}(k) &= \Phi_{GNSS}(k, k-1)X_{GNSS}(k-1) + W_{GNSS}(k) \\ Z_{GNSS}(k) &= H_{GNSS}(k)X_{GNSS}(k) + V_{GNSS}(k) \end{aligned} \right\} \quad (71.3)$$

In the equation, state vector and observation vector are taken respectively as:

$$X_{GNSS} = [e \quad n \quad u \quad V \quad H \quad a \quad \omega]^T \quad (71.4)$$

$$Z_{GNSS} = [e_z \quad n_z \quad u_z \quad V_z]^T \quad (71.5)$$

The position of local coordinate system is represented by e , n and u . Then the following matrices are specified: V , speed; H , heading; a , linear acceleration ω , angular velocity; $W_{GNSS}(k)$ the process noise; and $V_{GNSS}(k)$, the observation noise. The corresponding renewal equation is:

$$\left. \begin{aligned} \dot{e} &= V \sin H + W_1 \\ \dot{n} &= V \cos H + W_2 \\ \dot{u} &= W_3 \\ \dot{V} &= a + W_4 \\ \dot{H} &= \omega + W_5 \\ \dot{a} &= W_6 \\ \dot{\omega} &= W_7 \end{aligned} \right\} \quad (71.6)$$

The renewal equation for observation is:

$$\left. \begin{aligned} e_Z &= e + V_1 \\ n_Z &= n + V_2 \\ u_Z &= u + V_3 \\ V_Z &= V + V_4 \end{aligned} \right\} \quad (71.7)$$

In the equation above, R is the radius of the Earth. The filtering model is obtained through discretization of the state and observation renewal equations shown above. Because the state vector is nonlinear, it can be integrated with a extended Kalman filter for optimal filtering.

(b) Modeling for sub filters used in passive visual positioning

Sub filters used in passive visual positioning are designed in a image coordinate through the location-speed model. State vectors chosen are:

$$\mathbf{X}_{VISION} = [x_p \quad y_p \quad V_{xp} \quad V_{yp}]^T \quad (71.8)$$

In the equation above, x_p and y_p represents the coordinate of the carrier in the image coordinate system. V_{xp} and V_{yp} are speeds in the corresponding direction. So the resulting state model is:

$$\left. \begin{aligned} \mathbf{X}_{VISION}(k) &= \Phi_{VISION}(k, k-1)\mathbf{X}_{VISION}(k-1) + \mathbf{W}_{VISION}(k) \\ \mathbf{Z}_{VISION}(k) &= \mathbf{H}_{VISION}(k)\mathbf{X}_{VISION}(k) + \mathbf{V}_{VISION}(k) \end{aligned} \right\} \quad (71.9)$$

In the equation above, $W_{VISION}(k)$ represents the process noise while $V_{VISION}(k)$ the observation noise.

Because the carrier's position (x_{po}, y_{po}) and linear velocity v_{po} are available, an observation model can be established:

$$Z_{VISION} = [x_{po} \quad y_{po} \quad v_{po}]^T \quad (71.10)$$

$$\left. \begin{aligned} x_{po} &= x_p + v_1 \\ y_{po} &= y_p + v_2 \\ v_{po} &= \sqrt{V_{xp}^2 + V_{yp}^2} + v_3 \end{aligned} \right\} \quad (71.11)$$

Subfiltering can be completed based on Eqs. (71.10) and (71.11).

(c) **Modeling for subfilters used in fingerprinting (WLAN)**

The state vector in WLAN positioning is taken as:

$$\mathbf{X}_{WLAN} = [x_{we} \quad y_{wn} \quad V_{ew} \quad V_{nw}]^T \quad (71.12)$$

Taking positions in the east and north as indicated by WLAN positioning as observation vectors, and the corresponding observation equation is:

$$\left. \begin{aligned} \mathbf{X}_{WLAN}(k) &= \Phi_{WLAN}(k, k-1)\mathbf{X}_{WLAN}(k-1) + \mathbf{W}_{WLAN}(k) \\ Z_{WLAN}(k) &= [x_{we} \quad y_{wn}]^T = \mathbf{H}_{WLAN}(k)\mathbf{X}_{WLAN}(k) + \mathbf{V}_{WLAN}(k) \end{aligned} \right\} \quad (71.13)$$

In the equation

$$\mathbf{H}_{WLAN}(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (71.14)$$

$W_{WLAN}(k)$ is specified as the process noise, and $V_{WLAN}(k)$ the observation noise.

(d) **Modeling for subfilters used in positioning based on atmospheric pressure**

Since Baroreceptors can only provide altitudinal information, the output data is relatively simple. Therefore, the method of least squares can be adopted in output data filtering. The corresponding observation model is:

$$\Delta h_{bar} = \Delta h + W_{bar} \quad (71.15)$$

(2) **Design of the main filter**

Execute one time update step for each subfilters mentioned before through the equations below:

$$\begin{aligned} \hat{X}_i(k, k-1) &= \Phi(k, k-1)\hat{X}_i(k-1) \\ P_i(k, k-1) &= \Phi(k, k-1)P_i(k-1)\Phi^T(k, k-1) + Q_i(k-1) \end{aligned} \tag{71.16}$$

In the equation, $i = 1, 2, 3, 4$, respectively corresponding to subfilters for GNSS, visual positioning, WLAN positioning and positioning based on atmospheric pressure. Matrices are specified as follows: $P_i(k, k-1)$ the one-step-ahead prediction of the covariance matrix of the state estimate for subfilter i ; $P_i(k, k-1)$, the covariance matrix of the state estimate for subfilter i ; and $Q_i(k-1)$, the covariance matrix of the process noise.

Then, the coordinates of each methods are transformed to a unified geodetic coordinate XYZ systems. Herein, WLAN results and Baroceptor results form a 3D positioning result.

Fusion in the main filter is calculated:

$$\left. \begin{aligned} X(k) &= \Phi(k, k-1)X(k-1) + W(k) \\ Z(k) &= H(k)X(k) + V(k) \end{aligned} \right\} \tag{71.17}$$

The measured results are formed by GNSS positioning results, visual positioning results and WLAN and baroceptor based positioning results.

The states are location, velocity and acceleration of each results

$$\left. \begin{aligned} x(k) &= x(k-1) + v(k-1)t + 1/2a(k-1)t^2 \\ v(k) &= a(k-1)t \\ a(k) &= a(k-1) \end{aligned} \right\} \tag{71.18}$$

Fused results are showed below:

$$\begin{aligned} P_g^{-1}(k) &= \sum_{i=1}^3 P_i^{-1}(k) \\ \hat{X}_g(k) &= P_g(k) \sum_{i=1}^3 P_i^{-1}(k)\hat{X}_i(k) \end{aligned} \tag{71.19}$$

At last, the global filtering estimates $P_g(k)$ and $X_g(k)$ execute information distribution based on the following equations:

$$\begin{aligned} P_i^{-1}(k) &= \beta_i P_g^{-1}(k) \\ \hat{X}_i(k) &= \hat{X}_g(k) \end{aligned} \tag{71.20}$$

71.5 Testing and Analysis

Based on the requirements of the preliminary test, the new technology base of Chinese Academy of Sciences was chosen as a test ground where equipment including stations for differential GNSS corrections, WLAN signal nodes, devices for passive visual positioning, devices for correcting differences in atmospheric pressure, and communication devices was in place. Terminals tested include GNSS, WLAN and atmospheric pressure testing modules. They receive data such as corrections of differential GNSS, WLAN fingerprinting data, and atmospheric pressure measurements and obtain estimates of differential GNSS and WLAN positioning and atmospheric pressure-derived altitudes. Fusion of these estimates results in terminal-side estimation of the position.

Diagram for test environment is presented in Fig. 71.5.

Figure 71.6a, b are positioning track comparisons of actual measurements of single method positioning and fusion method positioning respectively, herein (a) shows results of outdoor scenario with (b) of indoor scenario. In the figure, the deep blue line represents fusion positioning results, the deep red line denotes high-precision reference, and in outdoor environments, we can see the two lines are almost in alignment, prevailing over other single method positioning results. Figure 71.6 shows that seamless handover between indoor/outdoor positioning within the tested area can be realized. Actual measurements reveal that seamless handover between indoor/outdoor positioning which ensures continuous positioning with navigation and positioning frequency of 1 Hz within the test area is possible using the architecture and algorithm of the multi-source fusion system mentioned before. Figure 71.6c shows comparisons of fusion positioning and differential positioning, herein, the blue line represents differential result, and the red

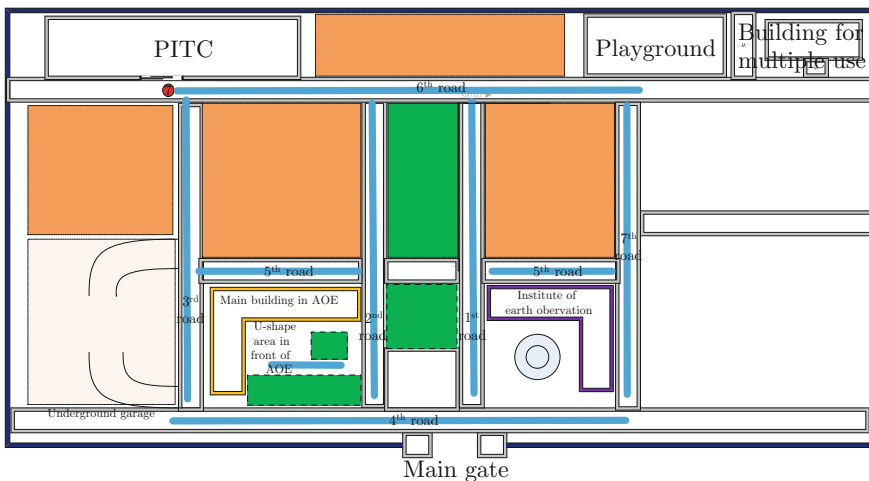


Fig. 71.5 Diagram for test environment

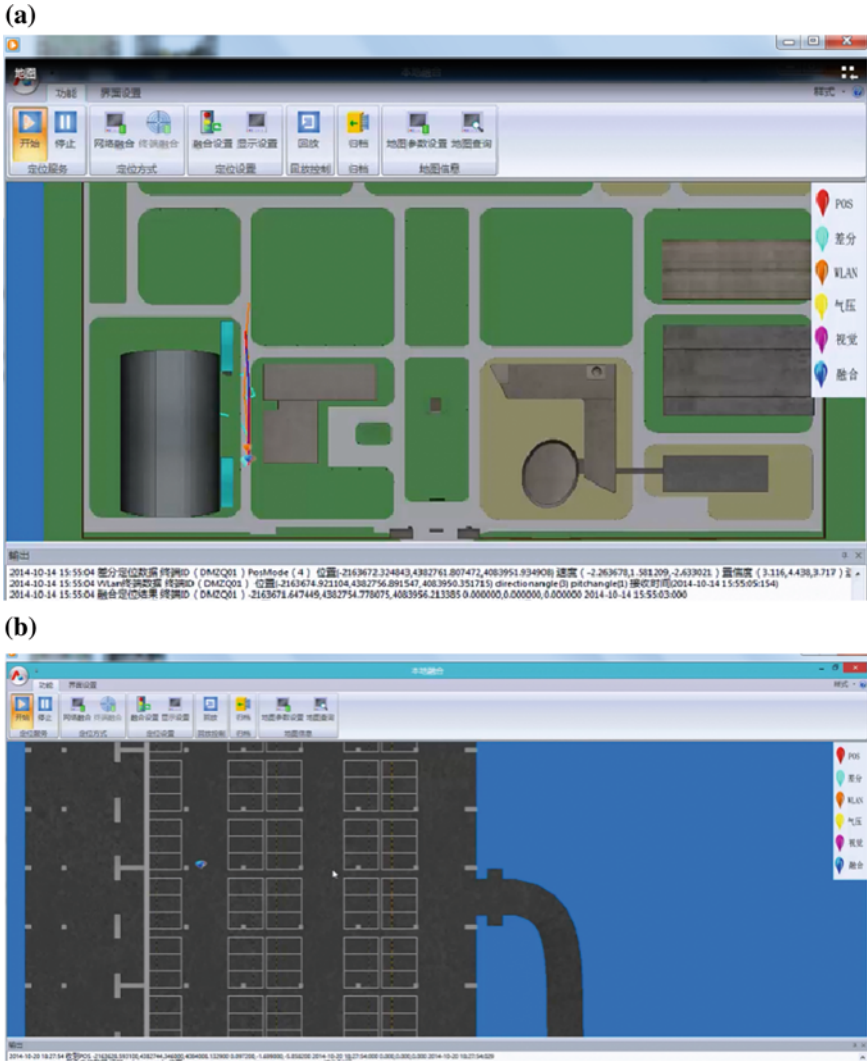


Fig. 71.6 Results for indoor/outdoor positioning. **a** Results for outdoor positioning. **b** Results for indoor positioning. **c** The fusion positioning result (the red one) and the GNSS positioning result (the blue one)

one represents the other result. From Fig. 71.6, we can see fusion positioning precision is obviously better than GNSS local differential precision, and guarantees continuous positioning even when GNSS data drifts apparently in urban valley.

The vehicle was running for 1 h within the test area with positioning results during this process gathered. The calculated errors for network-side and terminal-side positioning are presented separately in Fig. 71.7.



Fig. 71.6 (continued)

It must be pointed out that reference data is not available for underground garage. (The test adopted high-precision GNSS/INS integrated device for reference, which can not be used in underground garage because of drifting.) Therefore, qualitative analysis mainly focuses on the accuracy of surface positioning. Accuracy of positioning in the underground garage is analyzed through estimation.

Figure 71.7 is positioning error analysis diagram of fusion positioning, where red line is error standard line corresponding to 1σ (66.7 %) probability, and it shows that whether it is network-side or terminal-side positioning, multi-source fusion positioning is more accurate than any single positioning method, with positioning accuracy better than 1 m (1σ).

Indoor positioning is usually fulfilled through WLAN and baroreceptor-derived positioning. Its accuracy is affected by that of the WLAN and baroreceptor-derived positioning. (WLAN is responsible for accuracy of two-dimensional [horizontal] position estimation while baroreceptors for one-dimensional [altitudinal]). Testing of fixed points show that baroreceptor-derived positioning within the test area has a accuracy performance of better than 0.4 m, and WLAN positioning better than 3 m. Therefore the accuracy level for dynamic indoor positioning is about 3 m (1σ).

Moreover, comparison of the continuity performances of different positioning methods show that multi-source fusion positioning can improve continuity: it is 40 % better than differential GNSS, and 10 % better than WLAN positioning. This paper proposes an architecture and filtering algorithm for the multi-source fusion based system. Seamless indoor/outdoor positioning through the integration of

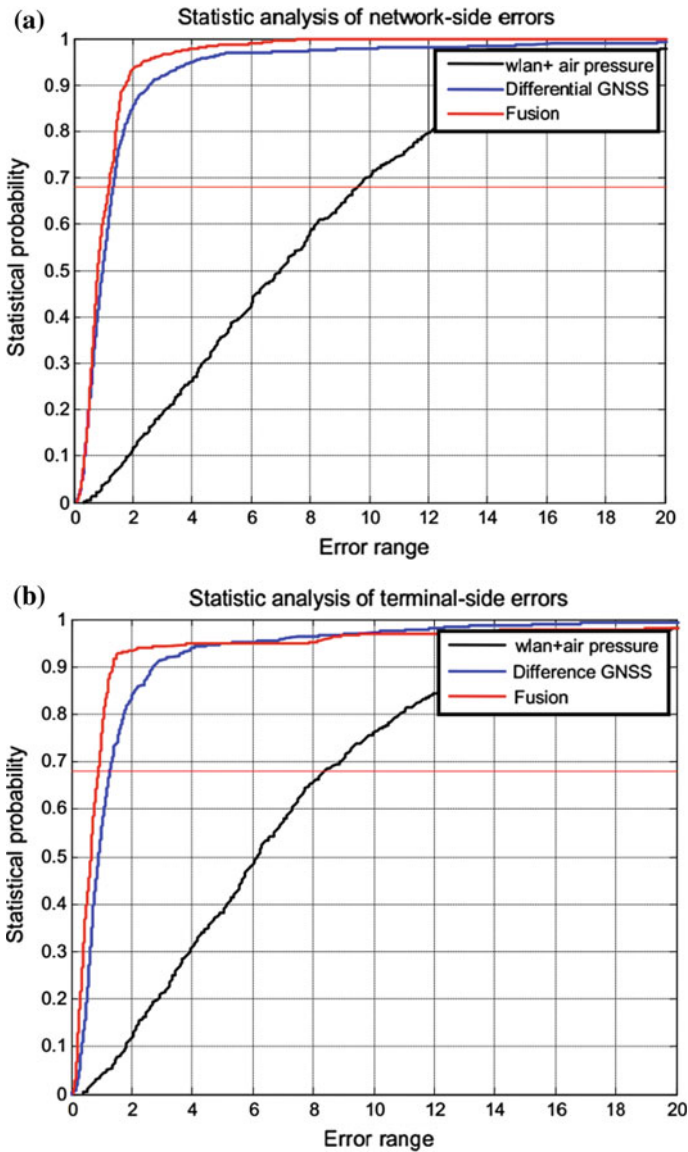


Fig. 71.7 Accuracy analysis for fusion positioning. **a** Error analysis for network-side positioning. **b** Error analysis for terminal-side positioning

multiple positioning methods was tested. Results show that this technology effectively improves the accuracy and continuity of GNSS and can meet users' demand for seamless indoor/outdoor positioning. Multi-source fusion positioning points to a new approach for seamless indoor/outdoor positioning, and is a case of the

application of multi-source fusion in navigation and positioning, hence providing a reference for the development of all-source navigation technology, opportunity navigation technology and PNT technology.

71.6 Conclusion

This paper analyses basic characteristics of navigation signal and communication signal, and proposes a novel signal system bearing basic features of communication signal, aiming at realizing deep fusion in signal layer. In this paper, the new signal's correlation function, power spectrum density, RMS bandwidth and positioning accuracy are studied, followed by comparative analysis of the new signal and BPSK modulated C/A code signal and BOC modulated BOC (4,1) signal, with the conclusion that BCC signal has better of navigation and positioning performance. Moreover, this study also puts forward signal synchronization algorithm based on signal features, simulates signal performance and tests and verifies key algorithm. The result indicates that single-carrier BCC signal's ranging accuracy is 6.06 m, and positioning accuracy is 10.8 m. The ranging and positioning accuracy will be better in multi-carrier conditions, which will be sufficient for users' basic positioning requirements. The analysis of communication signals confirms that the signal is able to integrate with existing communication signals, which is worthy of further in-depth study.

References

1. Deng Z, Yu Y, Yuan X et al (2013) Situation and development tendency of indoor positioning. *Communications* 10(03):42–55 China
2. Tian H, Xia L, Mok E et al (2009) Signals of opportunity assisted ubiquitous positioning and its key elements for outdoor/indoor environment. *Geomatics Inf Sci Wuhan Univ* 34 (11):1372–1376
3. Liu Y, Yang Z, Wang X et al (2010) Location, localization, and localizability. *J Comput Sci Technol* 25:274–297
4. Ger'onimo D, Sappa AD, Ponsa D et al (2010) 2D-3D-based on-board pedestrian detection system. *Comput Vis Image Underst* 114:583–595
5. Choi BS, Lee JW, Lee JJ et al (2011) A hierarchical algorithm for indoor mobile robots localization using RFID sensor fusion. *IEEE Trans Industr Electron* 58(6):1–10
6. Lu Y, Lai C, Hu C et al (2010) Path loss exponent estimation for indoor wireless sensor positioning. *KSH Trans Internet Inf Syst* 4(3):243–257
7. Weng SK, Kuo CM, Tus K (2006) Video object tracking using adaptive Kalman filter. In: *International journal of visual communication and image representation*, vol 17. ACM Academic Press, Orlando, pp 1190–1208
8. Hammes U, Wolsztynski E, Zoubir AM (2009) Robust tracking and geolocation for wireless networks in NLOS environments. *IEEE J Sel Top Sign Proces* 3(5):889–901

9. Song Y, Li QL, Sun FC (2009) Shannon entropy-based adaptive fusion particle filter for visual tracking. In: IEEE chinese conference on pattern recognition. IEEE Computer Society Press, Nanjing, pp 1–5
10. Maran S, Gifford WM, Wymeersch H (2010) NLOS identification and mitigation for localization based on UWB experimental data. *IEEE J Sel Areas Commun* 28(7):1026–1035
11. Filillidis A, Jain LC, Martin N (2000) Fusion of intelligent agents for the detection of aircraft in SAR images. *IEEE Trans Pattern Anal Mach Intell* 22:378–384
12. Yuan Y, Kam M (2004) Distributed decision fusion with a random-access channel for sensor network applications. *IEEE Trans Instrum Meas* 53:1339–1344
13. Shah SFA, Ribeiro A, Giannakis GB (2005) Bandwidth-constrained MAP estimation for wireless sensor networks. In: Proceedings of the 39th Asilomar conference on signals, systems and computers. IEEE Signals, Syst Comput Pacific Grove. pp 215–219
14. Guestrin C, Bodik P, Thibaux R et al (2004) Distributed regression: an efficient framework for modeling sensor network data. In: Proceedings of the international symposium on information processing in sensor networks. IEEE Computer Society Press, Berkeley, pp 1–10
15. Willett R, Martin A, Nowak R (2004) Backcasting: adaptive sampling for sensor networks. In: Proceedings of the international symposium on information processing in sensor networks. IEEE Computer Society Press, Berkeley, pp 124–133
16. Premebida C, Nunes U (2006) A multi-target tracking and GMM classifier for intelligent vehicles. In: IEEE international conference on intelligent transportation systems. IEEE Computer Society Press, Toronto, pp 313–318
17. Spinello L, Siegwart R (2008) Human detection using multimodal and multidimensional features. In: IEEE international conference on robotics and automation. IEEE Computer Society Press, California, pp 3264–3269
18. Oliveira L, Nunes U, Peixoto P et al (2010) Semantic fusion of laser and vision in pedestrian detection. *Pattern Recognit* 43:3648–3659
19. Jiang CH, Chen Z (2002) Multisensor fusion based on decision-making in INS/CCD/SAR integrated systems. In: Proceedings of IEEE Tencon'02, pp 695–698
20. Li H, Shen C, Shi Q (2011) Real-time visual tracking using compressive sensing. In: IEEE conference on computer vision and pattern recognition. IEEE Computer Society Press, Providence, pp 1305–1312
21. Mei X, Ling H (2011) Robust visual tracking and vehicle classification via sparse representation. *IEEE Trans Pattern Anal Mach Intell* 33:2259–2272

Chapter 72

Study on Quantum Inter-satellite Link Technology

Xian'an Zheng, Ying Wang, Yansong Meng and Zhe Su

Abstract It is a very important way to realize the satellite data injection, control and autonomous navigation through inter satellite link to communicate and range. Quantum technology has special advantages in realizing the secure communication, and has been gradually tested and applied. The technology based on the high precision ranging theory of quantum interference has become more mature. The design and analysis is of strategic significance to our future generation of global navigation satellite system inter-satellite technology development.

Keywords Inter-satellite link · Quantum ranging · Quantum interference

72.1 Introduction

Communications and ranging based on inter satellite link (ISL) is an important means of inputting multiple star system business data, control and maintenance of the system, maintenance of system accuracy, realization of autonomous navigation and system survivability enhancement. Without atmospheric effects, ISL particularly is suitable for optical communication. But poor laser communications privacy, information was easily obscured or blocked. Whether continuous laser ranging or pulsed laser ranging, the precision is limited to the electromagnetic wave noise limit, compared with quantum ranging. Relative to the laser and microwave, quantum inter satellite link has the advantage of confidentiality and accuracy.

At present, studies based on the entangled light quantum correlation properties of continuous progress, has gradually entered the application stage, and the related solutions and techniques have been shown to provide higher safety and precision.

X. Zheng (✉) · Y. Wang · Y. Meng · Z. Su
China Academy of Space Technology (Xi'an), Xi'an 710100, China
e-mail: 13992851132@qq.com

© Springer-Verlag Berlin Heidelberg 2015
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume III*, Lecture Notes in Electrical Engineering 342,
DOI 10.1007/978-3-662-46632-2_72

839

This paper shows, the design of the inter satellite link system, the research and the discussion based on the entangled photons inter satellite measurement communication scheme and related technology.

72.2 Quantum Inter Satellite Link Architecture

Quantum inter satellite link consists of three layers: the first is the classic channel that can be laser or microwave channel. Taking into account the use of photons as the media of information transmission, laser channel can be the first choice because of the ease of integration, less power consumption and volume. The second is the quantum range channel that needs to send the entangled photons, along the line of sight propagation. This link needs that building satellites must be visible. The third one is quantum teleportation channel, namely quantum communication channel. This link doesn't need that building satellites must be visible (Fig. 72.1).

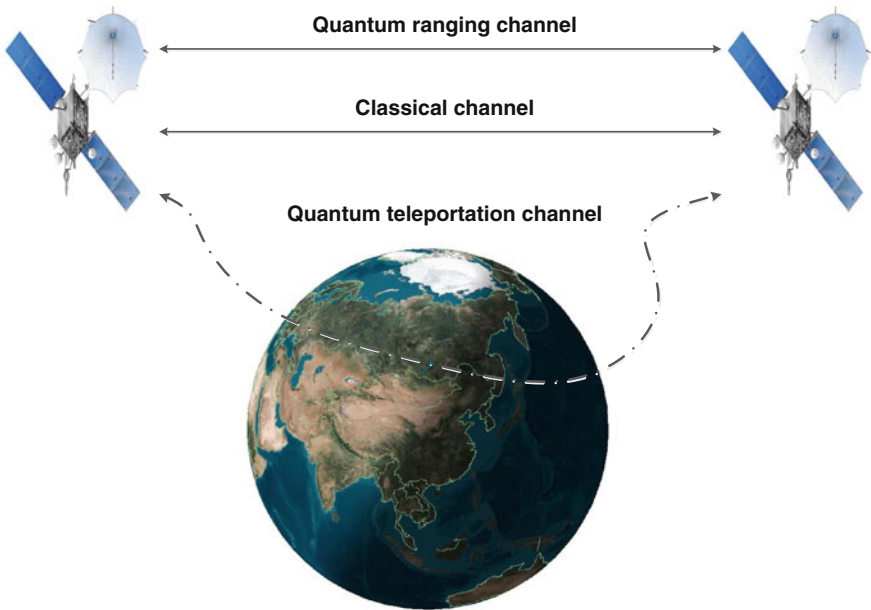


Fig. 72.1 Quantum inter satellite link

72.3 Quantum Inter Satellite Link Key Technology

72.3.1 *The Generation and Detection Technology of Entangled Photons*

Quantum entanglement is a kind of superposition state, and reflects the physical process cannot be differentiated. The theory of entanglement can occur in any physical quantities, such as energy, momentum, position, polarization etc. A typical two body entanglement of the superposition state, can be written as

$$\frac{1}{\sqrt{2}}(|a_1\rangle|b_1\rangle + e^{i\alpha}|a_2\rangle|b_2\rangle) \quad (72.1)$$

It reflects the state $|a_1\rangle|b_1\rangle$ and $|a_2\rangle|b_2\rangle$ cannot be distinguished. $e^{i\alpha}$ denotes the relative phase of two state. Entangled states are generated by making the indistinguishability. Practical technology usually uses polarization and frequency for generating quantum entangled state.

Detection of entangled photons based on single photon detection technology, mainly includes the single photon detection system operating in Geiger mode avalanche photoelectric diode semiconductor and the superconducting single photon detection system based on super conductor material. At present, the application of more is still the single photon detector based on avalanche diode. For the detection of weak optical signal, we need to reduce background noise and improve the gain.

For detection of entangled photon pair, the mature method is coincidence counting method. Coincidence counting measurement is used for association events occurring at the same time or measured in a short time. Limited by instrument capacity, two photons cannot be distinguished arrived at the same time. So we can set with a gate width, and if in the door wide time two photons are detected, we define they arrived at the same time. The algorithm makes the accuracy of time detection limited by the door width settings. Based on the coherence needs, the door width must be less than the coherence time of the light field.

Quantum coherence is a kind of statistically coherent processing; therefore the coherent time generally refers to the statistics of data acquisition time. The time of collecting data is longer, the final coincidence numerical estimation is more accurate and the accuracy is higher. But because of inter satellite high velocity, long time statistic makes the measurement value output frequency decreased, at the same time it does not increase linearly but increase the estimation accuracy, and on-board processing burden.

72.3.2 The Quantum of Inter Satellite Communication Technology

At present, the quantum communication is mainly related to quantum cryptography, quantum teleportation and quantum dense coding and other basic and applied research, and is one of the quantum technology field has to practice application.

Quantum cryptography is quantum key distribution using quantum state to realize secure communication, but its quantum state role is limited to classical association information establishing communication between the two parties, namely only plays the role of key and password. The research team led by Pan Jianwei of China's University of Science and Technology in Hefei in 2009, realize quantum cryptography network [1].

Quantum teleportation, its key is the connection properties by using quantum entanglement. The entanglement and the measurement is established between the two sides, so the quantum state of the message can be accurately transferred to another. Due to pass and no physical information in the traditional sense, it make theory the transmission without almost any time loss. In 2010, the Pan Jianwei's team implements the 16 km free space quantum teleportation, which verified the feasibility of satellite ground quantum communication [2].

Due to the quantum ranging is the use of high order correlation properties of quantum, so the light quantum phase, frequency and other parameters can still be used for ordinary communication. In addition, quantum communication information can also be measured by quantum interference, so each satellite only need to configure a receiving terminal and a mirror for inter satellite communication and ranging. In theory, quantum teleportation is not limited by the distance, and a vacuum ranging using femtosecond photonic transmission loss can also be controlled; sighting and tracking system can inherit laser equipment, therefore it is suitable for the establishment of inter satellite link.

72.3.3 Quantum Inter Satellite Ranging Technology

In theory, quantum entanglement is produced at the same time, also be detected. The realization of the spatio-temporal correlation of quantum ranging mainly depends on the entangled photons. Two photon entangled state is generated based on the different ways of pumping light generation can have a continuous optical pumping and ultra-short pulse pumped parametric down conversion. Continuous pump has long coherence length and performance of coherence is better. The generation time of two photon is random based on continuous laser pump, and it is with a certain probability of generating two-photon entangled. But ultra-short pulse generation by nonlinear crystal double photon pairs is determined by the pulse time, only when the pulse after passing through the nonlinear crystal the two-photon is generated. Compared with the continuous laser pump entangled photon pair, pulse

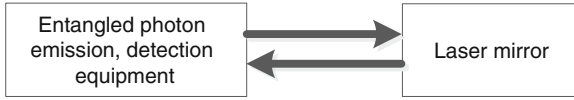


Fig. 72.2 Reflective ranging principle

laser pump photon entangled on the temporal and spatial performance is more controllable and easy to use. Therefore, quantum communication, quantum ranging use ultra-short pulse laser system for preparing entangled source. Because of the influence of the purity, the current detection capability of entangled state, in fact the pulse width influences the location precision.

Quantum ranging using the reflection type, as shown in Fig. 72.2, the target object configures angle reflector, the local system configures photon emission and detection equipment. There are mainly two kinds of schemes:

Scheme I: Strength method

When the pulse width is wide, we may not consider the electric field phase, only consider the strength, this situation can be called the intensity autocorrelation. The output is expressed as:

$$G(\tau) = \int_{-\infty}^{\infty} I(t)I(t - \tau)dt \tag{72.2}$$

where $I(t)$ and $I(t - \tau)$ is pulse intensity respectively, τ is the relative delay.

As shown in Fig. 72.3, in the scheme, the vast majority of processing is in the satellite A. Ultra short pulse laser generated by laser on satellite A, then through BBO crystal to generate entangled photon pairs, called the signal light and idle light. Signal light passes through the space transmission, reaches the satellite B, and is reflected back to detector 2. On the satellite A, internal idle light through a delay regulator that adjusts the delay, then is detected by detector 1. The two detector output is access to a coincidence counter. If the detector 1 and 2 at the same time to detect the pulse detector, it shows the counter has the peak value. Recording the amount of delay regulator, we can calculate the optical signal transmission distance. Among them:The signal light distance:

$$L_{signal} = L_A + L_B + r + L_r + \delta_s \tag{72.3}$$

The idle light distance:

$$L_{idler} = L_A + L_{delay} + L_r + \delta_i \tag{72.4}$$

where, L_A, L_B is transmission delay in satellite A, B; L_r is the public delay of transmitting and receiving. δ_s, δ_i is the path length error and measurement error of signal and idler in different paths. r is the real distance. L_{delay} is the delay of delay regulator.

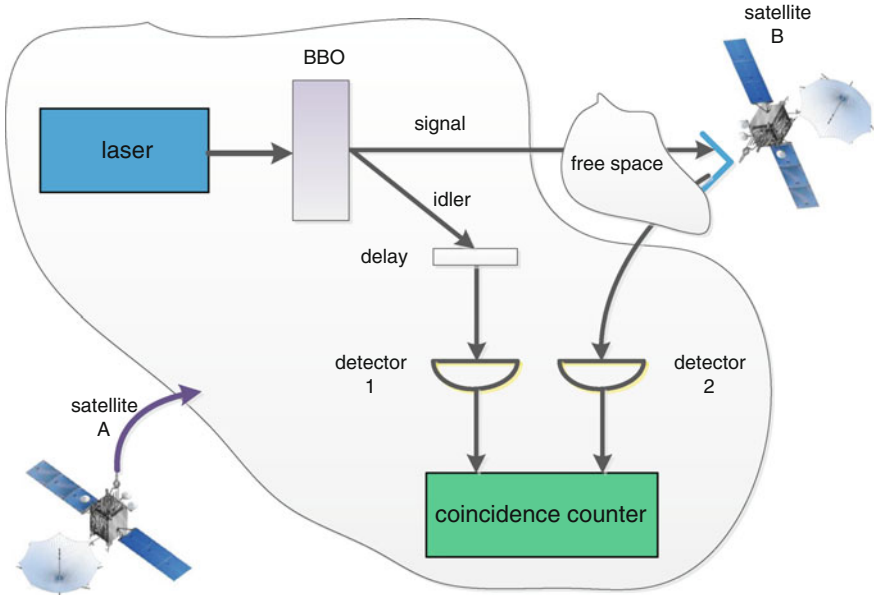


Fig. 72.3 Quantum ranging scheme I

When the coincidence count has peak, we define

$$L_{signal} = L_{idler} \tag{72.5}$$

Then the distance between the satellites A, B is

$$r = L_{delay} + \delta_i - L_B - \delta_s \tag{72.6}$$

The principle of this scheme can range using common laser pulse, at this time we can only use a laser pulse, then record the time of transmitting and receiving time for distance measurement. But after long distance transmission, the laser pulse energy is weak, so only by using the pulse shape it is difficult to accurately measure the arrival time. Another way is using second-order coherence properties, the accuracy can be less than the wavelength, but the existence of integer ambiguity problem is difficult to solve.

Analysis shows that, the quantum ranging scheme mainly uses synchronization of the entangled photon pairs generated and detected.

Scheme II: quantum interference method

Quantum interference can reflect the two order correlation effect of quantum, and can use various possible system states to analysis. From a number of possible states, coincidence measurement selects some to satisfy the specific association of state.

As shown in Fig. 72.4, by adding the 50:50 beam splitter in the two paths, two photon will have four possible paths: reflection ‘+’ reflection, transmission ‘+’

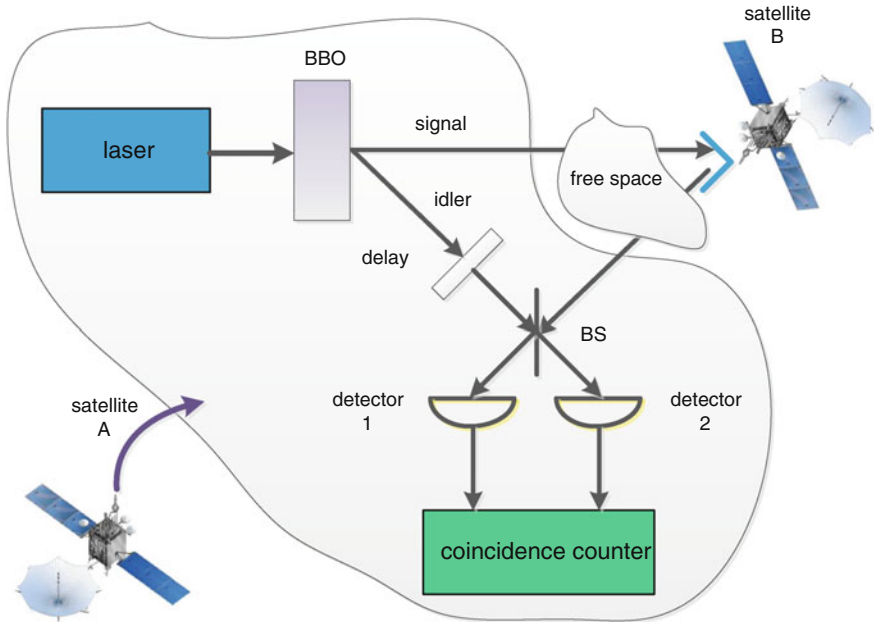


Fig. 72.4 Quantum ranging scheme II

transmission, reflection ‘+’ transmission and transmission ‘+’ reflection. Complete destructive interference makes some possibility into 0; add interference makes some other event probability becomes large. In the quantum ranging scheme II, HOM (Hong-Ou-Mandel) interference makes two kinds of situation probability become small, another large. With the counter we can distinguish depression, then measure the two photon arriving time. The use of femtosecond pulse laser get depression entangled photons can reach the accuracy obtained femtosecond laboratory [3, 4] (Fig. 72.5).

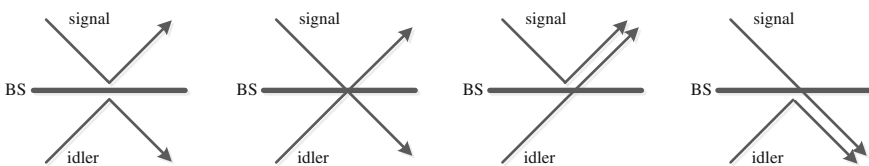


Fig. 72.5 Entangled photon reflection and transmission case through the beam splitter

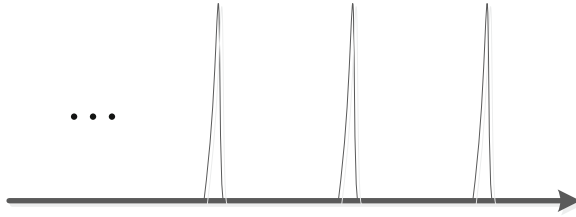


Fig. 72.6 Femtosecond pulse signal

72.4 The Quantum of Inter Satellite Link Signal

Quantum inter satellite link information carriers are photons, so system works on optical frequency band. According to the light in the atmosphere attenuation characteristic, the wavelengths of 850, 1060 and 1550 nm are better bands for the transmission [1]. Entanglement source band generally is 650–1080 and 1520–1560 nm. So we can select 850, 1060 or 1550 nm band as the frequency band for inter satellite link.

The target of link signal design is to meet the quantum of inter satellite communication and ranging request, considering the equipment can be realized, including power consumption and volume element. Entangled photon concentrated on the femtosecond pulse can improve the interference efficiency, or need long time to obtain an interference effect. A signal scheme as shown in Fig. 72.6, the pulse repetition frequency is determined by communication rate and average transmission power. When without using the quantum teleportation, we can use the OOK modulation for information transmission (Fig. 72.6).

In addition, quantum interference was carried out using properties of photon probability wave interference, so the appearance of cumulative probability needs of a large number of photon. One way is that a light pulse contains many quanta, so single pulse interference would be enough to achieve measurable degree of interference depression. Another way is that many optical pulse containing quantum transmit in a certain frequency, then the time accumulation can make interference depression appear.

72.5 Quantum Constellation Topologies and Link Network

72.5.1 Quantum Constellation Topology

Quantum inter satellite link signal is a kind of quantum signal based line of sight propagation signal. If there are 8 satellites for each track on the orbit, several link topology schemes are shown in Fig. 72.7. The transmission distance is affected by the impact of topology [5]. Quantum transport theory using single photon, so there

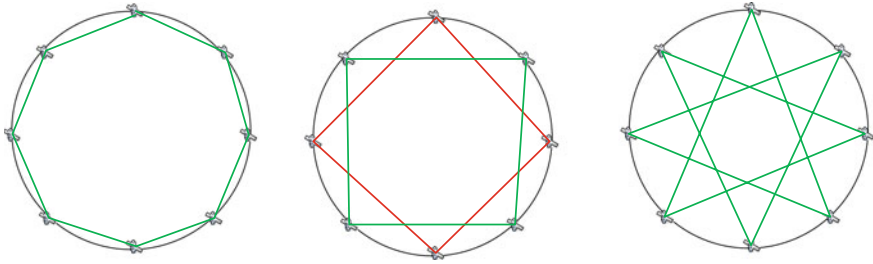


Fig. 72.7 Design of link topology in an orbit plane

is no problem of the divergence angle, but the single photon is easily absorbed or scattered by cosmic dust or particles, and detector has certain probability for detecting. Because the photons can be in femtosecond pulse, divergence angle is small undergoing a long distance after the beam.

Due to delocalization of quantum entanglement, the requirements of the communication link topology are greatly reduced, and invisible satellites can also be carried out key information communication.

72.5.2 *Quantum Link Networks*

Quantum inter satellite link network architecture in accordance with the hierarchical model can be divided into three layers: the physical layer, the data layer, network layer.

Preparation, transmitting and receiving, detection in the physical layer of light quantum, the main technology is the entangled state, squeezed state preparation, transmission to maintain and detection, focus on optic, laser and quantum devices.

The data layer relates to the information modulation and demodulation, the part is similar to microwave. But due to the use of quantum cryptography and quantum teleportation, there is greater difference of demodulation method, transmission frame structure and protocol.

The network layer is solving the routing problem, to realize the end-to-end quantum key communication, and is identical with the microwave link. But in quantum teleportation, when satellites are visible, the link is established and the quantum entanglement information is transferred. When invisible, still can use non localized characteristics of entanglement for direct information transmission, and in theory there is no time delay.

72.6 Capacity and Precision Estimation

At present, the quantum teleportation of inter satellite link communication is limited to the relevant technology, so communication capacity and cost also cannot be compared with the traditional channel.

Inter satellite ranging accuracy is mainly affected by the emission of noise, receiver noise, relativity, antenna phase center deviation, inter satellite distance and other factors. To present the device performance, considering only the distance of noise emission and receiving noise, inter satellite ranging accuracy can reach below 1 cm.

Satellite autonomous navigation accuracy is mainly composed of unit weight error and geometric dilution of precision (PDOP) decision. Geometric dilution of precision is decided by topological. Unit weight error of [6].

$$\hat{\sigma}_0 = \sqrt{\frac{V^T R^{-1} V + \delta\bar{X}^T P^{-1} \delta\bar{X}}{m}} \quad (72.7)$$

where, R^{-1} is the observation weight matrix, and inversely proportional to the measurement error; $\delta\bar{X}$ is the reference orbit; V is the residuals calculated by observation equation; P is error covariance matrix of forecast value; m is the number of observations.

When we do not consider the constellation rotation system error, the navigation accuracy can be up to 1 m. To reduce the error is mainly from the measurement precision and the inhibition of constellation system error.

72.7 Conclusions

Quantum inter satellite link from the theory of quantum mechanics, dependent on the laser technology, is the next generation of link technology after microwave link and the laser link. Quantum inter satellite link communication has the advantage of good safety and nonlocality. Quantum link advantage of ranging is not vague and high precision. Both can be combined, which focused on their respective advantages, so that quantum link between satellites become an advanced way of inter satellite link.

References

1. Chen TY, Wang J, Liang H et al (2010) Metropolitan all-pass and inter-city quantum communication network. *Opt Express* 18:27217–27225
2. Jin XM, Ren JG, Yang B et al (2010) Experimental free-space quantum teleportation. *Nat Photonics* 4:376–381

3. Jiang G (2012) Study of two photon interference. Graduate School of National University of Defense Technology, Changsha
4. Zhang Y, Quan R, Bai Y, Hou F, Liu T, Zhang S, Dong R (2013) Frequency indistinguishability investigation of generated coincident-frequency entanglement via ultra-fast pulsed sources. *Acta Phys Sin* 62(14) (Article ID 144206)
5. Wang X, Meng Y, Su Z, Wang G, Zhang L, Gong D (2014) A simple three-dimensional matrix method for global constellation intrasatellite link topological design. *Int J Navig Obs* 2014:17 p (Article ID 502158). <http://dx.doi.org/10.1155/2014/502158>
6. Zhonggui C (2012) Key technologies of autonomous operation for navigation satellite constellations using inter-satellite tracking data. Graduate School of National University of Defense Technology, Changsha

Part IV
Policies and Regulations, Standard
and Intellectual Property

Chapter 73

The Urgency and Principals of the Industry Policy Optimization for Beidou Satellite Navigation System

Junlin Yang and Xiangming Hu

Abstract As a strategic emerging and national defense industry, the satellite navigation system needs effective industrial policy to drive rapid development. The following four reasons can account for this view. Firstly, the development of satellite navigation system needs effective industrial policy guidance urgently to adapt the disorder even vicious competition in the beginning. Secondly, satellite navigation deserves the government protection and support as a comparative advantage industry. Thirdly, the government must adjust the market access mechanism of satellite navigation industry to achieve the optimal resources allocation and strongest competitiveness. Fourthly, the policy efforts of satellite navigation is far lower than the US, Russia or other developed countries with the disadvantages of low level, no system, too general and lack of policy tools. To perfect the Beidou satellite navigation system, the following principals should be considered: the government leading adapted with market economy system; the global view based on national conditions; the systematic design with overall propulsion; the intensive attention of theoretical research.

Keywords Satellite navigation system · Beidou · Industry policy

73.1 Introduction

The application of satellite navigation has become the third fastest global industry following the Mobile Communication and Internet, and now it is one of the most creative emerging industries. As a high technology industry, the rapid development

J. Yang (✉) · X. Hu
School of Public Management, Beihang University, Beijing 100191, China
e-mail: yjl13511063625@sina.com

X. Hu
e-mail: huxm9@sina.com

could not only drive national economic increment, but also enhance the adjustment of economic structure and increment quality. This industry is closely linked with information safety, defense safety and national safety. Therefore it is not only a strategic emerging industry, moreover, a strategic defense industry. It is the duty of current administration to boost the qualitative and rapid development of this industry [1].

73.2 To Realize the Necessary to Improve Industrial Policy of the Beidou Satellite Navigation System in Theory

Modern economics has proven government industrial policy is an important tool to promote the development of specific industries, and widely used by the Western countries. It is feasible and necessary to develop the industry of Beidou Satellite Navigation System with the help of policies.

73.2.1 To Realize the Necessary to Improve Industrial Policy of the Beidou Satellite Navigation System in the Opinion of Market Failure Theory

According the market failure theory, some disadvantages from the information lag, profit-driven and monopoly capital and other insurmountable flaws may cause some blind market behavior, spillover effects, or insufficient competition, which requires the government to develop industrial policy to correct market failures [2].

China's Beidou navigation system industry is in its infancy, compared with the current popularity of the United States global positioning system (GPS) in the Chinese market, Beidou satellite navigation system is still in a very vulnerable position in the market competition. In the current market environment, it is difficult for China's Beidou satellite navigation system to develop normally, especially to make it counterpart of applications such as GPS navigation system [3], if there is no effective industrial policy to guide.

In the development of our current Beidou industry, disorderly market competition, even vicious competition phenomenon inevitably occurs due to imperfection of industrial policy and supporting. Some companies invest blindly without full understanding of the information, and this is not only a waste of resources, but also a serious impediment to the healthy development of China's satellite navigation industry. Especially in competition with GPS navigation applications with inherent advantage, it is hard to compete on a similar platform. Beidou is related to China's information security, national security, defense security, technological innovation and a series of national development strategies with important spillover effects, the

market cannot solve the problem solely [4], there are special needs for the state to promulgate a series of policies to improve the industry protection and support, regulating the development of satellite navigation systems industry.

73.2.2 To Realize the Necessary to Improve Industry Policy of the Beidou Satellite Navigation System in the Opinion of Comparative Advantage Theory

In the opinion of Comparative advantage theory, the comparative advantage of a country's economy always change obeying certain rules, from the early scarcity of capital to the scarcity of land, to the relative scarcity of skilled labor while capital becomes relatively abundant and cheap, and so on. In era of globalization, the competitiveness of high-tech is the main competition, capital and technology-intensive industries have become comparative advantage industry. According to this theory, the development of comparative advantage industry cannot fully rely on the market, and thus it is necessary to develop appropriate government industrial policies to cultivate awareness, such as giving a tilt encourage in the present state-owned industries with a comparative advantage to promote faster development [5].

The Beidou navigation system industry is a capital and technology-intensive industries and needs to invest a lot of money and talent in the pre-construction, relying solely on private capital is unable, in addition, the capital return period of Beidou industry is quite long, not able to obtain a satisfactory income in short period. These two aspects determine the industry's development is inseparable from the necessary support of the government. In the contemporary globalized competitive market, industrial policy is the main form of government support for an industry. Therefore, the Beidou navigation system which has a comparative advantage in the industry needs tilt policy of the government for benign development.

73.2.3 To Realize the Necessary to Improve Industry Policy of the Beidou Satellite Navigation System in the Opinion of Industry Competition Theory

According the industry competition theory, the country should ensure that domestic producers have a special status in the domestic market, in order to make the country's industry obtain more production scale and technical advantages over competitor, thus making this scale advantage in the absence of protection transformed into lower marginal costs and higher market share. Its essence is to ask the government for their support and protect over important industries to achieve optimal allocation of resources [6].

Even though China's Beidou satellite navigation system has developed rapidly, but compared to the US GPS industry, there is a big gap whether in the development stage or in the market competitiveness. In 2011, US GPS industry scale has reached \$105.23 billion with an increase of 30.5 %. In China, though in recent years the average annual increase over 35 %, the scale of the industry is only about 100 billion Yuan. There is not a small gap particularly in product performance, quality and technology compared with the US GPS system. If the Beidou satellite navigation system is sold on the market directly without support of the governments, it is almost a fantasy to compete with the US GPS system with a high market share and a high degree of development. Therefore, it is necessary to adjust the market access mechanism to provide protection for its development, meantime a series of industrial policy measures should be adopted to support the application.

73.3 To Realize the Urgency to Perfect the Industry Policy of the Beidou Satellite Navigation System in Practice

Last few years a swift growth of navigation system application industry has been witnessed. Since the 11th Five-Year Plan, industry gather has been formed in Pearl River Delta, Beijing-Tianjin-Hebei Region, Central China and Yangtze Delta. To further promote its healthy development, the State Council released "National Program for Medium and Long-term Satellite Navigation Industry Development" in September 2013. The program made the overall long-term planning for the satellite navigation industry from the national level until 2020, which is an important milestone for Beidou industry policy. Whereas, compared with USA and Russia, our devotion is far from enough [7].

Take USA as an example, a series of regulations and policies have been released [8], including the laws from Congress, administrative order from president or defense department which cover the system construction, operation, maintenance, application and safeguard. The Congress establish the industry policy as a legislative body. Concrete resort includes: Title 10 of the U.S. Code, Section 2281, Global Positioning System, assigns the Secretary of Defense statutory authority to sustain and operate GPS for military and civil purposes; The statute directs the Secretary of Defense to provide civil GPS service on a continuous, worldwide basis, free of direct user fees; It directs the Secretary of Defense to coordinate with the Secretary of Transportation on GPS requirements and GPS augmentation systems, and to coordinate with the Secretary of Commerce and others to facilitate civil and commercial GPS uses [9].

The first policy concerning GPS, Presidential Decision Directive NSTC-6, was released in March 28, 1996. This directive establishes national policy for the management and use of the U.S. Global Positioning System and related U.S. Government augmentations [10]. In a series of orders since 1996, the Federal Communications Commission (FCC) has taken action to improve the quality and

reliability of 911 emergency services for wireless phone users, by adopting rules to govern the availability of basic 911 services and the implementation of enhanced 911 (E911) for wireless services, which pose as an important procedure to popularize GPS in civil application. The President authorized a new national policy on December 8, 2004 that establishes guidance and implementation actions for space-based positioning, navigation, and timing programs, augmentations, and activities for U.S. national and homeland security, civil, scientific, and commercial purposes. This policy provides guidance for development, acquisition, operation, sustainment, and modernization of the Global Positioning System. In February 2008, DoD Directive 4650.5 updates DoD PNT policy, procedures, and responsibilities, Expands and clarifies the authority and responsibility assigned to the Assistant Secretary of Defense for Networks and Information Integration/DoD Chief Information Officer, by designating that official as the OSD Principal Staff Assistant (PSA) for overseeing PNT policy, development, and coordination of the DoD position on all aspects of PNT for communication to external organizations. The above procedures lead to an enormous blossom in application, whose experience needs to be drawn by our country.

Compared with United States, Beidou navigation system is backward in technology, industry scale, and market share and industry policy [11]. Specific mainly in the following items: Firstly, the level of industry policy is relatively low. China has issued several supporting regulations, such as “Medium and Long-Term Planning for China Satellite Navigation Industry” by State Council and “12th Five Year Specific Plan for Navigation and Position Service Technology” by Ministry of Science and Technology. The laws and regulations concerning GPS in USA are normally issued by the Congress and the President respectively, rather than by an administration branch. Secondly, the policy is in lack of systematicness and coordination. Led by the United States industrial policy, navigation system can be a relatively complete engineering system. In terms of the subject formulating industrial policy, it has formed a complete system of policies from state laws and administrative regulations and then to department regulations; In terms of industrial policy content, it covers position navigation system applications from industry, industrial development and system operation standard, responsibilities of the relevant departments and division, detailed policies and regulations support each other to form a complete system of industrial policy. However, the policies of China’s Beidou navigation system just stay in the general planning and executive departments, not yet forming a relatively complete industrial policy system with the lack of corresponding support policies. Thirdly, content of industrial policy is not specific. Industrial policy requires being specific with strong operability. US satellite navigation industry policy are very detailed from the global positioning of GPS, the division of responsibilities of relevant departments, industry development and system operation specifications, specific support measures. The industrial planning of Beidou navigation system mainly made some provisions for major policy direction and principles of industrial development, in lack of details of the operation. In fact, it is not strong in operability and difficult to really play its due role. Fourthly, there is a lack of strong policy support tool for the Beidou satellite

navigation system industry. Despite industrial planning in the core content of industrial policy, the industry's capital investment (government procurement), tax incentives, investment and financing measures and achievements and other important content are also included, and specific provisions relating to the content need a certain degree of mandatory, otherwise it is difficult to play its due role in the policy. However, due to the existing industrial policy, Beidou navigation system is still in the planning stage, there is no policy measures to support the industrial development of specific provisions, which in fact the policy is far from enough to support the industry [12].

From previous analysis, we can see that the government plays a leading role in the exploitation and generalization of GPS industry drawing on the experience of western countries. With experiment over 20 years and expenditure of 30 billion dollars, it is unaccomplished depending on the market mechanism. In spite of the largest scale of navigation industry, USA owns the most systematic and thorough industry policy. It is the effective laws and regulations that promote the progress of the navigation system. China is falling far behind USA and other western countries, and it demands high attention to implement proper industry policies. However, it is the high-tech industry that determines the government and market should cooperate to achieve a leapfrog development, just as stated in the 3rd Plenary Session of 18th CPC Central Committee, the government should take more responsibility. In this opinion, to improve the industry policy of Beidou Navigation system application is not only necessary in theory, but also is urgency in practice.

In June 2013, some classified documents were disclosed to several media outlets by Edward Snowden, a former employer of Central Intelligence Agency. The PRISM was first revealed and provided insight that USA government gather information from Microsoft, Google, Yahoo, Facebook, Pal Talk, AOL, Skype YouTube and Apple. The astonishing incident indicates that there is a white war around us—the war of information [12]. Satellite navigation system has a deep concern with national information security, so it is urgent to take measures to improve industry policy [13].

73.4 Principles to Improve the Industry Policy of Beidou Satellite Navigation System

There are some basic rules to obey in formulating and improving policies, and there is no exception for satellite navigation system. Deep study of the basic principles of Beidou satellite navigation system in the process of designing industrial policy, is of great significance for effectively and correctly improving the relevant industrial policies in this area. Under current conditions, based on the spirit of developing China's high-tech industries and defense industry, with reference to foreign satellite navigation system industrial policy, improving China's industrial policy should follow these basic principles:

1. Principle of government leading. Industrial policy action is itself a government action in the development and improvement of the industrial policy, we must adhere to the principle of government leading. If the development of an industry is left by laissez-faire market regulation, it means no industrial policy. As mentioned earlier, our system Beidou satellite application industry is not only a strategic high-tech industry, but also closely related to information security, national security and national security. Faced with fierce global competition, we have to use the government's power to support the development of this industry and must establish such a concept, otherwise we'll timid in the development and improvement of the Beidou system failing to introduce effective measures. In the decision of 3rd Plenary Session of 18th CPC Central Committee on deepening the reform, one is that "market mechanisms must play the decisive role in the allocation of resources", second is "the government should play a better role." To implement industrial policy led by the Government is a manifestation of "the government should play a better role." During the current process of perfecting Beidou satellite navigation system, we should fully implement the "CPC Central Committee on deepening reform of several major issues," firmly grasp the strategic development opportunity, follow the law of development of market economy, the law of development of science and technology, accelerate the improvement, improve the relevant laws and regulations as the main policy to promote the Beidou satellite navigation system as a reliable institutional mechanisms [14].
2. Principle of adapting to market economy system. Contemporary industrial policy is to produce and develop under the market economy conditions, exist as a complementary mechanism of market mechanisms. Without the market mechanism, industrial policy itself would be meaningless [4]. The report "CPC Central Committee decision on deepening reform of several major issues," adopted at the Third Plenary Session of the eighth, made clear that the dominant role of the market in resource allocation. it is necessary to encourage moderate competition in developing Beidou satellite navigation system. In the course of these areas and moderate competition, we must adhere to the principle that the market plays a decisive role in allocating resources. The function of industrial policy is to regulate the market to ensure a good and legal environment. Improving the industrial policy must be adapted to the market economy system. The decisive role of market adjustment is crucial to the industrial development of the Beidou satellite navigation system [12] with a healthy environment of laws and policies [15].
3. Principle of thinking globally based on national conditions. Satellite navigation system itself is a globalized industry or global industry. Because the industry is closely related to information security, defense and national security, competition in this industry is fierce, especially between developed and emerging countries. Therefore, in improving industrial policy, we must have a global perspective, review the world and consider the development of relevant policy options from the perspective of international competition. With particular emphasis on the experience from Western developed countries and regions

(such as the US, Russia and the EU), some policy principles and spirit in the application of industrial development of satellite navigation systems can be absorbed as policy content for our own use. Meanwhile, China has its own national conditions, China's satellite navigation system application industry is backward, and faced with fierce competition from foreign advanced navigation system industry from the birth, which determines the application of China's satellite navigation system not only should draw on advanced foreign policy content, but also should have their own distinctive measures [16].

4. Principle of Systematic design and overall advancement. The policies of the Satellite navigation system should be an organic integrity considering the systematic, holistic and synergistic features when designing. To this end, we must strengthen top-level design and establish an effective and cooperative mechanism. Beidou satellite navigation system is one involved in various fields and multiple subjects, individual policy and system cannot meet the practical requirements of scientific advance, we must strengthen macro-level design and plan policies and regulations at the national level, to establish healthy running policies and regulations project, feasibility studies, drafting, review coordination and cooperation mechanisms.

The reality that Beidou satellite navigation system is systemic requires comprehensive policies and regulations must be the foundation and backbone, therefore, a comprehensive legal department should become the leader in the development of policies and regulations, the relevant functional agencies should give full play to promote it and the role of research institutions and experts is to provide timely and accurate advice to ensure that the policy promote the industry running scientifically. According to the systemic features of industry policy, in the development and refinement of the policy, several backbone laws should act as the foundation to promote policies improvement. Beidou applications could only own a broad prospect with expansion worldwide, in line with the laws and regulations of industrial development are the premise for domestic competition and entry into the international market.

Currently, there are more relevant policy and less promulgated laws and regulations, as a result, administrative means surpass the rule of law. To change this situation as soon as possible, we should make full use of open and transparency of laws and regulations, especially characteristics of high credibility and entry to the judicial process, to strengthen high-order legislation, speed up the development of laws and regulations to promote scientific and technological achievements through modifying transformation, satellite navigation and other regulations, while focusing on participation in international standards, the rule-making. On this basis, combined with more flexible concrete policies to guide, the Beidou navigation satellite system will continue to make new breakthroughs despite the stable development.

5. Principle of highlighting the focus. Formulation, improvement and selection of industrial policy must highlight key issues and focus on the healthy development to get rid of the constraints of the deep-seated contradictions and obstacles. The current process of improving Beidou satellite navigation system should take

full advantage of a new opportunity for deepening reform, to solve the fragmentation of responsibilities and internal ill-defined function problems. Several steps should be taken to establish efficient, clear and smooth management mechanism including rationalizing the system mechanisms and adjusting the policies and regulations. The target is to ensure that the industry in the new starting soon occupy the domestic market share and remain the strong competition in the international market. The key issues to setup the relevant regulations and policies include: policies on infrastructure development and regulations, policies on promoting technology, policies on intellectual property protection, policies on promoting the popularization, policies on ensuring system security, policies on encouraging competition in overseas cooperation. The following objectives strive to be achieved before 2020: to establish a system of laws with the features such as international cooperation and competition, level of clarity, scientific structure, cohesion coordination, open and compatible [17].

6. Principle of paying attention to the research and feasibility studies. The development, improvement and choice of industrial policy are very complex engineering system. To prevent failures of policy options, its scheme, contents, the timing of issue should be given full demonstration. Various administrations at all levels should fully grasp the opportunity to strengthen the study of theory and practice application. Drawing on our experiences and the development of related industries in other countries, learning from past lessons, paying special attention to the study of international and domestic market, concerning international GNSS industry developments and trends, to ensure that research results in real-time interact with industry decision-making and the latest research results immediately influence decisions and support the decision-making. Departments should emphasize the “outer-brain” function to effectively integrate research resources, including strengthening basic research, practical exploration, and introducing practical, effective and feasible policies.

References

1. Research group of development strategy (2007) Demonstration platform for satellite navigation application. Research report on the policies of GNSS (in Chinese)
2. Rodrik D (2004) Industrial policy for the twenty-first century. Harvard University, Cambridge
3. Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2007) GNSS—global navigation satellite systems: GPS, GLONASS, Galileo, and more. Springer, Berlin
4. Zhang D, Zhu R, Zhao C et al (2013) Build national satellite navigation system to promote the rapid development of Beidou satellite navigation industrial policy. In: Proceeding of 4th China’s satellite navigation academic annual conference (in Chinese)
5. Suri T (2011) Selection and comparative advantage in technology adoption. *Econometrica* 79 (1):159–209
6. Xu G (2011) Study on China’s high-tech industrial competitiveness. University of Jilin, Changchun (in Chinese)

7. Chen B, Li L, Zhang X et al (2014) Fast health scale—interpretation of national satellite navigation industry and long-term development plan. *Satell Appl* 1:016 (in Chinese)
8. Hessin RM (2009) U.S. space-based positioning, navigation and timing policy and program update. In: 4th international committee on GNSS
9. Zhu Y, Li X, Yang Y (2011) Considerations for accelerating compass construction from the viewpoint of international GNSS development. *Bull of Surv Mapp* 8:002 (in Chinese)
10. Global Positioning System(GPS) (2008) DoD's Report to Congress
11. Chen B (2013) Opportunity and change for Beidou Industry—review on 4th China's satellite navigation academic annual conference. *Satell Appl* 005:40–43 (in Chinese)
12. Zhang X (2006) Consideration on strategic focus of China's space industry. *Satell Netw* 9: 30–39 (in Chinese)
13. Yuan S (2014) Legal first Promote Beidou industry to take off healthily. *Satell Netw* 1: I0028–I0028 (in Chinese)
14. Cao C (2012) Promoting the emerging industry of Beidou with multi-measures and practical work. *Digit Commun World* 6:8–9 (in Chinese)
15. Zhang H (2012) The EU and US Legislations on GNSS and Their Inspirations for China's BeiDou Navigation Satellite System. *J Beijing Inst Technol* 4(Social Sciences Edition):019 (in Chinese)
16. Wu H, Xia D, He Z (2012) Thought on the application of Beidou. *Space Int* 4:32–37 (in Chinese)
17. Sun J (2010) Development road of Beidou navigation system. *Space Explor* 7:31 (in Chinese)